Trying to see and say where I am, have been and would want to go.

Hans Kamp

May 18, 2020

Contents

1 How I saw the world of Logic, Language and Cognition up to roughly 1980

My first work that I feel would have deserved publication (although it never was published; entirely through my own fault) is what I did in my 1968 dissertation. The dissertation was on a topic that arose from the course on Tense Logic that Arthur Prior offered at UCLA in the fall of 1965, my first semester there as a Ph.D student. The topic was squarely within the framework of Priorean Tense Logic and its main result can be stated quite simply: While Prior's most familiar Tense Logic only has the tense operators P ('it was the case at some past time or other that') and F ('it will be the case at some future time or other that'), other tense operators, which one might consider adding to the (P,F) calculus, came up as well in Prior's class, and for each such operator O a natural question is whether or not it could be defined in terms of P and F (in the sense of 'O(q_1,..,q_n)', where n is the arity of O, being equivalent to some formula from the (P,F) calculus in the propositional variables q_1,..,q_n). If yes, then the addition of O would be harmless but also ultimately useless. If not, then there would be a real point to its addition: the calculus would acquire greater expressive power and therefore might be able to deal with logical and linguistic phenomena that the (P,F) calculus could not. As it turned out – this is the main result from the dissertation – there is a pair of binary operators, S(INCE) ('it has been the case uninterruptedly that q since some time when it was the case that p') and U(NTIL) ('it will be the case uninterruptedly that q until some time when it will be the case that p'), that are 'complete' in the sense that there is a large and natural class of tense operators O each one of which is definable in terms of S and U; the class is that of all operators that can be semantically defined with standard means (the relation $\prec$ of temporal order together with variables for the propositions that the operator operates on).[1]

Connected with this result, and the initial motivation that led to it, was that S and U make it possible to introduce some kind of metric into Prior type tense logics, in spite of the fact that the temporal structure that is directly captured by these tense logics is a purely topological one. (Only the order of time is relevant to the semantics and logic of these logics; they have no

---

[1]The result holds only for 'continuous' orderings, among them the integers and the real numbers, but not for e.g. the rationals. However, as was shown by Jonathan Stavi, it can be improved by adding a further pair of binary operators. In this further extended calculus a formula can be given for each operator O in the semantically defined class which correctly characterizes O(q_1,..,q_n) for all linear orderings, including those that are not continuous.

automatic way to express anything about the size of temporal intervals, such as, for instance, that the time during which a proposition q_1 was true was twice that during which some other proposition q_2 was true.) That gives the (S,U)-calculus a lot of extra power, of a kind that promised to be useful in the analysis of natural language, where so much of what we say has to do with how long things last or go on for. But the joy was not to be of long duration. Even during my time at UCLA, and thus while my dissertation was in the making, I had stumbled on another reason why the (P,F)-calculus wasn't good enough as a framework for dealing with the semantics and logic of natural language; and this reason was as much a problem for the (S,U)-calculus as it was for the (P,F)-calculus. This was the problem that is posed by 'now' when it too is treated as a tense operator. One feature of 'now' is that often it gets us back to the utterance time even when it occurs embedded within the scope of one or more other tense operators. This 'reaching back to the utterance time' is another phenomenon that Priorean tense logics like the (P,F)-calculus and the (S,U)-calculus are unable to deal with. Needed for this are logics whose semantics is given by 'double-indexed' truth definitions: a formula is specified as true relative to pairs of times <t,t'>, where t is the current evaluation time and t' the utterance time. In such a semantics 'Now q' is true when evaluated at t of the pair <t,t'> iff q is true at the utterance time t'.[2]

My dissatisfaction with Tense Logic as a framework that could be integrated into some more comprehensive semantics framework was cumulative. It is possible to express in the tense logics that were available by the beginnings of the seventies – e.g. an (S,U)-calculus with double indexing – a substantial part of the repertoire of things that can be said in English and other natural languages. But to do this one has to use tense operators in ways that have little if anything to do with the natural language tenses of which the tense operators were meant to be the formal counterparts. As a consequence, the 'logical forms' that result when the tense operators are used in these ways often bear no formal resemblance to the natural language constructions those

---

[2]It should be noted here that Montague never advocated the operator treatment of 'now' and had himself treated it as an indexical individual constant. The same is true of Kaplan, for whom 'now' is one of the paradigm indexicals, and for whom it also has the status of an individual constant. (In fact, some of the things that Kaplan says about indexicals, in *Demonstratives* and elsewhere, requires non-trivial reformulation when 'now' is to be treated syntactically as an operator instead.) I myself also abandoned the idea that 'now' is an operator not long after I had worked out the details of the 2-dimensional sermantics that such a treatment seemed to require. In my case the return to treating 'now' as an individual constant was part of giving up Tense Logic as a frame for natural language semantics altogether. More about this below.)

forms represent. As one would put it nowadays: the tense-logical framework offers no framework for a plausible syntax-semantics interface for those constructions. And while it is true that using the (S,U)-calculus instead of the (P,F)-calculus creates room for stating the truth conditions of a much larger array of natural language constructions, the switch from (P,F)-calculus to (S,U)-calculus makes the interface problem a good deal worse.

Further explorations of tenses and other devices of temporal reference in natural languages made Prior type Tense Logic appear ever more unsuitable for the study of natural language semantics and logic. One important factor was my realization how common it is for 'now' and its counterparts in other languages than English to refer not to the utterance time (as assumed in the work mentioned above), but to some past 'reference point' or 'perspective point'. It couldn't be excluded off hand that some new clever form of multi-indexing could deal with such cases. But playing around with ever more complex variations in this domain no longer felt like it was going to do more than plug one leak until the next one would be discovered.[3] Work on other aspects of semantics and logic – adjective semantics and vagueness, logic-based translation between formal and natural languages, non-assertoric speech acts like permission-granting utterances and commands – reinforced this skepticism. But a crucial experience was the problem I was asked to have a look at in the summer of 1978 by my later colleague Christian Rohrer, who at that time was professor for Romance Linguistics at the University of Stuttgart. The problem is one about the two French past tenses Passé Simple and Imparfait. These tenses differ in ways that present no difficulties to native speakers of French, but are not easy to explain to those whose native language does not have a comparable distinction. Among these are in particular German students who aspire to become French teachers at German high schools, where they must be able to convey this difference to (their high) school students. And what you need to explain this difference to youngsters who learn French as a second language, is something like a proper theory of what the difference is.

Not surprisingly there had been a number of attempts by German Romance philologists (and likewise by Scandinavian and Dutch philologists, who saw themselves confronted with the same problem) to identify and clearly describe the differences between Passé Simple and Imparfait. One of these differences

---

[3]Playing around with multiple indexing had become something of a little logic industry by the mid seventies. But I have never been able to see more in this, than an exercise pursued for its own technical interest.

was that between 'punctuality' and 'durativity': the Passé Simple, it was observed, was used to talk about 'punctual' events and the Imparfait about temporally extended, 'durative' events or states. Rohrer thought that a more precise articulation of punctuality and durativity ought to be possible within the model-theoretic framework for studying natural language semantics that had been initiated in the work of Montague and that was becoming a new paradigm for doing natural language semantics; and since I had been working on the semantics of time and tense within this framework, he decided to draw my attention to the Passé Simple-Imparfait problem.

The idea was that the models used in the model-theoretic semantics of tensed discourse can be thought of as abstract characterizations of possible worlds that extend through time and that contain the events and states that the sentences talk about. Punctuality and durativity should be given mathematically precise definitions as properties of the events and states that Passé Simple and Imparfait sentences are used to talk about. It became soon clear to me, however, that a distinction between properties of the events and states in the models couldn't capture the distinction that Romance philologists were after. Instead, the notions of punctuality and durativity that one needed to explain the difference between Passé Simple and Imparfait had to be understood as operative at a conceptual level – that at which the sentence or discourse unfolds its own story, with its own temporal progression and its own temporal granularity – a granularity typically much coarser than that of the time structure of any model that can be regarded as a realistic model of a possible world through time. Because of this difference in granularity, an event that had a non-punctual temporal extension in the model representing the actual world could nevertheless appear punctual at the level of discourse time. And in that case it could be described by a Passé Simple sentence in spite of its actual durativity.

Two points emerged from these considerations that proved to be of decisive momentum for my future work. The first was that this is what a more precise theory of semantic interpretation needed:

- To provide an explicit account of how the interpretation of discourses consisting of sentences with Passé Simple, Imparfait and other tenses leads to discourse structures that correctly represent the content of those discourses and that define the discourse time structures within which the events described by Passé Simple clauses are punctual and the events described by Imparfait clauses are not.

6

But in the context of this first, quite general desideratum, a second point proved to be crucial as well. Among the interpretation rules that have to be spelled out in an algorithm for turning French sentence sequences into discourse structures satisfying the desiderata above are those for the tenses, and in particular those for Passé Simple and Imparfait. But it turns out that these principles – which capture so to speak the meaning of these tense forms, as the contributions they make to the discourses within which they occur – are typically *sentence-transcendent*:

- The contributions these and other tenses make to the discourse content are not restricted to the sentences in which they occur but serve, partly or wholly, to connect their sentences with the preceding discourse.

Both these points were novelties for me. And especially the second gave me headaches. That it could be a central feature of a type of sentence constituent to connect its sentence with other parts of the surrounding discourse wasn't among the possibilities in my book at that time (and I think not in the books of many of my colleagues either). Since its origins with Montague's work, formal semantics had been understood as a matter of accounting for the meanings of single sentences. When sentences are combined into a multi-sentence discourse, the meaning of the whole would arise out of the meanings of those individual sentences. That might involve various pragmatic principles, but the constituents of the different sentences would at that point have done their work, making their contributions to the sentence contents that needed to be combined into the meaning of the discourse, and no further consideration of those individual constituents should be needed at this level of discourse integration. That the semantic effect of a sentence constituent should reach beyond the horizon of its own sentence felt like heresy.

Such scruples may sound strange to someone doing linguistics in 2020. But at the time they made me feel that the idea of trans-sentential semantic contributions made by sub-sentential constituents just couldn't be right – that there must be something beyond the point to which my thoughts had led me which was still eluding me. It was only when about a year and half later I started to think – I cannot now remember exactly why – about Geach's donkey sentences that it dawned on me that anaphoric pronouns were in this respect like tenses: They could not only find their antecedents within the sentences of which they themselves were part, but also in earlier sentences. Their semantics too, it seemed, could reach beyond the sentence boundary.

A further point emerged at this point. Anaphoric antecedents of pronouns often are indefinite noun phrases, and that is possible both when indefinite antecedent and pronoun occur in the same sentence, as in Geach's donkey sentences, but also when they occur in different sentences (so-called 'donkey discourses'). And donkey discourses appear to have a property for which there is no obvious room within a sentence-based semantics: the sentence with the pronoun and the sentence with the antecedent are *contentwise connected*, by the anaphoric link between the pronoun and its indefinite antecedent, in a manner that renders them 'logically inseparable': the proposition expressed by the two sentences together takes the form of an existentially quantified conjunction, which as a rule cannot be 'factored' as the conjunction of two logically independent propositions expressed by the respective sentences individually.

These discoveries about donkey pronouns gave me heart. We now had at least two kinds of expressions – tenses and pronouns – that were able to make semantic contributions by reaching beyond the bounds of their own sentences and whose semantic properties could not be properly described unless their trans-sentential contributions were explicitly brought into the theory. Furthermore, the story about donkey pronouns seemed a good way to put my case for the need of a trans-sentential natural language semantics before the formal semantics forum. A number of semanticists were thinking hard about donkey sentences at that time. The formal semantics of tense and aspect was a more recondite part of the formal semantics landscape, without some like an established, widely shared sentence-only formal semantics against which the trans-sentential claims about tenses could be pitched.

However, if a plea for a discourse semantics for anaphoric pronouns was going to work, then it would of course not be enough to restrict attention to 'donkey discourses'. The account should apply to donkey sentences as well, and in such a way that the two types of case, the donkey sentences and the donkey discourses, would be accounted for in essentially the same way. It was the attempt at such an account that led to DRT as it is probably familiar to most of those to whom it is familiar in any form: An account of how to build semantic representations for a small fragment of English with pronouns, indefinite and universally quantifying noun phrases, conditionals, relative clauses and a handful of nouns and transitive and intransitive verbs. It is the first formally worked out version of DRT in which there are DRSs (DRT's discourse representations) with complex DRS conditions, which enable these representations to capture the content of universal quantifications and conditionals and in which an explicitly stated DRS construction algo-

8

rithm constructs such DRSs from syntactic structures of sentences from this fragment.

The paper in which this account was written up (*A Theory of Truth and Semantic Representation*) does not say anything about temporal reference or the perceived analogy between pronouns and tenses, especially with regard to their sentence-transcendent capacities. Another paper, which appeared in the same year (*Événements, réprésentations discursives et référence temporelle*, 1981) is devoted to the Passé Simple and the Imparfait. Here too a construction algorithm is spelled out, for a small fragment of French, which contains these tenses as well as French's third tense form that often behaves like a simple past tense, the Passé Composé. But this fragment does not contain anything that gives rise to DRSs with complex DRS conditions. The account of donkey anaphora in *A Theory of Truth and Semantic Representation* is mentioned in a footnote, but not elaborated on in the paper.

In fact, there is one respect in which the analogy between tenses and pronouns that was such a central motivation behind the development of DRT was not properly worked out in publicly accessible work at that time. The common aspects of the sentence-transcendent roles of tenses and pronouns were visible clearly enough, I think, from the two papers; and so should have been the connection between pronouns with sentence-transcendent and pronouns with sentence-internal antecedent. What was missing, and has been missing to this day at least as far as my own work has been concerned, is an explicit DRT-based treatment of tenses whose interpretation is sentence-internal, such as the tenses of finite complements of attitudinal verbs and *verba dicendi*, or those occurring in relative clauses and in temporal subordinate clauses.[4]

The most important methodological implications that I associated with this work at the time – and this is still very much the way in which I see these things today – are (i) the conclusion that semantics cannot be restricted to the level of the individual sentence: The same kind of compositional principles that are responsible for the meaning of single sentences (in those cases where sentences have self-contained meanings on their own) are equally involved in the meaning construction for multi-sentence discourses. And (ii) it is possible to structure discourse representations in such a way that they

---

[4]A treatment of such sentence-internally interpretable tenses is planned as final part of Chapter 4 of the UT Semantic Notes, a documentation of bottom up DRS construction for a wide and growing natural language repertoire.

can serve the double purpose of (a) correctly representing the discourse parts from which they have been constructed and (b) serving as 'discourse contexts in the interpretation of the next sentence, a general principle to which I have sometimes referred to as 'Content equals Context': The DRS constructed from the sentence or part of the discourse preceding a given sentence S correctly captures the truth conditions of the sentence of discourse from which it has been constructed (according to the model-theoretic semantics for the DRS language) and at the same time its representational structure enables it to serve as discourse context in the DRS construction for S.

1.1 Tasks and Nature of Logic

I conclude this section with a few remarks about the views about logic that I had when I came to UCLA and about how these changed, starting with the work for my dissertation. In the final section, Section 7, I will return to this topic, in the light of the various things that I will bring up in Sections 2 - 6, leading up to my momentarily final personal assessment of the place and roles of logic today. The remarks that follow right here are about the first decade of what looks to me now as a very gradual shift extending over most of my career.

When I came to UCLA in 1965 I thought of myself as a logician. My dissertation project was one situated squarely within mathematical logic as the discipline was generally understood in those days: try to discover new decidable subclasses of first order predicate logic, by using the method of Semantic Tableaux. (At that time the Tableaux Method was still quite new, but I had been brought up with it, so to speak, as a masters student in the Institute for Logic and the Foundations of the natural Sciences that had been founded only a few years before by E.W. Beth, one of the two inventors of the Semantic Tableaux; at Beth's Institute the Tableaux Method already had the status of a kind of canon, and was used to teach the basics of mathematical logic to students who had never seen any logic before.) I never got very far with that project, since already in my first semester at UCLA Prior's course on Tense Logic deflected my interests towards what then became the topic of my dissertation. However, I never really gave up on the original topic; but as time has gone by my doubts have grown that anything particularly useful would have come out of it. The Tableau method has proved of great value and is that to this day, but in application to new logics rather than to classical logic in its classical, predicate logic form.

The switch from a topic squarely within mathematical logic to Tense Logic

bore within it the germs of a different view of what formal logic was – what it was about and what it was for. One important change of vision, which grew stronger and more conscious in the course of the following years, was one that went from a more absolute to a more relative and instrumental view of logic. I had come to LA with a view of formal logic according to which there was one true notion of logical validity, even if it was (still?) a matter of ongoing debate which notion that was – Classical, Intuitionistic (or Minimal in the sense of Johannson's Minimal Logic, as an outside candidate). This was the central question about the foundations of logic, and where the deep convictions of different members of the logic community showed themselves for what they were. But it was a true debate – in the long run only one of the different sides could be right.

My intercourse with Tense Logic overturned this rigorous perspective on formal logic, as aiming for the absolute: settle once and for all what are *the true* notions of valid logical inference and logical truth, and then try to find out as much about these notions as possible, a project to which each of the four main branches of mathematical logic, Model Theory, Proof Theory, Recursion Theory and Set Theory, made its own contribution. The first dent in this picture that Tense Logic produced in me had to do with the fact that what formulas of Priorean Tense Logics come out as logically valid depends on assumptions about the structure of time, for instance on whether time is to be taken as branching or linear, as discrete or dense, or whether or not it is assumed to be continuous. How are we to choose between these various options? Tense Logicians at the time didn't know; and since they didn't know, they kept all options open, by investigating how validity varies with these different assumptions about the structure of time, leaving the decision what the structure of time 'really' was to someone else. But who should be making those decisions, and on what basis? Since I could not see any principled way of resolving these questions, Tense Logic became for me the first example of a branch of formal logic which I felt I knew well and where it was very unclear that there was just one notion of logical validity. Perhaps it was right to see the pool of alternative Tense Logics not as a preliminary towards the one true tense logic, but rather as a set of alternatives from which different 'logic customers' – people who needed a tense logic within the wider context of some scientific or philosophical project – could pick and choose according to their individual needs. The logician's task ended with creating the pool.

This more instrumental view of the function of formal logic, as making available various formally precise logical systems that can then be employed in various contexts, and where different systems may be more suitable for differ-

ent purposes, is very different from the strongly absolutist perspective which was mine when I left Amsterdam. The relativistic perspective has grown ever stronger with the years, to a large extent because so much of the work within formal logic during the last half century has been concerned with the logical analysis of quite specific problems or domains of scientific investigation. This is not, I should add right away, to say that the absolutist perspective has lost its justification; the instrumental and the absolutist, I believe, both have their legitimate position. But more about this later.

2   DRT and Dynamic Semantics

DRT has been with me since its first formulation in 1980/81, and has been an active preoccupation for almost all this time. So it has, to some extent, also been a filter through which I have seen the world of language and logic at large, and quite possibly to a much greater extent than I am aware of or would be happy to admit. That makes it the most appropriate point of departure for these reflections on where I stand today and how I got there.

DRT has gone through many changes over the years. It has been used as the basis for other formal approaches to the meaning and logic of natural language, quite a few of which bear witness to their connection with DRT in their names. It also met, more or less from the start, with a good deal of resistance, and some of the revisions that have been proposed have been designed to deal with some of those objections. As a first step towards organizing the following discussion of objections, revisions and other extensions of DRT let us recall the basic structure of the DRT approach towards natural language semantics.

The core of any particular version of DRT is its *DRS language.* DRS languages are formal languages designed in the spirit of what are nowadays the standard presentations of Predicate Logic and the Typed Lambda Calculus. That is, the specification of a DRS language L_{DRS} involves:

(C_{DRT} 1) a recursive definition of the well-formed expressions of the language ('C' here stands for 'component'), and

(C_{DRT} 2) a model theory for the language, consisting of a specification of (a) the class of models and (b) a definition of truth or satisfaction that connects each well-formed expression with each model, by specifying the semantic value of the expression in the model (or in some part of the model

when the model is, for instance, an *intensional model*, consisting of a family of extensional models, each of which represents a different possible world).

In addition, the use of L_{DRS} in the semantic treatment of some natural language fragment L involves:

(C_{DRT} 3) an algorithm for converting sentences and 'discourses' (sequences of sentences) of L into DRSs from L_{DRS}. (These algorithms are usual referred to as 'DRS construction algorithms'. They always presuppose a formal syntactic specification of the fragment L that is the target of the analysis. I will take this for granted in what follows and not refer to it again.)

In the working life of the linguist who uses DRT as a tool of semantic analysis it is C_{DRT} 3 that takes up most of one's time and energy. It is here that the actual linguistic analyses are found – in the form of the syntax-semantics interface principles that characterize the semantic contributions of the different syntactic constituents of sentences of L, and of the discourses that can be formed out of those sentences. But the other two components of the DRT approach are no less important, even if those tend to be less of a daily preoccupation for the linguist: While it is true that different DRT languages, with their respective syntax and model theories, will be chosen for different applications, and that often changes will be needed when a given DRS language is used in application to a new fragment or specific semantic phenomenon, there is in practice much less variation here, and the changes, if necessary at all, are often local and minor. More often than not, when the semanticist wants to give an account of a natural language fragment or of some particular linguistic construction, he has to choose a suitable DRS-language just once. And that is then where all the hard work begins. But even if this is what absorbs most of the energies of the working linguist, I repeat nevertheless: the forms of the DRSs of the L_{DRS} that is used in a given application are of crucial importance to what those DRSs can tell us about the expressions from which they are constructed; and so is the model theory for L_{DRS}, both the definition of the model class and the semantic value definition. For it is also in virtue of the model theory that the expressions of the given fragment are assigned their 'intentions'[5], the functions that determine their semantic values for each possible world or model.

---

[5]I am using the word 'intention' here in more or less the sense that it has with Brentano: that which links linguistic expressions to their denotations in the real world and its possible alternatives.

Important for our discussions is going to be this: As I have described the DRT approach, as involving the three components C_{DRT} 1 - C_{DRT} 3, it is independent of any position one might want to adopt with regard to what it 'means' for an application of the approach to some NL fragment L to assign (via its construction algorithm) a DRS to some sentence S or discourse D from L. As the last section should have made clear, the original motive for developing DRT was the conviction that discourse interpretation takes the form of computing semantic representations for discourses incrementally, with the DRS K computed on the basis of each initial segment <S1,...,Sn> specifying the correct truth conditions for that segment (as determined by the model-theoretic semantics for the DRS language L_{DRS} that is being used when it is applied to K), while on the other hand the form of K often plays an essential part in constructing the incrementation K' of K that represents <S1,...,Sn,Sn+1>. If the approach can be made to work (in the sense that it delivers the intuitively correct truth conditions for discourses and their initial segments for some non-trivial fragment L), then this success can be seen as some evidence that the understanding of such discourses by speakers of L involves the formation of content representations that have some of the structural properties of the DRSs that the theory assigns to discourses of L.

2.1 Objections to DRT

How plausible are such inferences from truth-conditional success to the cognitive relevance of the form of DRSs? That is difficult to say, since a proper methodology for evaluating such claims has been missing.[6] And it should not be surprising that some people have been of the opinion that the basis for inferences to the mental relevance of DRS structure is far too slim. But the point to be made at this point that whether one accepts or rejects such inferences is not by itself an argument for or against the DRT approach as legitimate way of doing natural language semantics. So long as it helps to get truth-conditionally satisfactory treatments of significant parts of natural language right, and especially when it succeeds in places where alternative methods have failed, then prima facie that should be a stroke in favor of the method, irrespective of possible implications for mental representation.

This doesn't mean that there couldn't be other objections against DRT, irrespective of its success in getting the correct truth conditions. But these

---

[6]The work by Brasoveanu and Dotlacil that I will have something more to say about later on suggests a way out of this unsatisfactory methodological predicament.

should be sharply distinguished from the 'mental relevance' issue. Our next task is to discuss two main types of such further objections, which have been prominent more or less since the time when DRT first saw the light. The first type of objection is against the 'representational' role that is played by DRT's DRSs (irrespective of the question whether their representational form has any psychological relevance). And the second objection is directed against its claims that what we need is a discourse semantics, rather than a semantics restricted to single sentences; and with that comes the objection to the incremental dimension of DRT, in which the semantics of the preceding part of a discourse is an input to the semantics of the next sentence. Supporters of the second objection have for the most part also rejected the representational aspects of DRT. Their position was essentially that there is nothing wrong with the central assumptions of Montague Grammar: (i) that the semantics of natural languages can be given by a direct mapping from the syntactic structures of well-formed expressions of the natural language fragment L to their semantic values in models and (ii) that semantics can and should be limited to single sentences. I will have a fair amount to say about the first type of objection but little about the second.

1. The first type of objection is against the 'representational' role that is played by DRT's DRSs. Note well, this is now not be misunderstood as an objection against flirtations with mental relevance of DRSs. In practice, the objection that I want to address – the anti-mental-representation objection – may not always have been clearly distinguished by those who launched them from the objection against using any kind of representations or logical forms as intermediaries between natural language expressions and the models in which those must find their semantic values; and I think I may to some considerable degree have been at fault in this, through the rhetoric I adopted both in oral presentations of DRT and in writing. But the distinction can and should be made.

The objection that is at issue right now has to do with the role that DRSs play in the specification of DRS construction algorithms. The first such algorithms were particularly apt to provoke such qualms. They were 'Top Down' algorithms in which the syntactic structure of the sentence S to which a DRS is to be assigned (perhaps as the update of a DRS that has already been computed for the discourse segment preceding S) is traversed from the top (the S or CP node for the sentence as a whole) to the various leaves of the syntactic tree. It is only when the leaves have been reached and dealt with and the DRS (or DRS update) K has been completed that we have a structure to which the model theory for the given DRS language assigns well-

defined truth conditions. But no model-related semantic values are assigned to the intermediate outputs of such a construction algorithm, those structures that are reached when the algorithm has arrived at some non-minimal constituents of S (nodes of the tree that are between the S-node and its terminal nodes). There was a reason for setting up the construction procedure in this Top Down fashion, having to do with the anaphoric possibilities for pronouns that the first explicit version of DRT (that of *A Theory of Truth and Semantic Representation*) was particularly concerned with. But in the long run that proved to be not a good reason, and the very phenomena that seemed to favor the Top Down approach initially are now dealt with in a way that has no need of Top Down construction and where in fact Top Down construction is more of a hindrance than a help. (More about this in Section 3 on Presupposition.) In retrospect I regret having stuck it out with Top Down DRS construction for as long as I have. Already in the second half of the 1980s it was proposed (by Asher, see his (Asher 1993), and Zeevat (Zeevat 1989)) to switch to 'Bottom Up' construction algorithms, in which DRSs or DRS-like structures are built from the leaves to the top of the syntactic tree, following the same path as recursive definitions of the values of syntactic constituents in models familiar from Montague Grammar and also from the standard way of doing model theory for formal languages (including the DRS languages of DRT). The reason for carrying on with Top Down algorithms for as long as I did –*From Discourse to Logic*, which appeared in 1993 (and which is still the most prominent presentation of DRT today) makes use of Top Down algorithms throughout, in spite of the proposals by Asher and Zeevat that were several years old at that point and had been known to us (my coauthor Uwe Reyle and me) pretty much from the time they were made – was motivated by nothing better than inertia. *From Discourse to Logic* had been in making for many years and recasting everything it says about DRS construction, which occupies by far the largest part of the book and affects pretty much everything else, was simply infeasible given the deadline that we had set ourselves. In fact, in retrospect that was the only realistic option, just as it appeared to us at the time. Otherwise the book might never have come out.[7]

---

[7]It was only when presupposition was incorporated into DRT that I came to see Top Down DRS construction as something that definitely had to be given up. But in the DRT 'update' (Kamp, van Genabith & Reyle 2011), where presuppositions are explicitly introduced as components of DRS languages, the question of DRS construction is hardly discussed, let alone that a Bottom Up construction is as much as outlined there. ((Kamp et al. 2011) was written nearly ten years before it appeared, but even around the beginning of the present century it was clear to us that Top Down DRS construction had to be replaced by Bottom Up construction some time.) For more on Presupposition in DRT see

A problem with Top Down construction algorithms, we observed, is that only for the final product does the approach provide a systematic evaluation in terms of truth conditions. No such evaluations are offered for the structures obtained at intermediate stages. In this respect Top Down construction algorithms differ from the semantic value assignments defined in Montague Grammar, where each syntactic constituent of syntactically well-formed sentence is assigned a value in each model, and not just the sentence as a whole. When one has the view that what goes on in the mind of an interpreting speaker is in any case a computational process that transforms representations into other representations, then the lack of model-related evaluations of intermediate stages may not seem all that much of a drawback. But if such an intuitive justification of the DRT approach is to be set aside, as I suggested we do for the purpose of the present discussion, then a construction algorithm without semantic evaluation of parts or stages may seem much more suspect. (This also explains why it is not all that easy to keep the objection against drawing any conclusions about the cognitive relevance of DRSs from its truth-conditional adequacy separate from other objections against the representational character of DRT treatments that make use of Top Down algorithms. The objection we are discussing right now will be felt much more strongly by someone for whom the mental connection is a dead end from the outset.)

For DRT approaches that use Bottom Up construction algorithms the matter is different. In fact, they can be formulated in such a way that the intermediate representations, which the algorithm constructs for all the well-formed constituents of sentences, get assigned model-theoretic values too, and not just the sentence as a whole (or the DRS for the discourse of which it is the last sentence). But for such a Bottom Up form of DRT there remains a possible role that representation can take. Bottom Up algorithms may still rely on the representational forms they assign to the daughters of a syntactic mother-daughters configuration in the way they specify how these representations are combined into the representation of the mother node. (This is indeed the case for the Bottom Up algorithm referred to two paragraphs above.) And that then still constitutes a difference with the compositional operations of MG and of the syntax-semantics interface architecture of 'mainstream formal semantics', where the compositional operations are operations involving *semantic values* and not syntactic operations on terms that denote those values.

Section 3.

17

Let me elucidate this point in a little more formal detail. The main operation of Montague style formal semantics is functional application, of the function that is the semantic value of one of the two daughters of a given mother node to the semantic value of the other daughter. When instances of such applications are written down as part of the semantic value definition for some natural language fragment L this has to be done in some meta language which has to provide among other things terms to denote the various semantic values. The term formalisms that are standardly used for specifying these semantic values are versions of the Typed Lambda Calculus (such as for instance Montague's Higher Order Intensional Logic or the version that Muskens developed as part of his embedding of DRT within Type Theory[8]). Suppose that the term for the daughter whose value is the function has the form $\lambda x.\tau$ and the value for the daughter whose value is the argument has the form σ. Then the term denoting the value of the mother is denoted as $(\lambda x.\tau)(\sigma)$. Such complex terms are often hard to make sense of, especially when they are the result of several cases of functional application. But usually this problem – this is just a transparency problem for the developer and the users of the semantic theory – is mitigated by the possibility of simplifying complex application terms by subjecting them to reduction rules of the Lambda Calculus, and in particular to the rule of β *conversion*. All of us who teach formal semantics spend time on getting our students to know how to carry out β conversions, thereby turning nearly impenetrable terms into ones that can be made sense of much more easily.

One common but mildly deplorable effect of such exercises is that many students come away from them with some ill-articulated notion that β conversion is what most of compositional semantics is about. Of course it isn't. It is o.k. to use β *conversion* because it is the syntactic reflection of functional application: a term $(\lambda x.\tau)(\sigma)$ and the term that results from carrying out β *conversion* on it always denote the same semantic values. So when $\lambda x.\tau$ denotes the function f (a semantic value of the type of a function) and σ the entity d (also a semantic value), then the term resulting from applying β *conversion* to $(\lambda x.\tau)(\sigma)$ denotes the result of applying the function f to the argument d.

In fact, suppose the theory is formulated in such a way that the terms it uses to describe the semantic values of nodes that play the function part in functional applications that yield the value of the mother node are always

---

[8]For more on Muskens' proposals see Section 2.2 and Section 4.1

of the form $\lambda x.\tau$. Then we state the clause for functional application in the recursive definition of $[[\alpha]]_M$ (the semantic value of expression α in model M) in the following way. Suppose that the syntactic configuration $[[\alpha_1]_{D_1} [\alpha_2]_{D_2}]_M$, consisting of a mother node M with daughter nodes D_1 and D_2, is one of those that are interpreted as functional application and that D_1 contributes the function and D_2 the argument. Then the clause can be stated like this:

(1)   Suppose that the value of D_1 is (given as) $\lambda x.\tau$ and that of D_2 as σ.
      Then the value of M is given by the term $\tau[x/\sigma]$.

      ($\tau[x/\sigma]$ is the result of applying β reduction to $\lambda x.\tau(\sigma)$.)

The point of this reformulation is this. One could think of (1) as a 'representational' specification of how the semantics of the mother node is determined by the semantics of its two daughters, one in which the 'term representing' the semantic value of the mother node can be obtained from the representing terms of the daughters by subjecting them to a syntactic operation. But as we already know, there is no need to proceed in this way. We can also define the semantic value of the mother as the result of applying the semantic value of one of the daughters to the semantic value of the other daughter.[9] Going the first, syntactic, way is possible, but it isn't necessary. And because it isn't necessary it is harmless from the point of view of someone who believes in 'strict compositionality' as entailing that the recursive clauses of the semantic value definition can all be stated as operations on semantic values.

There is no reason, however, why all representational formulations of composition principles can be replaced by operations on the denoted semantic values in the way that (1) can be replaced by the statement that the semantic value of M is the result of applying the semantic value of D_1 to the semantic value of D_2. Replacement is not possible in a situation like the following:

(2) $[[R_{D1}]]_M = [[R'_{D1}]]_M$, $[[R_{D2}]]_M = [[R'_{D2}]]_M$, but $[[R_{Mo}]]_M \neq [[R'_{Mo}]]_M$.

In this situation there clearly cannot be an operator O that is defined on the semantic values of the representations and that yields the right semantic

_____

[9]Of course, this formulation must have a way of saying what the semantic value of the mother M is, and using a term from the Lambda Calculus is a convenient (and the familiar) way to do this: the semantic value of the mother node is $(\rho)(\sigma)$, where ρ is a term denoting the semantic value of D_1 (the 'function daughter') and σ a term denoting the semantic value of D_2 (the 'argument duaghter'). But that doesn't alter anything to the fact that here the value of M is given by an operation on the *semantic values* of the daughters.

value for both occurrences of M. For the operator gets the same argument values from D_1 and from D_2 for each of the two occurrences of the configuration, so if it returns the right value for M in the first configuration occurrence, then it will also return that value for the second occurrence, and that will then necessarily be the wrong value.

That the possibility of irreducible representational composition clauses isn't just a matter of idle speculation is suggested by examples like that in (3).

(3) a. One of the ten balls is not in the bag, and in fact it is under the sofa.

b. Nine of the ten balls are in the bag, and in fact it is under the sofa.

These two sentences differ semantically. The first is a coherent statement of there being one of ten balls missing from the bag in question, and its being under the sofa. The second cannot be interpreted in this way, because here it is impossible to interpret the pronoun *it* as referring to the missing ball. One aspect of the syntactic structure of each of these sentences is the formation of the conjunction of the first and the second clause via the sentence connective *, and in fact*.[10] Thus the syntactic structure of (3.a) can be given as in (4) (with a 'Curried' treatment of conjunction as first applying to the second conjunct and the result of this then applying to the first conjunct); and likewise for (3.b), except that the first conjunct is now a different sentence. Let us refer to this second syntactic structure as (4').

(4)

```
                        S
              _____/  _____
             /                    \
            S                    2ndConj
            |                   ____/  \____
One of the ten balls is not in the bag   \
                           Conj           S
                            |             |
                      , and in fact   it is under the sofa
```

_____
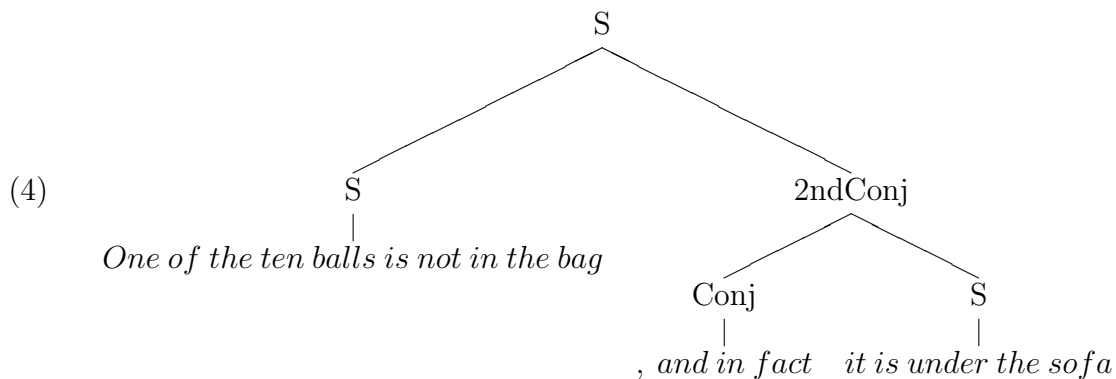
[10] A question may be raised about the semantics of *, and in fact*: Is this just the plain logical conjunction (the one typically represented as & or as $\wedge$)? Or is it something different? This question is irrelevant to the point at issue, so long as we assume that *, and in fact* makes the same semantic contribution to (3.a) and to (3.b). Even if this were a problem, it plays no part in the next example.

(4) is a syntactic configuration consisting of a mother node S and two daughter nodes S and 2ndConj and the syntactic structure of (3.b) is another instance of this configuration. So (4) and (4') are two occurrences of this configuration, belonging to the two expressions (3.a) and (3.b) (which in this case are identical with the two occurrences rather than containing them as proper parts).

This is how the pair of sentences (3.a) and (3.b) can be seen as a case where a representational account of semantic composition cannot be reduced to an account involving only operations on semantic values. First, observe the following. The second daughters of the two occurrences are identical. (They are identical surface strings and there is no reason to assume that the two instances of this surface string have different syntactic structures.) So the semantic contribution that the second daughter makes to (3.a) and (3.b) can be assumed to be the same too. The first daughter of (4) differs from the first daughter of (4'). But it may be argued that these two sentence structures nevertheless have the same semantic value: They express the same proposition, one that is true in a world in which there is a set of ten balls, exactly nine of which are in some given bag, while one is missing. If this proposition is the semantic value of both the first daughter of (4) and of (4'), then semantic value composition cannot give us what we want. For suppose the semantic values of the two daughters of the outer S node in (4) determine the semantic value of their mother node (the outer S node, which is the node of the entire sentence (3.a)), the proposition that of the ten balls nine are in the bag and one is not and that that last ball is under the sofa. Then, since the semantic values of the two conjuncts of (4') are the same as those of the to daughters of (4'), the prediction is that (3.b) has this meaning too. But we already saw that this is wrong.[11]

The pair in (3), and likewise the alternative version in (5) of the footnote, give

---

[11]The pair of sentences in (3), with its conjunction *, and in fact*, may seem rather artificial. The original example, due to Partee, involves pairs of sentences, as in (5)

(5)   a.   One of the ten balls is not in the bag. It is under the sofa.

       b.   Nine of the ten balls are in the bag. It is under the sofa.

This version can be taken as an illustration of the question at issue, provided we think of the well-formed expressions of (the given fragment L of) English as including also sequences of well-formed sentences connected by periods ('.') and that such sequences are built syntactically by combining a given sentence or sentence sequence with a new sentence that gets tacked on to it as (another) .-conjunct. Then the sentence pair in (4.a) and that in (4.b) can be seen as two occurrences of the same mother-daughters configuration, which presents the same problem for a semantic values recursion as the pair in (3).

us a negative answer to the reducibility question only if there is a representational account of it that does not lead to the intuitively wrong prediction that the intuitively correct interpretation of (3.a) is also a possible interpretation of (3.b). This is one of the achievements of the original version of DRT presented in (Kamp 1981). The solution to this puzzle offered there may be too well-known to bear repeating in full. But it will helpful for the general line of thought in these reflections, to recall its central features. As in (Kamp 1981) we will focus on the sentence pair version in (5).

The first and most important point about the DRT treatment of the sentence pairs in (5) is that their first sentences yield distinct DRSs $K_{a,1}$ and $K_{b,1}$, which DRSs determine the same proposition (as they should according to the discussion above). But these DRSs each play their part in the interpretation of the next sentence of the pair – an interpretation that takes the form of trying to extend the given first sentence DRSs $K_{a,1}$ and $K_{b,1}$ to extending DRSs for the full sentence pairs (5.a) and (5.b). And $K_{a,1}$ and $K_{b,1}$ differ in a way that explains why this extension is possible in the case of (5.a) but not in the case of (5.b). The crucial difference is that $K_{a,1}$ has a discourse referent introduced for the indefinite DP *one of the ten balls* which can serve as 'antecedent' for the pronoun *it* of the second sentence. $K_{b,1}$ has no such discourse referent. (Instead it has a discourse referent for the set of nine balls that are in the bag, but that discourse referent is no good as an antecedent for *it*.)

The way in which DRT manages to account of the difference between (5.a) and (5.b), and likewise between (3.a) and (3.b), crucially depends on this representational difference between $K_{a,1}$ and $K_{b,1}$ – that $K_{a,1}$ has a discourse referent for the missing ball whereas $K_{b,1}$ does not. Switching back to the sentences in (3): it is this difference that enables the DRT account to explain why the content of (3.a) doesn't automatically qualify also as the content of (3.b). If it wasn't the DRSs $K_{a,1}$ and $K_{b,1}$ themselves that enter into the determination of the semantic values of (3.a) and (3.b), but instead the propositions determined by these DRSs, then we would be stuck with the unwanted prediction that (3.a) and (3.b) mean the same thing. It is by relying on a representational difference between $K_{a,1}$ and $K_{b,1}$ that is not reflected by the propositions they determine that DRT manages to avoid this prediction.

But can we see this as a conclusive demonstration that we need semantic representations with representational properties that are not fully reflected by the semantic values they determine, but are essential to the predictions that our semantics makes? That isn't nearly as evident as it seemed to me

when I came up with the account of (Kamp 1981) of which the last couple of paragraphs reviewed the essential features. First, the crucial difference between (3.a) and (5.a) on the one hand and (3.a) and (5.b) on the other is that in the former the pronoun *it* can find an antecedent in the first conjunct or sentence whereas in the latter it cannot. Perhaps this difference is one that falls within the competence of another module of the grammar than the one dealing with the compositional semantics, perhaps a module that is part of its syntactic component and one that already has done its work as part of assigning the sentences their syntactic parses, in which the difference is explicitly recorded, for instance by a coreference marking between *it* and *one of the ten balls* in the syntactic structure for (3.a), which would be absent from the syntactic structure for (3.b). It isn't easy to see how this could be a matter of syntax, for one thing because the syntactic principles hat have been found to constrain pronominal anaphora in English and other languages (as in Chomsky's Binding Theory) do not seem to apply here – they have nothing to say about the difference between (3.a)/(5.a) and (3.b)/(5.b). But even if we think of the pronoun anaphora module as operating at a point where the semantics has already run its course, predicting for both the a-cases and the b-cases the semantic value that only the cases can have, the module might conceivably act as a filter that rules out the semantic value assignment to the b-cases as somehow involving the assumption that *it* can be understood as referring to the missing ball, which in these cases is impossible.

There is a family of approaches along these lines, in which sentences and discourses with anaphoric pronouns like the *it* of (3) or (5) are preprocessed before the compositional semantics comes into action. A brief discussion of this approach is given in the next section, Section 2.2.1. But there is also a quite different approach, which avoids the need for a representational solution to the problem that such examples seem to pose. It consists in redefining the notion of semantic value in such a way that the first conjuncts of (3.a)/(5.a) and (3.b)/(5.b) no longer have the same values; the new semantic values capture precisely that difference between these sentences that accounts for the difference of the conjunctions. This approach has been pursued vigorously in the eighties and nineties. It is now most widely known as 'Dynamic Semantics'. We will look at this approach in more detail in Section 2.2. But a few words should devoted it right away.

First, refining the notion of semantic value may have been motivated initially by the desire to provide a strictly compositional treatment for examples like those in (3) and (5) adopted by those who have wanted to preserve strict semantic compositionality. But the new semantic values the Dynamic Se-

mantics introduced for this purpose have proved to be remarkably fruitful in a much broader sense. In particular they have led to a new way of thinking about many of the foundational issues in semantics and logic. It is this new set of concepts and this new way of thinking about semantic and logical issues that form the core of Dynamic Semantics as we know it today.

It should also be noted that the way I have been presenting the issues in this section is potentially a little misleading. When Dynamic Semantics set out to develop its own account of the phenomena illustrated by donkey sentences and donkey discourses, the account presented in (Kamp 1981) made no attempt to be strictly compositional in the sense discussed here. The claim was rather that such an account was probably impossible, but that there was no compelling reason for expecting that it would be possible and that a strictly compositional wasn't really needed – nothing speaks against the assumption that the systematic relation between perceivable form and meaning without which language could not function involve formal aspects of the semantic representations that mediate between overt form and truth conditions . I still am of this persuasion today. But it is conceivable that if DRT had adopted Bottom Up DRS construction from the start, the Dynamic Semantics reaction would have been different, and this part of the recent history of formal semantics and logic would have taken a somewhat different course. I am not quite sure of this, since Heim's File Change Semantics (Heim 1982,1988), which was developed at essentially the same time as DRT and which deals with the same phenomena involving donkey sentences and donkey discourses, is considerably closer in spirit to Dynamic Semantics than is DRT, and I assume that Heim's work was known to Dynamic Semanticists early one.[12]

So much, for now, about objections of the first type. About objections of the second type I will be briefer. The issue raised by the second type objections is first and foremost an empirical one: Are there multi-sentence discourses whose truth conditions cannot be described as composed of propositions that are expressed by each of the sentences of which the discourse is made up? Some of the discussions focused on 'donkey discourses' like (6).

(6)   Pedro owns a donkey. He beats it.

---

[12]For what I know, the crucial assumptions of FCS antedate DRT by a year or so, but the two theories were developed independently to a point where when Heim and I put two and two together, they had matured to the point where it could be ascertained that they made the same predications about the phenonema in the intersections of their respective domains.

"Are the truth conditions", it was asked, "really what DRT (and other Dynamic theories) take them to be – those according to which, for instance, the truth conditions of (6) are that there exists a donkey that Pedro owns and beats? Isn't the presence of the anaphoric *it* an indication that the speaker must have had some particular donkey of Pedro's in mind when she used the indefinite *a donkey*?" The question should not be taken lightly. Often the presence of a subsequent pronoun that must be construed as anaphoric to a given indefinite appears to be a strong clue that the speaker had a particular instance of the indefinite in mind and that it was about that instance that she wanted to make a statement – that the speaker used the indefinite *specifically*.[13] And perhaps it is possible to develop a sentence semantics in which the semantic value assigned to a sentence containing an indefinite may be made to depend on the context, in such a way that the sentence gets a different value in a context where the indefinite is used specifically than when it is used non-specifically. In such a semantics 'Pedro owns a donkey', with specific use of *a donkey* might get as its semantic value the proposition that Pedro stands in the ownership relation to the particular donkey that the speaker has in mind. A similar line might then also be taken for sentences containing pronouns, such as for instance 'He beats it': In a context where the speaker thinks of Pedro when using *he* and of the donkey that she thought of when she used 'a donkey' when she now uses *it*, it might be argued that this endows the second sentence with a proposition to the effect that the first of these two individuals owns the second. In this way we would get intuitively plausible truth conditions for the discourse 'Pedro owns a donkey. He beats it.' as the conjunction of the two propositions that are the semantic values of its two sentences in the context set by the referential intentions of the speaker. Note well, however: If this can be made to work at all, then only if we are prepared to make semantic values dependent on what goes on in the speaker's mind in the way indicated. Not impossible, but perhaps not really the price that every objector to a semantics that extends beyond the single sentence would be happy to pay.

In addition there is a problem about donkey discourses in which the sentences are marked as having 'inferential evidentiality', which in English is typically conveyed by the use of epistemic 'must', as in (7).

(7)   'Someone must have been here recently. His footsteps are right there

Perhaps it could also be said of a speaker of this sentence that she has a particular individual in mind when she uses 'someone' in the first sentence,

---

[13]See for instance (Farkas 2002)

namely the unique person who left the footprints in the place she is referring to with 'right there' in the second sentence, even if the use of 'must be' suggests that the speaker cannot identify this 'someone' in any other way. But if this example is to be dealt with along the lines of a 'single sentence only' semantics, then an even heavier burden would seem to rest on the notion of 'having an individual in mind' than is apparent from(6). It is important not to forget that many 'donkey discourses' are more like (7) than like (6).

The point of these last two paragraphs has been that even if it is right that whenever a pronoun in S2 is anaphoric to an indefinite in S1 the combination S1.S2 expresses the conjunction of two propositions each of which is about the same particular instance of the indefinite in the first sentence – something that is a point of debate in its own right –, heavy reliance on speaker's intentions is needed if such an account is to be made to work in detail. I do not wish to suggest that speakers' referential intentions have no place in semantic theory altogether. (I may be more tolerant on this point – promiscuous, as some may want to put it –than many others, including supporters of the second type of objection. The role of speaker's intentions in a theory of linguistic meaning will be discussed later on.) But I do not think that the speaker's referential intentions should come into the semantics of sentence pairs with an indefinite in the first sentence and a pronoun anaphoric to it in the second – not in the way it would be necessary to appeal to them in a theory that treats discourses as conjunctions of mutually independent propositions expressed by their individual sentences.

2.1.1   The E-Type Approach to Donkey Anaphora

One approach to donkey pronoun problems, I noted in the last section, is to subject the sentences and discourse in which they occur to some kind of preprocessing. All the versions of this approach that I am aware of involve the elimination of the problematic pronouns from their sentences (or their syntactic structures, but this is a distinction that is typically not made) by replacing them with suitable definite descriptions. The first version of this approach was first proposed by Evans and was arguably anticipated by Geach, the one who introduced donkey pronoun puzzles from the scholastic literature to the 20-th century logic community, in his use of the term 'pronoun of laziness': Donkey pronouns are stand-ins, or abbreviations, for definite descriptions, which give explicit descriptions of the entities the abbreviating pronouns should be understood to refer to. The pronouns that are to be accounted for as abbreviations of definite descriptions Evans referred to as 'E-type pronouns', and this how they are usually referred to in this

family of approaches, usually referred to as the 'E-type approach'.)

To repeat, according to the E-type approach, the semantics of sentences and discourses with donkey pronouns can be given by first replacing these pronouns by the descriptions for which they are proxies and then deal with the semantics of the results of these substitutions by applying whatever analysis is available for dealing with those results. This presupposes among other things that a viable semantics is already in place for the substituting descriptions. Exactly what Evans thought about this I do not know, but I suppose he may have assumed that a Russellian analysis of definite descriptions would do for this purpose. In any case, today, after many decades of further exploration of nominal reference for English and other languages, the proper treatment of the descriptions for which donkey pronouns are supposed to be abbreviations – one might refer to these uses of definite descriptions as 'donkey descriptions' – can be seen to pose as many problems as the pronouns they are invoked to account for.

The conclusion to be drawn from these considerations has to be that a proper E-type treatment (for some fragment L containing donkey sentences and/or discourses) must satisfy the following two requirements:

(i) There must be a procedure for determining, for each occurrence of a donkey pronoun in a donkey discourse or donkey sentence, a definite description for which this pronoun token can be regarded as proxy. Evans already pointed out some of the problems with this. The problems have also been noted and commented on by others.

(ii) There must be a satisfactory account of the semantic contributions made by the donkey descriptions that replace donkey pronouns according to the vaious versions of the E-type approach.

Different versions of the E-type approach (including the 'D-type approach' of Neale (Neale 1990) and others[14], have addressed ways in which either of these requirements can be met. However, many do not really come up with any very specific proposals at all on either count. One consequence of this lack of specificity is that it tends to be difficult to assess what any of these proposals have to say about the question whether a genuine discourse semantics is needed or a semantics restricted to single sentences would suffice.

---

[14]For present purposes the D-type approach can be considered a variant of the E-type approach

As an illustration of the care that needs to be taken in drawing any conclusions about this question, consider once more the donkey discourse (6), repeated below.

(6)   Pedro owns a donkey. He beats it.

The first step of an E-type analysis of (6) is to replace the pronouns by suitable descriptions. The descriptions used in (8) have been made up ad hoc, and are not the results of applying some well-defined procedure. But I think these descriptions are in concord with what advocates of the E-type approach would suggest for (6). The description chosen in (8) to replace *it* is one that has actually been proposed for donkey discourses of this form. Pronouns like *he* in (6), which are anaphoric to proper names, are not discussed in the E-type approaches known to me, no doubt because such pronouns were never seen as problematic. The description *the person named 'Pedro'* will do as well, I presume, as any other one that picks out the referent of the antecedent – here *Pedro* – uniquely.

(8)   Pedro owns a donkey. The person named 'Pedro' beats the donkey Pedro owns.

Since there is to my knowledge no fully articulated method for obtaining the semantics of (8) that we can rely on here, we have to make certain guesses as to what truth conditions the method would assign to the two sentences of (8). There can be little doubt, I believe, that, when stated in Predicate Logic, the truth conditions of the first sentence should be as in (9).

(9) $(\exists y)(\text{donkey}(y)\ \&\ \text{own}(p,y))$,

   where p is an individual constant denoting the referent of *Pedro*.

But what about the second sentence of (8)? An old-fashioned translation into Predicate Logic which uses Russell's theory of descriptions leads to the formula in (10), which can be simplified to that in (11):

(10) $(\exists u)(\text{person}(u)\ \&\ u = p\ \&\ (\forall v)(\text{person}(v)\ \&\ v = p \rightarrow v = u)\ \&\ (\exists w)(\text{donkey}(w)\ \&\ \text{own}(p,w)\ \&\ (\forall z)(\text{donkey}(z)\ \&\ \text{own}(p,z) \rightarrow z = w)\ \&\ \text{beat}(u,w)))$

(11) $(\exists w)(\text{donkey}(w)\ \&\ \text{own}(p,w)\ \&\ (\forall z)(\text{donkey}(z)\ \&\ \text{own}(p,z) \rightarrow z = w)\ \&\ \text{beat}(p,w))$

Note that (11) logically entails (10). So if we take the semantics of (6) to be given by the conjunction of (9) and (11), then its truth conditions are given by (11) on its own. These truth conditions are different from the ones predicted by DRT and other dynamic approaches in that they entail that there is a unique donkey that Pedro owns, and that it is this donkey that he beats). And perhaps that is right, at least for this particular example: Perhaps the presence of the anaphoric *it* does convey uniqueness upon the donkey that is talked about in the first sentence, and that wouldn't be conveyed in the absence of the anaphoric *it*.

Note also that in this analysis the proposition expressed by (6) is the conjunction of two propositions, separately expressed by the two sentence of which (6) is composed. But this comes at the cost of reincorporating, as it were, the proposition expressed by the first sentence into the proposition expressed by the second sentence.

As I said, any version of the E-type approach involves a 'preprocessing' step in which donkey pronouns are replaced by descriptions. That involves choosing a suitable description for each such pronoun and an account of how that is done has to be part of every such version. In the two examples discussed above – the *he* and the *it* in the second sentence of (6) – these descriptions could be constructed from material in the first sentence. Evans I noted above, was already aware that this doesn't work in all cases. Sometime the description needs to be recovered from the non-linguistic context, and in these cases an algorithm that only looks at the words of the preceding sentence of discourse doesn't stand a chance from the start. But even when the description can be constructed just on the basis of the preceding sentence of discourse, what are the general principles that permits these extractions? A fundamental difficulty one is facing here is that it is hard to see how the right description can be chosen if one doesn't have some kind of interpretation of the sentence containing the pronoun. At the very least an E type account would require some fairly complicated interleaving of the compositional semantics envisaged with the preprocessing steps that prepare the input for this compositional component. And as far as I can see, these steps will require access to and use of information that would qualify them as pragmatic according to criteria endorsed by many.

Bickering about what is semantics and what pragmatics may not be a very fruitful preoccupation. But the interleaving of theory modules that deal solely with the syntax semantics interface and modules that take other information into account is something that as far as I can see any account of

natural language meaning will have to countenance. Below in Section 3 I will present a form of presuppositional DRT in which such interleaving is explicit and essential. Semantic processing of a new sentence starts with the application of a module of pure syntax-semantics interface, the output of which is a so-called 'preliminary' sentence representation. This preliminary representation contains explicit semantic representations of the presuppositions triggered by the processed sentence. In a next step or series of steps these presuppositions are then resolved, often with the help of information that is widely taken as characteristic of pragmatic modules. After presupposition resolution has turned the preliminary representation into one that is presupposition-free this latter representation can then be merged with the discourse context. Both the first and the third phase of this process make use of the representational form of the representations involved. More about this in Section 3.

This is all I have to say about objections to DRT of the second kind, which claim that discourse meaning can be accounted for by a combination of sentence semantics and certain pragmatic theory components. But more needs to be said about systems of Dynamic Semantics, which were developed from a conviction shared with DRT that something like an incremental semantics of discourse is needed, but decided that DRT is not the right way to go about this. The next subsection discusses some aspects of the 'Dynamic Semantics Research Program' and its conceptual implications.

2.2 Dynamic alternatives

So far I have spoken of Dynamic Semantics as a program that aims to do semantics without any essential use of representations, but that is like DRT concerned with the incremental interpretation of multi-sentence discourse. (Perhaps it would be better to speak of the incremental assignment of semantic values in the present broader context.) But this is a way of characterizing Dynamic Semantics from the perspective of DRT and it might well be thought that this perspective is biased. In fact, it has often been stressed, and this prominently also by the developers of Dynamic Semantics themselves, that DRT isn't dynamic semantics at all. The thought goes, in a nutshell, as follows:

The incrementality of discourse interpretation, which DRT and Dynamic Semantics are equally concerned with, is captured in DRT by using the DRS K for the antecedent discourse as input to the construction of the semantic representation for the next sentence, and this construction takes the form

of building a DRS that extends K. But what for many formal semanticists is the essence of semantics – articulating the denotation relations between expressions and the world or situation they can be used to talk about – is not really different from what it has always been in Montague Grammar: for each expression E and each possible model M E denotes in M a *semantic value*: an entity from M or built from elements of M. This semantic value definition is just as 'undynamic', or 'static' in DRT as it is in the model theory of formal languages that we owe most of all to Tarski and in Montague Grammar.

Dynamic Semantics is different in this respect. It too treats the relation between expressions and models as central. But this relation no longer takes the simple form of expressions denoting in models single semantic values. Rather, expressions are analyzed as *transition vehicles* between semantic values; they transform input values into output values.[15] Hand in hand with this input-output concept of expression meaning goes a refinement of the semantic values. These are of finer granularity than the ones traditionally used in static semantics and designed to capture certain aspects of the incremental dimension of discourse interpretation. The new values are known as *information states*. Exactly what kinds of entities these are varies between the different versions of Dynamic Semantics. This has to do with the range of phenomena each given version is meant to account for. But in all cases information states capture some aspect of information that goes beyond pure propositional content, by containing some kind of record of previously introduced entities (just as DRSs keep such a record trough the discourse referents in their Universes). And because information states carry information of this kind, the output states that an expression E associates with a given input state can capture what new entities E introduces into the discourse of which it is part. What is (radically) new about Dynamic Semantics is this combination of a relational conception of linguistic meaning and some more refined notion of 'semantic values' between which these relations hold.

The relational dimension of Dynamic Semantics should be distinguished from the question of doing away with 'semantic representations'. This is clearly recognizable in what I believe to be the two first systems of Dynamic Semantics, *Dynamic Predicate Logic* and *Dynamic Montague Grammar*. Both

---

[15]One could still say that expressions denote semantic values, but now meaning by that that the denotation of E in M is an input-output relation, a set of pairs each of which consists of an input state and a corresponding output state. While such terminology is occasionally employed, we should guard against using it to downplay the fundamental difference between the static and dynamic approach to semantics.

were developed by Groenendijk and Stokhof ((Groenendijk & Stokhof 1991*a*), (Groenendijk & Stokhof 1990), (Groenendijk & Stokhof 1991*b*)). Dynamic Montague Grammar is in the spirit of Montague Grammar in that it directly defines a relation between expressions of the natural language fragment for which it is designed and the models in which these find their semantic values. But the semantic relation is defined in terms of input-output relations, the hallmark of Dynamic Semantics as I understand it and use the term in these reflections. Dynamic Predicate Logic develops the hallmark relational notion of meaning for a formal language whose syntax is that of First Order predicate Logic in its standard form.[16] The formulas f DPL can then be used as logical forms for expressions from the natural language fragments, but the details of these applications are by and large left to the discretion of the natural language semanticist who adopts DPL for such purposes.

The points of MG and DPL as alternatives to DRT go in a similar direction but are nonetheless different. MG shows that what the first versions of DRT were trying to accomplish can be accomplished without using representations like DRSs. DPL shows that these purposes can be accomplished by using a logical form formalism that is indistinguishable from Predicate Logic – the source of logical forms that was almost universally in use before Montague and still much used within philosophy as a semi-formal tool when it comes to clarifying logical aspects of natural language meaning; all that is needed is to give this formalism a relational semantics.

Both MG and DPL are thus, each in their own way, demonstrations that the original goals of the DRT of (Kamp 1981) can be accomplished without the specific assumptions that DRT was making, and thus as refutations of the claims made in (Kamp 1981) that these goals could not be achieved without those assumptions. But they do not show that there couldn't be other reasons why DRSs might be useful as intermediaries between natural language expressions and the models in which these determine their values. In fact, there are many 'hybrid systems' between DRT and Dynamic Semantics – systems that make use of some DRS language, whose DRSs are assigned to the sentences and discourses of the natural language fragments under study, but where the semantics for this DRS language is relational in the sense of Dynamic Semantics. The first systems of this kind were developed my Muskens ((Muskens 1991), (Muskens 1994*b*), (Muskens 1994*a*), (Muskens 1996)). But

---

[16]This semantics makes it possible for an existential quantifier with a given syntactic scope (defined in the usual way in terms of the syntactic form of the formula to which it belongs) to reach semantically beyond that scope, so that it can bind syntactically free occurrences of its bound variable in subsequent conjuncts.

many others have built on Muskens' work, developing systems of 'compositional DRT', systems that vary as a function of the phenomena they are intended to handle. Some of these systems will be discussed in Section 4.

It should be noted in this connection (as Groenendijk and Stokhof do explicitly in their (Groenendijk & Stokhof 1991a)) that something like the input-output relations that are central to Dynamic Semantics can be found in the DRT version of 'A Theory of Truth and semantic Representation'. But in (Kamp 1981) the relational dimension is limited to the semantics for complex DRS conditions, such as those that are used to represent conditionals and universal quantifications. For instance, the model-theoretic clause for conditional DRS conditions of the form $K_1 \Rightarrow K_2$ is to the effect that:

(12) Given an input assignment f $K_1 \Rightarrow K_2$ is satisfied if for every extension g of f that satisfies K_1 there is an extension h of g which satisfies K_2

This satisfaction clause speaks of relations between assignments (or 'embedding functions', as they are called in DRT), between f and g and between g and h. In DRT it is possible to limit the use of such relations between assignments to the semantics of complex DRS conditions because discourse updates are dealt with by 'DRS updating', i.e. extending a given DRS with the contributions made by a new sentence. DRT was focused on getting the truth conditions of discourses right (and not just of individual sentences), so a truth definition for the completed DRSs for the entire discourse was all that it took to be required. The important insight of Dynamic Semantics was that there is an important sense in which the mechanisms that are needed for the interpretation of sentence-internal structure – mechanisms, in particular, that can deal with sentence-internal donkey anaphora (i.e. with donkey sentences) are also responsible for the sentence-transcendent aspects of discourse interpretation. This, among other things, is one way of bringing out even more forcefully that sentence-internal and sentence-transcendent donkey anaphora are manifestations of the same phenomenon.[17]

---

[17] The similarity of these two manifestations is further enhanced in Dynamic Semantics by treating conditionals is constructs involving dynamic conjunction. What is represented in DRT as $K_1 \Rightarrow K_2$ is expressed in DPL by $\neg(\phi \ \& \ \neg\psi)$, where ϕ and ψ correspond to K_1 and K_2 and $\&$ is dynamic conjunction, defined by '$\phi \ \& \ \neg\psi$ connects information states i and j (in the sense that j is a state that $\phi \ \& \ \neg\psi$ can output when given i as input) iff there is an information state k such that ϕ connects i and k and ψ connects k and j. (I am using here the notation '$\&$' for the dynamic conjunction of DPL, as in (Groenendijk & Stokhof 1991a), rather than the now generally used dynamic conjunction symbol ';'.) Both a donkey pronoun in the consequent of a conditional (or in the nuclear scope of a quantifier) that is anaphoric to an indefinite in the antecedent of the conditional (or in the

It more or less follows from the last paragraph that it must be possible to formulate a relational semantics also for DRS languages. In particular, the operation of updating one DRS K (the current discourse context) with the representation K' for the next sentence (an 'improper DRS' in the terminology of (Kamp & Reyle 1993)) should now lead to an updated DRS K'' which connects the information states i and j iff there is an information state k such that K connects i and k and K' connects k and j. It is not difficult to show that this condition is satisfied when a relational semantics is given for the DRSs of 'A Theory of Truth and Semantic Representation' using the information states of DPL. In such a 'relational' version of DRT, the syntactic operations involved in updating K with K' correspond to the semantics of dynamic conjunction in the same way that lambda conversion stands to functional application. In fact, we can represent the update of K with K' as '$K; K'$', as way of making explicit how its semantics depends on that of K and that of K'.[18]

2.3 Information States of computers and Information States in us

The central idea of Dynamic Semantics is that linguistic meaning can and should be analyzed in terms of transitions from input to output information states. An inspiration for this conception – both Groenendijk & Stokhof and Muskens testify to this [**references**] – is the use that is made in a general theory of the workings of of computers of the *internal state* that a computer is in at any point in the course of a complex computation that it is carrying out. The theory is abstract in that it abstracts from the individual properties of particular computer hardware. What all computers have in common according to it is that they can be described as devices whose computationally relevant states consist in values assigned to a (typically large) set of variables. Processing by the device consists in changing the values of these variables, with some changes taking place at each computing step. A computer program consists of a sequence of instructions for changing values of some of the variables, and running a program on a computer consists in the computer executing the successive instructions of the program for changing its internal

---

restrictor of the quantifier) and a pronoun that is anaphoric to an indefinite in a preceding sentence are now analyzed as situated in the second conjunct of a dynamic conjunction with the indefinite situated in the first conjunct.

[18]This is one of the things accomplished in the work by Muskens mentioned above. Muskens does more, however. He defines a version of the Lambda Calculus in which DRT, with a relational semantics for it, can be embedded. More about this in section 4.

state, from a given starting state as point of departure.[19]

Since the phenomena that led to Dynamic Semantics had to do with entity-introducing and anaphoric noun phrases, the notion of an internal state as it is used in the theory of computation looked like a good model for the information states that language interpreters are in when they are processing inputs containing such noun phrases. Moreover, formal semantics, in the form in which it had been conducted since the founding of model theory in the work of Tarski for formal languages like the Predicate Calculus and the transfer of that approach to the semantics of natural languages by Montague, was ready made for such a notion of information state, you could say, since it makes a central use of the notion of *satisfaction* of a formula in an model by an *assignment* of values to the variables of the language in question. Indeed, the information states of DPL just are such variable assignments. And when information states are simply assignments, not all that much needs to be changed when we go from a 'static semantics to a Dynamic Semantics. All that is needed is to promote assignment changes, which in the satisfaction definitions of static semantics are involved in the clauses for the quantifiers, as the norm, i.e. as actually or potentially involved in all clauses of the compositional semantics. (DRT can be seen as representing an intermediate stage in this transition form static to dynamic in that it deals with the semantics of complex DRS conditions in terms of relations between embedding functions.) As noted, different systems of Dynamic Semantics make use of different notions of information state; and I think that there is wide agreement within Dynamic Semantics community that there are better choices than plain variable assignments. A little more about this later on.

Here is a thought suggested by the analogy between internal states of computers as assumed in Dynamic Logic and the information states that are involved in human interpretation of noun phrases. Dynamic Logic, we noted, was motivated by the desire to abstract away from details of hardware implementation that detract from what is essential to the ways in which digital computers process information. Can we see Dynamic Semantics as a similarly general theory of human language processing, which abstracts away from details of how we process (discourse containing) noun phrases? The thought can take different forms. First, it may be that different speakers of the same language (e.g. English) go about processing the same linguistic constructions in ways that differ in their details. Second, it may be that the

---

[19]This general theory of computing is now mostly referred to as 'Dynamic Logic'. For overviews see e.g. (Harel 1984), (Eijck & Stokhof 2006).

same speakers process the same constructions in different ways depending on the contexts in which they encounter them. And thirdly, the very same speaker may differently process the same linguistic construction in the same contexts, depending on experience, or perhaps on extraneous factors like concerns that compete with the current interpretation task. Let me embroider a little on this third possibility. What I have to say also applies, mutatis mutandis, to the other two.

The thought has much to do with the notions of linguistic competence and linguistic performance. A dominant picture in linguistics, prevalent since the generative revolution due to Chomsky, is that what is common to all mankind is an inborn capacity for acquiring human language. The 'language engine' that each of us is equipped with form birth can be activated through exposure to any of the different human languages that are spoken around the globe (or, per impossibile, any of those that were spoken within the past), and when that happens the vast majority acquires the grammar and the core vocabulary of that language with astonishing speed. And the competence that the learner has thus acquired can then be applied to the wide range of different ways in which human beings use language, in production and reception. But exactly how does the competence get realized in this wide range of performances? Let me, in speculating about this, focus on reception – on language interpretation – since these are the performance processes that semantics has been primarily concerned with. (Even semantic accounts that present themselves as neutral with respect to the production-interpretation distinction start from the 'raw linguistic input' – sequences of sounds or letters, perhaps 'tokenized' as a sequence of words and morphemes, and then articulate how the forms of these sequences determine their 'meanings'.) The first idea, more or less a repetition of what I have already said, is that the grammatical competence that an interpreter has acquired when learning her language need not fully determine the details of how she arrives at the interpretation of a given spoken or written input that her grammar determines. But if the interpretation processes that human interpreters perform are not fully determined by competence and input, what does determine them? I don't have any suggestions here, except that the details of such interpretation processes are the result of the accumulated linguistic experiences of the given interpreter, not only during the acquisition of her competence, but during the sum of all her exposition to and use of language. This total experience will generally exceed by a vast margin what has gone into her competence acquisition and change over time as details of use and exposure change; all kinds of heuristics may be taken on board in the course of these ongoing language-related experiences.

I am not sure that this is what governs linguistic performance. But if it is, then the question that needs to be asked is how these diversifying forces do not lead to interpretations that deviate from the form-meaning relation that is part of the interpreter's competence. This is where the next thought comes in: language processing isn't just the application of a process dictated by previous linguistic practice, and which therefore may differ even for the same interpreter, the same input and the same context, but involves in addition the application of a 'checking' module that directly implements the interpreter's competence and checks whether the interpretation that the heuristically based processor has come up with is in accordance with the grammar of the language (as including both syntax and the syntax-semantics interface). I am assuming here that checking whether a proposal for the syntactic structure and semantic interpretation of a given input – in the form, say, of a syntactic tree and semantic representation recursively determined by that tree – satisfies the constraints imposed by the grammar is much easier (or computationally efficient) than using the grammar to compute tree and semantic representation from the input. In fact, heuristically based performance and checker may well operate hand in hand on-line, perhaps along the lines of prediction-driven syntactic parsers (chart parsers, like the Earley parser,(Earley 1970)). Moreover, so long as the heuristic processor operates in accordance with the grammar, the overhead cost incurred by the checker may be minimal. It is only when the checker filters out proposals that disagree with the grammar that significant extra cost will arise, as then the proposed structure will have to be corrected, possibly requiring more or less serious backtracking. But even if this does happen ever so often, the processor-cum-checker architecture may be the more efficient design over all. It is also conceivable that several heuristic processors work in parallel, with as output the first structure proposal that is passed by the checker.

To repeat, the thoughts about competence and performance I have allowed myself to indulge in in this section are mere speculations. If true, they would seem to have non-trivial consequences not only for language processing but quite possibly much more widely for human cognitive behavior that is constrained by some kind of norms. What I will have to say about production in Section 5 can also be seen as a variation on the theme of the present section.

2.4 Dynamic Semantics with DRT. Linguistic and Non-Linguistic Contexts

A number of systems of Dynamic Semantics retain DRT insofar as they adopt a DRS language as the one for which they give a relational semantics, with input and output information states. (Such systems, you might say, are like DPL, except that the object language syntax is that of the chosen DRS-language rather than that of Predicate Logic.) More about such 'Dynamic-Semantics-with-DRT' systems in Section 4.

The question that such systems raise is what role is left for the DRS-languages they adopt. Recall that much of the work on pronoun anaphora that in DRT is contributed by DRSs as Discourse Contexts is taken over in Dynamic Semantics by the information states that expressions of the object language can update. DRSs are now (among) the expressions that produce updates of information states, just as the object language expressions of Dynamic Semantics systems that assume other object languages, like DPL, or Dynamic Montague Grammar. Are there still any good reasons for preferring DRS-languages in this role?

A positive answer to this question would presumably have to do with remaining advantages in dealing with the syntax-semantics interface for natural languages: by translating natural language first into a DRS or DRS-like structure and then relying on a model-theoretic semantics for the DRS-language still has benefits, even if the semantics is given in the form of input-output relations between information states. I believe that at least some of the 'Dynamic-Semantics-with-DRT' systems have been motivated in just this way. This should become clear from the discussions of such systems in Section 4. But in connection with the donkey-type phenomena on which we have been concentrating in this section a case for such systems is not so easily made, and the reason for that is one for which I have largely responsible myself. In the presentation of DRT in (Kamp 1981) I went out of my way to that what the theory meant to accomplish was not a full-blown account of pronoun resolution, but only to provide a framework within which a detailed account of resolution might be developed and which imposes certain constraints on what resolutions are possible (stated in terms of DRT's notion of accessibility). That much can easily emulated in other ways, for instance by treating pronoun resolution as a decision about the choice of variables in the way that this has been proposed for DPL, where resolution takes the form of choosing the same variable for the pronoun and the antecedent to which it is resolved in the DPL formula that serves as logical form.

But there is nevertheless a difference here between DRT[20] and, for instance, DPL. While original DRT has nothing to say about resolution beyond the constraint imposed by accessibility – i.e. that the discourse referent for the pronoun's antecedent must be accessible from the position of the pronoun – the framework it offers allows for the addition of further anaphora resolution principles.[21] I do not know how proponents of DPL would handle such further constraints on pronoun resolution. But as far as I can see this would require separate modules, which motivate variable choice, and that would make use of information of the partially constructed logical form in some manner or other. (I believe this problem arises equally for other 'Dynamic-Semantics-with-DRT'-systems, but probably in different ways, varying with the details of how pronoun antecedent relations are handled in those systems.)

What DRT does have to say about anaphora crucially depends, we saw, on the form of DRSs, which enables them to serve as discourse contexts for the interpretation of pronouns in sentences or clauses that are to be interpreted next. This double function of DRSs, as content identifiers and as discourse contexts, endows the theory with a self-containedness that from a general perspective of theory construction has a certain appeal: All that a DRT-based treatment of a natural language discourse needs to take into account is that discourse itself. Even though context is needed for pronoun resolution, the contexts that are needed, according to the original DRT approach, are the result of the DRSs that have already been constructed from part of the discourse that is being treated.

But what when the contextual information needed to resolve an anaphoric pronoun (or other anaphoric noun phrase) is not contained in the discourse context? When we look at occurrences of definite noun phrases in texts[22], then the majority do not refer to entities that have been mentioned before. Not really surprising, but the point needs making. Pronouns are an exception in this regard, in texts they are almost always anaphoric to some other

---

[20]as presented in (Kamp 1981); but the observation applies equally to all later DRT incarnations

[21]This is so in particular for the Presuppositional DRT that I have been using myself since the late nineties. For Presuppositional DRT see Section 3.

[22]In the terminology I am using here the definite noun phrases of English are (i) pronouns, (ii) definite descriptions, (iii) proper names and (iv) demonstrative phrases (those beginning with, or consisting of, the words *this*, *that*, *these* or *those*). This four-fold classification is not universal. Many other languages have different type of repertoires of definite noun phrases, with different semantic and pragmatic properties.

noun phrase that also occurs in the same text and in the vast number of cases to a noun phrase that occurs earlier.[23] But this is not so for other definite noun phrases, and in particular not for definite descriptions. More than half of the definite descriptions one finds in most texts are 'discourse new', i.e. they refer to entities that the text has not mentioned up to that point. This fact has been cited as objection against Heim's File Change Semantics (Heim 1982,1988), in which definite descriptions are treated as 'familiar' noun phrases - noun phrases that must find their referent in the context in which they are used. By itself that is right and unproblematic, but a difficulty arises when the only context available is the discourse context. To the extent that FCS is committed to this assumption – and this regard DRT is no different[24] – then we have an evident conflict with the way in which many definite description occurrences get their references (Fraurud 1990).

The obvious response to this observation is that evidently there is more to context than just the discourse context. What the theory – FCS, but by the same token DRT – needs is a richer notion of context. Formally it isn't difficult to provide this. In DRT we can assume that the interpretation of a sentence or discourse starts with a non-empty DRS, which contains the contextual information that is needed to interpret the discourse-new noun phrases occurring in it. And a Dynamic Semantics system can assume that the initial input state contains this information. (In a Dynamic-Semantics-cum-DRT system this would be an output state of the starting DRS, when this DRS is applied to the information-free information state.) The drawback of any of this is that from a theory that may have been false but capable of making non-trivial predictions we move to one that is neither. Or at least this will be so unless we have more to say about how the non-empty starting DRSs or information states are determined in actual interpretation situations. But that involves much more than just looking at language.

There is no simple solution to this problem. What we need is a way of assessing what information is available to an interpreter on the basis of what he knows on the basis of previous experience, and more specifically what knowl-

---

[23]I am treating kataphora, where the anaphoric expression precedes its antecedent, as a form of anaphora.

[24]The only justification why this objection was formulated at the time as an objection to FCS rather than to DRT was that the original presentation of DRT had nothing to say about definite descriptions. The only noun phrase types considered were pronouns indefinite descriptions and proper names. Had DRT included definite descriptions while also restricting the relevant contexts to the discourse context, it would have been open to the objection just as much.

edge he shares with the speaker/author, and on which the speaker relies in her choice of definite noun phrases. A step in the right direction – or so I would like to claim – is the notion of an *Articulated Context*, as developed in (Kamp 2019*a*) and other publications. Articulated Contexts are complex information-carrying structures which contain the current discourse context as one of their components, but also have other components which store information that the interpreter has accumulated through earlier experiences as well as, in the case of face to face communication, information about entities in the immediate environment in which the exchange takes place (a component that is needed to deal with the deictic uses of definite descriptions and pronouns). Articulated Contexts, in which information is distributed over the different context components according to its source, not only serve as containers of information that is needed for discourse-new definites, but also to account for certain differences between the different definite noun phrase types, such as in English (and a range of other languages) between definite descriptions and pronouns.

Something like the notion of an Articulated Context is needed to deal with the phenomenon of discourse-new definites. And the structure of Articulated Context is also helpful when it comes to describing, at a first, largely informal level, what the different possible uses are of th different types of noun phrases (e.g. which can be used anaphorically and which deictically). But the difficulty, to repeat, is to give the notion operational bite: How are we to find operational criteria to determine what Articulated Contexts are available to which interpreters when? The problem that confronts us here is an instance of a more general difficulty: What can we say about form and content of information in the minds of a human agent? In the case at hand we are talking about information that is used in utterance and text interpretation. This kind of case may be special in that what we know about language interpretation and linguistic meaning allows us to abduct some of the properties the mental information structures must be like. But there is a serious danger of circularity in this, as the mental information structures ought to account for the results of linguistic interpretation,and that requires that we should have some language-independent access to what those mental information structures are like. This, you might say, is the fundamental predicament of armchair psychology and for any form of armchair linguistics that tries to build on it. I will return to this point in Section 5.

3   Presupposition

Starting in the nineteen nineties, DRT has undergone fundamental technical and conceptual changes through the incorporation of accounts of presupposition. The crucial insight here goes back to Van Der Sandt, who in his 'Presupposition Projection as Anaphora Resolution' (Van Der Sandt 1992)[25] notes that non-projection of linguistic presuppositions is subject to the same kind of logically motivated structural constraints as the resolution of pronouns and other anaphoric noun phrases: An anaphoric expression can find an antecedent within its sentence only when this antecedent occurs in a position that is accessible from that of the pronoun; likewise a presupposition doesn't project (it doesn't emerge as a presupposition of the complex sentence that contains its trigger) when it is verified by information found in this same positional relation to the presupposition's trigger.[26] Van Der Sandt concluded that presupposition and anaphora are two sides of the same coin – presupposition justification[27] and anaphora resolution are aspects of interpretation that are subject to the same logically based configurational constraints.

(Van Der Sandt 1992) is primarily concerned with the presupposition projection problem, which was the primary concern of semanticists in the days when the paper was written, as well as with the question how and when accommodation can come to the rescue in those cases where a presupposition cannot be verified without it. But the structural analogies between pronoun resolution and presupposition justification obtain just as much for the transsentential cases, those where the pronoun's antecedent or the presupposition verifying information is in the antecedent discourse. Here too there is a question of accessibility: The antecedent or presupposition justifying information must be available at the top level of the antecedent discourse and not in some subordinate position. But even when in discourse level cases of presupposition or anaphora resolution accessibility constraints are satisfied there is more to say about the role that such resolutions play in the incremental interpretation of discourse than is easily visible from sentence-internal cases.

---

[25]See also (Van Der Sandt & Geurts 1991) and in particular (Geurts 1999).

[26]We already noted that this constraint is known as the 'accessibility constraint' of DRT. The constraint was originally defined in terms of the structure of DRSs. Its fundamental nature was demonstrated by Chierchia and Rooth (Chierchia & Rooth 1984), who derived it from more fundamental logical properties of logical forms for complex sentences.

[27]For the introduction of the term 'presupposition justification' see (Kamp & Roßdeutscher 1994).

At the level of discourse, presupposition and anaphora resolution aren't just extra tasks that make interpretation more difficult for the interpreter. They also serve to connect the semantic contribution of the sentence containing the anaphoric expression or presupposition trigger in the right way with the antecedent discourse where the resolving information has to be found. In other words, presupposition and anaphora resolution help the interpreter in establishing discourse coherence. For pronouns this is an old observation, going back to (at least) Halliday and Hasan (Halliday & Hasan 1976). But the observation applies to anaphora and presupposition resolution across the board.

The incorporation of presuppositions into DRT that I have been working with for the past two decades also goes back in its essence to Van Der Sandt's original proposal: For each new sentence of a discourse, first a *preliminary representation* is constructed in which all presuppositions triggered within the sentence are explicitly represented, essentially as DRSs that are adjoined to its non-presuppositional DRS.[28] Moreover, anaphoric pronouns and descriptions are included among the presupposition triggers. (In fact, all definite noun phrases are treated as triggers of so-called 'identification presuppositions', presuppositions whose resolution yields a specification of the noun phrase's referent.) The construction of the preliminary representation of a sentence only makes use of information contained in the sentence itself. It is only during the next stage, when the presuppositions of the preliminary representation are resolved, that the discourse context, the representation of the antecedent discourse, is brought into play. Sometimes the discourse context will be all that is needed to resolve the presuppositions of the preliminary representation. But there are also cases where this is not so, we have noted, for instance with definite descriptions that are discourse-new. To deal with such cases a notion of context is required that also contains information that is not part of the discourse context, such as is found in the Articulated Contexts mentioned in the last section.

Even the Articulated Context, as it is available to the given interpreter, may not suffice to deal with all presuppositions of the preliminary representation. In some of these cases the interpreter may have to resort to accommodation, by improvising information that he does not have but that it is reasonable for him to assume given that the speaker expressed herself the way she did and that she wouldn't have unless she knew that the presuppositions carried

---

[28]This is somewhat of a simplification; sometimes presupposition representations come with their own subordinate presuppositions and sometimes presupposition representations are adjoined to a subordinate DRS of the main non-presuppositional DRS.

by her words were satisfied. But in pretty much all cases where the discourse context provides at least some of the information that is needed to resolve a presupposition the resolution imposes a connection between the sentence containing the trigger and its predecessor or predecessors in the discourse. In short, presupposition is an effective means of enforcing discourse coherence and making it visible to the recipient.

In fact, adding a presupposition trigger to a sentence for the sake of underscoring an evident connection between the sentence and the preceding discourse is often a conventionalized requirement. Classical examples of this are the presuppositions triggered by words like *again*, *too/also*, but also the use of a definite as opposed to indefinite descriptions. Failing to mark the connection that justifies the presupposition that introduced by such a presupposition trigger, by failing to insert the trigger into the sentence, is then itself a source of incoherence. This coherence constraint is Heim's 'Maximize Presupposition' principle (Heim 1984)]: When there is a choice between two expressions that differ only in that one triggers a presupposition that the other does not, and this presupposition is clearly satisfied in the given context, then choose the expression that triggers the presupposition.

The architecture of the described form of presuppositional DRT is one in which semantic and pragmatic interpretation stages alternate: Construction of preliminary representation is a process that satisfies many of the criteria that have been associated with semantics as distinct from pragmatics: It is concerned only with single sentences and only with the information that can be derived directly from the syntactic form of the sentence. Turning a preliminary representation into a full representation, on the other hand, by resolving its presuppositions (after which, in a further step, the resulting sentence representation may then be integrated into the discourse context,) has some of the salient features of pragmatic processes: Reliance on information in the discourse context and/or other contextual sources and a wide range of inference mechanisms in molding such information into the form that presupposition verification requires. Such an architecture comes much closer to widely assumed views about the general form of language interpretation and linguistic meaning, as a process in which phases of pure sentence semantics alternate with phases of discourse integration in which there is access to contextual information and modes of reasoning of all kinds.

As noted, presupposition resolution is often impossible without recourse to

accommodation.[29] But accommodation isn't always possible even when it is needed. The conditions under which it is possible, and the form it may then take, vary from one presupposition trigger to the next. How they vary is, as far as I can see, the part of presupposition theory that still is the least well understood. But one aspect of the use of accommodation in presupposition resolution appears to be quite clear, and it is one that should be mentioned here, since it relates to an important aspect of meaning to which I want to devote a couple of lines for its own sake. Among the presuppositions for which accommodation is comparatively easy are the identification presuppositions of definite descriptions – both their existence presuppositions and their uniqueness presuppositions. In this regard the identification presuppositions of definite descriptions are like those triggered by *again*, which are also quite unproblematically accommodated when necessary.[30] But there is also an important difference between the identification presuppositions of definite descriptions and the presuppositions triggered by words like *again* or *too/also*. Whether the presuppositions triggered by *again* or *too/also* is or isn't resolved, the non-presuppositional function of the sentence – e.g. the proposition asserted – is well-defined. When A says to B: "Yesterday John fell asleep during the lecture again", then the asserted proposition is that there was an event yesterday of John falling asleep during the lecture. For the utterance to pass as felicitous it must be possible for B to assume that this has happened before: on other days in the past John also fell asleep during the lecture. But even if B has good reasons to assume that this never did happen before, the non-propositional content stands. (In fact, it is only by determining what the non-presuppositional propositional content is of a sentence containing *again* that it is possible to identify the content of the *again* presupposition.)

For sentences with definite descriptions the matter is different. Suppose for instance that A says to B: "I just met the piano teacher of my daughter." This utterance must be construed as expressing a proposition that is (among other things) about the piano teacher of the speaker's daughter. I take it that

---

[29]Most of what I know and understand about presupposition accommodation I owe to David Beaver. See in particular (Beaver 2001) and the references therein to Beaver's own work. Three is no explicit reference to this debt in what follows. But those who no me well enough will be aware of this debt too.

[30]The only real obstacle against accommodation of either kind of presupposition is when the interpreter takes himself to have evidence that the presupposition is false, e.g. that there isn't anything that satisfies the content of the description or that there was an event or state of the kind described by a sentence with *again* that preceded the event or state that the sentence reports or says something about.

until it has been established that there is a unique such individual – in the way that is accomplished by resolving the description's identification presupposition – no well-defined propositional content can be determined. In other words, in this case resolution of the presupposition is a necessary precondition to determining the non-presuppositional content. And when resolution has succeeded, the propositional content will *depend on* the outcome of that resolution, since it is this outcome that is needed as direct object argument for the verb *meet*. When the identification presupposition can be resolved on the basis of information that is available in the given context, then this dependence is unproblematic. It is just an instance of the general assumption made by all semantic frameworks that treat discourse interpretation as incremental, viz. that the propositional content of later parts of the discourse can depend for their propositional content on earlier parts. But when resolution requires accommodation, then the matter is different. The accommodated information is new – it is not part of the context as it was given to the interpreter. But without it the non-presuppositional propositional content is not properly defined. So if the information that the interpreter takes home from his interpretation of the utterance is to be well-defined the accommodated information will have to be part of it.

It is important, however, that in the information structure that the interpreter ends up with the accommodation is kept distinct from the non-presuppositional propositional content that depends on it. That is, the information structure as a whole must be articulated into these two distinct parts. Keeping the parts distinct is necessary because they typically play different roles in the further use of the information. For instance, only the non-presuppositional content can be suitably negated by the interpreter, e.g. by him saying something like: "No, that isn't true."

Such articulations of information content are central in work on the distinction between *at-issue and not-at-issue content* (see e.g. (Potts 2005), (Simons, Tonhauser, Beaver & Roberts n.d.), (Murray 2014)). According to this work the result of interpretation (e.g. of an assertion made by means of an indicative sentence) is normally a complex of propositional contents, one of which is the at-issue content, while the other parts, all of them non-at-issue content, are connected with the at-issue content in various ways (e.g. as presuppositions or as implicatures). The case we are discussing is a special and comparatively simple example of this, in which there is besides the at-issue content just one non-at-issue part (viz. the accommodation that the interpreter has made). But what is noteworthy about this example is that the information structure is not a combination of two independently defined

propositions, which are related to each other by some information-theoretical connection, but a complex in which the proposition determined by one part is defined only in the context provided by the other part.

A framework that has been designed to deal with examples like the one under discussion – that of an information complex in which one part derives its content from a binding relationship to some other part – is the *Layered Discourse Representation Theory* of Geurts and Maier (Maier 2013). LDRT is an extension of DRT whose representations (its 'LDRSs') can represent information as consisting of different bits that are situated at different 'information levels'. Among the levels that can be represented in LDRSs are in particular the level of presupposition and the level of assertion. And LDRT is set up in such a way that a bit of information that is primarily located at one level can be referentially connected to a bit that is primarily located at a different level and in this way contribute to the content of this other bit. For the example we have been discussing, in which accommodation has been used in the resolution of the identification presupposition of a definite description, LDRT seems perfect. A question to which I do not have the answer but would like to find one is how well-suited LDRT is in general for the information structures that the at-issue-not-at-issue community advocates.

A little more about LDRT will be said in the next section. LDRT was developed for the purpose of dealing with information structures we have been discussing. And for these it seems just right.

4 DRT-based Extensions and DRT-related Alternatives

This section is a brief aperçu of some of the systems to which DRT has been in one sense or another a kind of ancestor (as incidentally recorded in their names, all of which end on 'DRT'). The little I will have to say about each of the systems I will say anything about at all will be far too short to do proper justice to any of them; at best what I will say may serve as a stimulus to look at all or some of these systems more closely by consulting the publications in which they were originally proposed, pretty much all of which can be downloaded from the web. But the survey is also not meant to be complete; some systems will only be mentioned, and some I may have missed altogether.

From the perspective of these reflections the systems to be discussed can

be divided into two types. On the one hand there are systems of Dynamic Semantics that retain DRS notation in some capacity, but which define the semantics of DRSs in terms of input-output relations between information states. On the other there are systems that are more like the original versions of DRT in that they do not rely on a relational semantics for the semantic representations/logical forms that the theory assigns to natural language expressions. These systems differ from original DRT in that they add new aspects of the semantics and pragmatics of natural language and extend the representations they use accordingly. I begin with a few remarks about systems of the first kind.

4.1 Systems of Dynamic Semantics in which DRSs play some role

At the end of Section 2.2 we noted that there is one connective that is directly suggested by the relational semantics of Dynamic Systems. This is 'dynamic conjunction', standardly symbolized as ';', the semantic of which is relative product: $\alpha;\beta$ holds between information states i and j iff there is an information state k such that α holds between i and k and β holds between k and j. In Section 2.2 ';' was discussed as conjunction of DRSs. But dynamic conjunction is available for the formulas of any object language that a Dynamic Semantics system may adopt (such as, for example, the language of Predicate Logic in DPL).

Dynamic conjunction is one of the 'natural' operators in languages with a relational semantics. Another one is negation, '$\neg$', the operator that turns a formula ϕ into one that says of an input state i there is no way of transforming i into any output state j via ϕ. The combination of ; and $\neg$ can be used to simulate dynamic versions of the other standard connectives of Predicate Logic. For instance, the dynamic implication $\Rightarrow$, which in DRT is expressed by complex conditions of the form '$K \Rightarrow K'$', can be expressed as '$\neg(\phi ; \neg\psi)$'.[31]

---

[31]This observation is related to the fact that in first order DRT negation is the only operator needed to get the full complement of standard classical operators ((Kamp & Reyle 1993), Ch. 2). The reason why in DRT negation is all one needs has to do with the fact that there conjunction and existential quantification are structurally defined, conjunction via set membership – two DRS conditions can be 'conjoined' by putting them into the same condition set – and existential quantification by inserting the 'variable' (i.e. discourse referent) that is to be existentially bound into the right DRS universe. Simulating conjunction as joint set membership was actually not a good thing, since it

The input-output semantics of Dynamic Semantics systems enables us to define dynamic conjunction and negation irrespective of what assumptions are made about the in- and output states. In particular, these operators can be defined even when information states are identified with variable assignments. But when richer notions of 'information state' are chosen, it is often possible to define further operators, which would not be made available by simpler notions. Striking examples of this are the states used in Brasoveanu's PCDRT ('Plural Compositional DRT') and the *infotention states* of (Bittner 2014). Let me say a few things about PCDRT, as an example from the family of Dynamic Semantics systems that use DRS languages as object languages.

PCDRT is a good example of how the structure of information states provides the basis for the constructs of the DRS language adopted as object language and can serve as a guideline for how this DRS language is chosen. In fact, when the motive is (as it usually is in the design of these systems) to capture certain linguistic phenomena, the primary analysis that is offered of those phenomena is often at the 'meta-level', at which information states, and relations between and operations on them can be described with the familiar tools of classical logic and set theory. But part of the aim is to provide an object language (a DRS language for members of the family of systems we are discussing at the hand of PCDRT) in which this analysis can be expressed: The object language should make logical forms available for sentences and discourses in which the phenomena manifest themselves, so that we get the meta-level analysis of those phenomena back when the information state-based semantics for these logical forms are spelled out in full, by applying the relevant clauses of the semantics definition for the object language. (For systems of the kind we are considering the result of this spelling out is a meta-level description of the usually complex relationship between input and output states.)

The information states of PCDRT are richer than those of many other members in its family. They are rich in more than one respect, and it is worth our while to look at those different respects a little more closely. PCDRT's infor-

---

loses the asymmetry of dynamic conjunction that manifests itself in anaphoric behavior: a pronoun in the second conjunct of a natural language conjunction can be anaphoric to an indefinite in the first conjunct but not the other way round. Section 1.5 of 'From Discourse to Logic' was an awkward way of making up for this omission. It is possible, of course, to fill the hiatus by using $\Rightarrow$ to define dynamic conjunction, just as in Dynamic Semantics ';' can be used to define dynamic implication, viz. as '$\neg(\phi \Rightarrow \neg\psi)$'. (Kamp & Reyle 1993) should have mentioned this at least as an exercise. But we failed to do so.

mation states build on ideas that can be found in earlier Dynamic Semantics systems, and in the first place on the proposals of Muskens. Muskens' information states are defined for the version of Type Logic within which he develops DRT. One of the types of this Type Logic is that of 'discourse referents', which as we noted earlier, may be thought of in the present context as the variables, or 'registers', of the abstract theory of computation. Elements of this type take on values in different states – those that are identified with assignments in the first and simplest versions of Dynamic Semantics. A slightly different way of looking at this is to think of the value function as a 2-place function, from pairs of states and discourse referents to values. For many purposes this is a helpful way of thinking about how states determine values for discourse referents and so is the graphical representation of such functions as matrices, with the values occupying the cells of the matrix and, usually, the states corresponding to the rows of the matrix and the columns to the discourse referents.

One aspect of the richness of information states has to do with the ontological sorts that are represented by the values filling the cells of such matrices. These sorts are associated with the discourse referents that are admitted by the states. This is a dimension of richness and complexity that Dynamic Semantics systems share with DRT. (Kamp & Reyle 1993) discusses a number of distinct DRS-languages, in Ch's. 1 & 2, Ch. 4 and Ch. 5, which differ from each other among other things in what sorts are represented by the discourse referents that they make available. While the DRS-language of Ch's. 1 & 2 only has discourse referents for individuals, the languages of Ch. 4 also have discourse referents for sets of individuals and those of Ch. 5 discourse referents for times, events and states. With each new sort for which a category of discourse referents is adopted comes the obligation to say what the ontological properties of this sort are and its relations to other sorts that are represented by discourse referents from the system, an obligation that has to be discharged in one part by a proper definition of the class of models and in another by the truth definition of the DRS-language relative to models of this model class. Essentially the same obligations are incurred by the designer of a Dynamic Semantics system. (Recall in this connection that these systems also make use of models of some model class, since it is from such models that the values must be drawn that occupy the cells of the information states of the system.)[32]

---

[32]A special problem, which is of fundamental importance for the relationship between language and mathematical logic (of the kind that I was brought up with as a graduate student, see Section 1): When we adopt discourse referents for sets of individuals and

The DRS languages of PCDRT also admit discourse referents for the mentioned sorts. But in addition they also include a further sort, that of possible worlds. On this point PCDRT has made the same choice as a growing number of other dynamic systems. The issue goes back to the early days of Montague Grammar, when Gallin (Gallin 1975) proposed his so-called 'Type 2 Logic' as an alternative to Montague's Higher Order Intensional Logic. Type 2 is a system which treats possible worlds as regular citizens of the ontological universe, on a par with other kinds of individuals. In DRT terms this means that one adopts discourse referents for possible worlds, just as one has discourse referents for 'ordinary' individuals and may have discourse referents for times or events.[33]

PCDRT is among the systems that treat possible worlds in this way. This makes it possible to add a discourse referent for the actual world, as well as discourse referents for other possible worlds, so that information can be represented about what might have been as well as about what is – just as discourse referents for the current moment and for times in future and past make it possible to record information about what is the case now and what was or will be the case at certain times before and after now.

Most importantly of all perhaps, the information states of PCDRT are not single states – single rows of the value matrix – but the entire matrix, with all its rows. This makes it possible for an information state to capture the

---

allow those discourse referents to occur, like other discourse referents, in the universes of subordinate DRSs, for instance in the universes of the antecedent DRSs K of conditional DRS conditions $K \Rightarrow K'$, then our DRS-language is no longer a version of (perhaps many-sorted) *first* order logic, but of *second* order logic, with the consequence that there can be no sound and complete proof procedure for such a DRS language. For an argument see (Kamp & Reyle 2011), Section 6.1. This is not the first argument to show that even comparatively simple occurrences of plural noun phrases in natural language statements render those statements essentially second order (i.e. irreducible to a first order condition). The moral here is that on the intuitively most natural understanding of the statements in question, a core part of English and other human languages is beyond complete semantic analysis in first order terms. I will have no more to say in these reflections about this aspect of the relations between formal logic and natural language.

[33]The need in DRT for discourse referents representing possible worlds was first made explicit in the dissertation of Roberts (Roberts 1987); Roberts shows that there are multi-sentence discourses in which the first sentence introduces a possible situation or world which then is picked up in the next sentence. Further arguments in favor of possible worlds as 'full citizens of the ontology' can be found in the dissertation of Stone (Stone 1998). As the interest in and work on modal constructions in natural languages has been increasing over the past decades, the view has grown that we cannot do without treating worlds as entities that languages have the means to refer to and quantify over.

full range of possible verifications of sentences with indefinite noun phrases, with the different verifying values for the discourse referent representing the indefinite filling the cells in the column determined by this discourse referent in the different rows. It is then possible in PCDRT to collect all those values into a set that is the value of a new discourse referent for this set, an operation that can be used to deal with many complex cases involving singular and plural anaphoric pronouns, among them the famous Hintikka examples, like:

(13) Every child in the orphanage got a present for Christmas. They opened them right away.

By virtue of their finely articulated structure PCDRT's information states make it possible to define operations on them – so-called 'max'-operations – that transform them into new information states that also have discourse referents for the pronouns *they* and *them* in this example, whose values are the set of all the children in the orphanage and the set of all the presents those children got. It is then also possible to add syntactic operators to the object language of PCDRT – the DRS-language for whose expressions PCDRT provides a relational semantics in which its information states are the input and output states – that represent these operations on information states. The semantics of such operators is given in advance, so to speak, so this is a natural and sound strategy for extending the expressive power of DRS-languages.[34]

One general question that is raised by Dynamic Semantics cum DRT systems is what the respective roles are of their information states on the one hand and their DRS languages on the other. Part of the problem is that much of the work that has been done on and with systems of this kind indicates a primary interest in well-motivated and well-defined logical form languages, and less with the details of how natural language expressions are

---

[34]As far as I can see, there is no obvious limit to how much structure can be incorporated into the information states of Dynamic Semantics systems. For one more example of additional structure, in (Brasoveanu 2011) the states (i.e. the rows) of information states are treated as stacks. That is, the order in which the discourse referents of an information state are arranged is part of the information state's identity. (In other words, information states are not invariant under permutation of their columns.) This extra structure makes it possible to define operations on information states which can be used in semantic analyses of constructions in which pairs of noun phrases express quantifications in which a single quantificational operator binds two variables at once; examples of this are sentence-internal uses of *different* and *same*, as in "Every student chose a different topic" and "Every student chose the same topic", with the readings that the topics chosen by any two students were different or that there was a single topic that all the students chose.

mapped onto logical forms from those languages. It is enough if it is clear
for practical purposes how the logical forms can be obtained for the primary
examples of the constructions of one's primary interest. I assume however
that there is no point of fundamental disagreement here: The developers
and users of these systems would agree that natural language applications of
their systems should include a detailed compositional algorithm for turning
expressions from the relevant natural language fragment into expressions of
the DRS-languages that their systems make available. (There may not be
full agreement about what such an algorithm should be like, but that is a
further matter.)

Let us assume that a satisfactory application of such a system to some natural
language fragment comes with a natural-language-to-logical-form conversion
algorithm for the expressions of this fragment. Let us also assume that it is
also possible to give a 'static' semantics for the DRS language of the system,
which is consistent with the system's input-output semantics in that it makes
the same prediction about truth conditions. (This assumption shouldn't be
taken for granted and would have to be checked for each Dynamic Semantics
cum DRT system; but it holds for many of the systems that have actually
been proposed.) The question is then: 'What if any speaks in favor of the
input-output state semantics as distinct from its static alternative? I per-
sonally do not find it easy to come up with an answer to this question that I
consider satisfactory, but that may just be the reflection of a personal bias.
But here is a suggestion: The central conception underlying Dynamic Se-
mantics is that the use of language in talking about the world – or about
certain parts of it, or alternatively about non-actual worlds, like those con-
structed in works of fiction – results in *imposing on this world or world part
w a language-related structure*. The structures imposed by an expression E
of the object language of the system are the output states that E associates
with the input states it is given.

To make this a little more precise recall that the relational semantics of Dy-
namic Semantics is model-theoretic no less than the static semantics used in
standard versions of DRT. In either case the expressions of the chosen object
language are related to models. It is always a model M from the model
class specified by the system to which the expressions of the system's object
language are related, either via a truth or satisfaction definition (the static
approach) or via an input-output relation (the dynamic approach). In the
latter case an important part of the structure that the output states impose
on the reality represented by M has to do with which entities from M can
appear as values of the same discourse referent of the output information

state.

There is of course more to the process dictated by an expression E when it is let loose on some given input state. In order that it is possible for E to put its structural imprint on the states it outputs M must be such that this imprint is possible. According to all definitions of input-output state relations known to me some of the steps that lead from input to output state are *tests*. These are steps which check whether certain conditions are compatible with the information stored in M. If not, then the transition from input state to output state aborts. (The standard way of representing this in Dynamic Semantics is that the output state in such cases is the impossible information state, a state that cannot be true of any model.) It is through the tests that are dictated by a sentence or discourse E that E shows its truth-conditional content. But when the tests involved in using E to update are compatible with M, then it is possible for E to structure information from E in the manner specified by the notion of information state used by the given DynamicSemantics system. In that case E's output states will take the form of matrices the cells of which contain elements from M or values that are built from elements of M.

As a rule the output states of expressions from the object languages of Dynamic Semantics systems structure only part of the information contained in the models that provide their material. In this regard they are like the situations of Situation Semantics. But a difference is that situations are not assumed to have a specifically linguistic structure. For instance, the situations associated with a model M can be defined as partial sub-models of M. (A partial sub-model of M is determined by some subset U' of the universe U_M of M, but it is not a full sub-model in that it may lack in formation whether an n-tuple $< a_1, .., a_n >$ of elements of U' does or does not satisfy an n-place relation R that is defined in M.) But the partiality is not a reflection of gaps in what E says about M. Its source is not so much linguistic, but rather of a more general epistemic origin, or even a matter of metaphysics: of what it is for the world to contain information.

One of the original motives for Situation Semantics – the semantic properties of so-called *naked infinitive perception sentences*, like (14)

(14)   John saw Mary cycle past.

– lost much of its force once it became widely agreed that verbs are best analyzed as describing events or states and that perception verbs like the

saw of (14) should be analyzed as taking events as direct object arguments. But one way to look at this is as a shift from one new ontological category – that of a situation – to another that of an event structure, or event complex. Event structures have become a topic in their own right within the theory of tense and aspect. One thing I would like to become clearer about than I am is how the various notions of situation, event structure and the information states assumed in different systems of Dynamic Semantics are related to each other. Do we need versions of each of these three categories; or are they the outcomes of different approaches to what are ultimately the same language-related concerns, and thus partly or wholly interchangeable? More about Situation Semantics can be found in Section 7.4.[35]

I conclude this section with an observation about DRS construction. One of the most important insights of generative syntax, which has played a central part in the work of Chomsky since his earliest contributions from the nineteen fifties, is that syntactic structure does not perfectly align with left to right sequencing: The right syntactic parse of a sentence is in general not just a matter of imposing brackets on it as linear sequence of words and morphemes; to go from linear segmentation to syntactic structure certain 'movements' of constituents are often needed. If DRS construction for sentences is understood as the compositional construction of a DRS from a syntactic structure that cannot be expected to perfectly align with left to right segmentation, then such constructions cannot be automatically rewritten as the stepwise building of the DRS as the word+morpheme sequence is traversed form left to right. This is a feature of all syntax-semantics interfaces in DRT that assume syntactic structures for which there is no perfect alignment with left to right segmentation. This is true in particular of any of the DRS construction algorithms that I have ever put mind and hand to.

But from the perspective of actual linguistic processing such non-alignment is a problem – can a compromise be found in that parts of what the left to right processor has already seen is put in some kind of store until its turn comes in the construction of a syntactic or semantic parse? Dynamic Semantic cum DRT systems have addressed at least certain aspects of this question. Their proposals for DRS construction involve alternations between steps which in-

---

[35]Other examples of Dynamic Semantics cum DRT systems are van den Berg's system for dealing with plurals (Van Den Berg 1996), the work of Dekker on discourse referent stacks [**reference**], the work by Vermeulen on Reference Systems (see e.g. (Vermeulen 1994)) or the Partial Compositional Discourse Representation Theory of Haug (Haug 2014). I am debating whether to include something about these systems in a future extension of the present document.

troduce new discourse referents – for instance to represent the contributions made by Determiner Phrases (noun phrases that function as arguments to verbs and other predicates) – and so-called 'tests', steps which verify whether possible values for discourse referents satisfy certain simple or complex predicates. Even if it isn't possible to formulate DRS construction algorithms that consist of such alternations in a strictly left to right fashion, as a series of steps that are triggered as the processor traverses the sentence sequence from left to right, the distinction between the two kinds of steps is nevertheless conceptually and technically importance. The conceptual importance can be summarized as follows: the test steps are what a DRS-based semantics shares with model-theoretic semantics for formal languages like the Predicate calculus and natural languages along the lines of Montague Grammar. The discourse referent introducing steps are those what distinguishes the DRT approach from these more traditional forms of model-theoretic semantics.

What that distinction comes to in more detail can I believe be helpfully illustrated when we make the following special assumptions: (i) The DRS language of our Dynamic Semantic cum DRT system has atomic conditions of the forms '$x = y$', where x and y are discourse referents, and $P^n(x_1, .., x_n)$', where $x_1, ..,$ x_n are discourse referents and P^n is a primitive n-place predicate of the language. (ii) The models M of the system are intensional models $< W, U, F >$, consisting of a (i) a non-empty set W of possible worlds, (ii) a non-empty universe of individuals U and (iii) an interpretation function F, which assigns each primitive n-place predicate P^n for each $w \in W$ as *extension in M at w* a subset $F(P^n, w)$ of U^n. (iii) The information states that the system makes use of are sets of tuples of the form $< w, a_1, .., a_n >$ for some fixed number n. (That is, for any one information state n will be fixed; but for different information states n may vary.) An information state I *relative to* a model M ($= < W, U, F >$) is an information state such that $w \in W$ and $a_1, .., a_n \in U$. Moreover, we consider only constructions of *simple* DRSs, DRS whose DRS conditions are all atomic.

Given these assumptions, the effect a discourse referent-introducing step on an information state I relative to a model M is that each tuple $< w, a_1, .., a_n >$ is replaced by the set of all tuples $< w, a_1, .., a_n, a_{n+1} >$, where a_{n+1} is any element of U. A test step involving an atomic DRS condition $P^n(x_1, .., x_n)$ consists in checking for each $w \in W$ whether $< a_1, .., a_n >$ belongs to $F(P^n, w)$. When this is the case then the tuple $< a_1, .., a_n >$ is retained in the output information state J of the step; when this is not so the tuple is eliminated. (When all tuples $< w, a_1, .., a_n >$ are eliminated for some given w, then this world w is no longer represented within J and therewith no longer a member

of the propositional content determined by J.)

In this special set-up the distinction between the two kinds of steps is particularly clear. Discourse referent introducing steps (also called 'random assignment' steps in the literature) introduce a new discourse referent, as the new n+1st position of the tuples $a_1, .., a_n, a_{n+1}$, without any constraint on its values. Test steps are like instances of the familiar satisfaction clauses for predications involving primitive predicates from traditional model-theoretic semantics. They prune sets of assignments (here of individuals to sets of discourse referents) by throwing out those which do not agree with the extensions specified by the model.

4.2 Other DRT-based Systems

This section is about further approaches that are based on early versions of DRT in one way or another. The theories to be touched upon are:

- UDRT (Underspecified Discourse Representation Theory)

- SDRT (Segmented Discourse Representation Theory)

- Lambda DRT (Λ Discourse Representation Theory)

- LDRT (Layered Discourse Representation Theory)

- PDRT (Presuppositional Discourse Representation Theory)

4.2.1 UDRT

UDRT (Reyle 1993) was developed as DRT's answer to the question what should be understood by *underspecification*, where, roughly speaking, an underspecified representation of some bit of information is a partial description of it, from which some of the information that is meant to be represented is still missing. UDRT's explication of this notion is based on the idea that underspecifications are always relative to some given formalism for representations that are not underspecified. More specifically, given any given DRS language L for full (i.e. not underspecified) representations it is usually possible to develop underspecification formalisms L_{UND} *for* L. L_{UND} will always be an extension of L. Furthermore for each representation R of L_{UND} there is a non-empty set SP(R) of representations from L such that for each R' $\in$

SP(R) R is an underspecification of R'. When R is itself a representation of L, then SP(R) = {R}. When this is not so, i.e. when SP(R) ≠ {R}, then SP(R) has at least two elements.

Underspecified representations from L_{UND} may result from language interpretation when the input is ambiguous in ways that cannot be resolved by the interpretation algorithm for L, or at least that cannot be resolved right away. If a representation R from L_{UND} results in this way (where R is a genuinely underspecified representation, which doesn't belong to L), then this indicates that the interpretation is not yet complete; that it still needs to be completed, by being 'updated' to one of the full representations in SP(R). The value of an underspecification regime is in the method or methods it specifies for replacing R by a member of SP(R) on the basis of information that may become available at some later point. Without such a specification method the regime is incomplete.

A truth-conditional semantics for the representations of L entails a truth-conditional semantics for the representations of L_{UND}, since each R from L_{UND} is truth-conditionally equivalent to the disjunction of the representations in SP(R). This semantics determines, as usual, a relation of logical consequence for L_{UND}. So L_{UND} provides the basis for inferences that can be drawn from its representations in spite of their being underspecified. Such inferences may themselves assist in the updating form underspecified to fully specified representations.

Note that for a given representation language L it is in principle possible to develop any number of different corresponding underspecified representation systems, which deal with different ways in which information needed for a full representation can be missing.

The options that UDRT makes available for representation construction in which underspecified representations can serve as intermediate stages on the way to full representations are orthogonal to all other aspects of semantic representation that are addressed by DRT and the DRT extensions listed above. This means that it is in principle possible to formulate underspecified systems for any of these other systems. One example of this is the system of Underspecified Segmented Discourse Representation Theory proposed in (Schilder 1998).

4.2.2 SDRT

SDRT ((Lascarides & Asher 1993), (Asher & Lascarides 2003) among many other publications). The basic motivation for SDRT is the observation that discourse relations are an inalienable dimension of discourse interpretation: in order that the interpreter can identify a discourse segment as coherent he has to establish a discourse relation between any discourse unit in the segment and at least one other discourse unit, and these discourse relations then become an integral part of the resulting interpretation.[36] Discourse units, in the terminology of SDRT and other theories of discourse coherence, are in the first place the successive sentences of the segment but also the clauses that make up complex sentences, as well as other parts, such as gerunds and certain other complex noun phrases, and quite often also units composed of several sentences. (The definition of what the discourse units of a given segment are is among the tasks for such a theory. It may sound that this concession renders the enterprise circular but in actual fact it doesn't.)

SDRT has been using DRSs as semantic representations of discourse units. In the original and simplest versions of SDRT discourse representations take the form of graph-theoretical structures in which these DRSs are the nodes and the rhetorical relations between them have the status of labeled edges. (In other words, these species are multi-graphs.) In subsequent versions even more complex representational forms have proved desirable, for one thing because some discourse units are proper parts of other discourse units.[37]

One general problem for theories of discourse relations is to account for how discourse relations are recognized: For instance, what is it for one see that the discourse relation between two consecutive discourse units is Causal Explanation (the second discourse unit provides a causal explanation for the first) rather than Narrative Progression (the second unit describes an event that can be understood as naturally following the event described by the first unit)? SDRT has made an unprecedented effort to spell out the logic of the

---

[36]To my knowledge the first compelling case made for the central importance of discourse relations for discourse coherence was made by Rhetorical Structure Theory. See (Mann & Thompson 1987).

[37]In order that a discourse is experienced as coherent the interpreter has to be able to identify a certain minimal set of discourse relations. But the discourse relations recognized by the interpreter often exceed this minimum. One consequence of this is that interpretations of the same discourse or text by two different interpreters may differ in that one interpreter recognizes a discourse relation between two discourse units that the other hasn't, even though both interpreters have identified enough discourse relations to accept the discourse as coherent.

inference processes that interpreters use to identify discourse relations. The effort has been invaluable in revealing, through its very careful probing, how very difficult and complex this problem is.

One of the most important distinctions within the set of discourse relations assumed in SDRT is that between *coordinating* and *subordinating* relations. Discourse units connected by a coordinating relation are situated at the same 'level' of the discourse representation; when discourse units D1 and D2 are connected by a subordinating relation, then D2 is 'below' D1. This distinction implies a further order like structure to SDRT's representations: A discourse unit D occurring as a node in the representation can dominate one or more other nodes, to which it stands in subordinating relations. If D directly dominates two or more discourse units, then those will be at the same level. But each of them can stand in a dominance relation to yet other discourse units. In this way the SDRT representation, of which more and more gets built as the discourse or text is processed, looks much like a steadily growing tree. At each stage this tree structure will have a *right frontier*, consisting of those nodes that are not connected by a coordinating relation to a discourse unit to their right. That is, a node D belongs to the right frontier at this stage if D is not connected by a coordinating relation to a discourse unit to its right (i.e. one that occurs later in the discourse). (But note well, a right frontier can consist of several nodes, so long as these form a chain in which each next node is directly dominated by the node directly preceding it.) An important constraint on discourse coherence is that when a new discourse unit is to be connected via some discourse relation with the representation that has already been built, the connection must be with one of the nodes on the given representation's right frontier. One of the interesting questions that is entailed by this 'right frontier constraint' is whether it also applies to other discourse-related interpretation processes, such as the trans-sentential resolutions of anaphoric noun phrases. (**Question: How much is known about this at present?**)

Discourse relations interact with many other phenomena that manifest themselves at the level of multi-sentence discourse, such as in particular presupposition and anaphora resolution. Asher, Lascarides and a number of other researchers have explored a variety of such interactions within the SDRT framework.

A further challenge for theories of discourse relations is the following distinction between discourse relations. Some discourse relations are or entail relations between the events or states that are described by the discourse

units between which these discourse relations hold. But others are purely 'formal': relations between the forms of two discourse units, like 'parallelism' and 'contrast' (such as, for example, when unit D1 is of the form 'A V-ed B' and D2 of the form C V-ed D). In speech such formal relations often require special prosody, with accents on the contrasting elements A, B, C and D, and especially on the latter two. When such a formal relation holds between two discourse units, this tends to be enough for discourse coherence, and no other relation between the two units needs to be perceived. But formal discourse relations do not contribute to the *semantic content* of the discourse in the way that the event-related ones do, as for instance Causal Explanation, which entails that the event described by D2 cannot have been later in time than the event described by D1. In other words, while these latter relations are relevant at two levels, that of discourse coherence and that of discourse content, the formal discourse relations are only relevant at the level of coherence. This seems another useful application for LDRT. For more on LDRT see below.

4.2.3   Lambda-DRT

Lambda DRT is a version of the Typed Lambda Calculus in which there are types for both discourse referents and for DRSs. Since this version contains, as is standard for versions of the Typed Lambda Calculus, variables for all types, it has variables of the DRS type. These variables can be lambda-abstracted and the resulting lambda abstracts can be turned, in the familiar way, into universal quantifications over variables of this type. Such quantifications are terms of type t and can, like other types of type t, be made into the complements of semantic and attitudinal predicates (like 'true' or 'x believes'). Such predications can in their turn be made parts of DRSs. This makes it possible to build DRSs that make quantificational claims about arbitrary DRSs, and thus involve quantification over domains that may include themselves.

The resulting system has an expressive power that is unlike any of the other formal systems considered in these reflections. Because of its self-referential capacities its semantics has to be stated very carefully so as not to run into the contradictions of paradoxical self-reference right away. But the system has the syntactic means for representing self-reference no less, and as far as I can judge, the danger of paradoxical self-reference remains. This is a system that I would like to understand much better than I do. It does not seem to have been used much for the more homely purposes of analyzing constructions found in natural languages, perhaps because its potentials in a

different direction renders it excessively demanding for the every day needs of the working linguist.

It is important to distinguish Lambda DRT from versions of DRT in which lambda abstraction is permitted over variables of certain types, such as type e or type $< e, t >$. (Kamp et al. 2011) shows that terms formed from DRSs via lambda abstraction over variables of type e is straightforward. (It can also be useful, for instance when some of the higher order commitments of MG (such as the treatment of DPs as of type $<< e, t >, t >$) are to be built into one's DRT language.)

4.2.4 LDRT

In previous sections LDRT (Maier 2013) has been mentioned twice. Section 3 pointed out how LDRT provides a way of representing complexes consisting of an accommodation made towards satisfaction of the identification presupposition of a definite description and of a non-presuppositional part which makes a claim about the description's referent. And in Section 4.2.2 it was suggested that LDRT might be used to express the distinction between discourse relations that contribute to both discourse coherence and truth-conditional content and those relations that contribute to discourse coherence only.

As noted in Section 3 the idea behind LDRT is that linguistic acts and the linguistics forms that they make use of can often be described from different theoretical perspectives. Each perspective is concerned with certain properties that interact with each other in ways this perspective investigates. That doesn't mean that there is no interaction between properties studied by different perspectives, but the perspective-internal interactions tend to be particularly salient can often be profitably studied independently of the interactions between properties belonging to different perspectives, and theories that concentrate on the properties of one perspective while ignoring others are often an obvious and also a good place to start.

A second general feature of linguistic forms and the acts that consist in producing them is that they have a linear structure, in the obvious sense that form and act have a beginning and an end and a well-defined order that leads from beginning to end. In other words, we can dissect each form or speech act into smaller or larger segments. And it is a general (if somewhat imprecisely stated) truth about the different perspectives from which linguistic forms and acts can be studied that each has to start with its own notion

of proper segmentation. The general strategy is then to define the relevant properties for the smallest segments according to the chosen segmentation regime and to explain how relevant properties of longer segments result from the properties of their subsegments.

When it comes to studying the interactions between the properties that were previously only investigated perspective-internally it is most helpful to align the properties that are studied from the different perspectives as they apply to the different segments of given linguistic expressions or their utterances, and in such a way that the alignment is easily surveyable by those interested in perspective-transcending interactions. Aligned data presentations have had a great deal of attention over the past two decades and remarkable advances have been made in this domain. Typically such data presentations consist of multiply annotated left to right presentations of utterance or uttered form with the different annotation levels arranged horizontally and where ideally the users can restrict their view of the data by selecting and juxtaposing those levels in which they are interested. The constraints that such presentation modes are subject to have to do in part with the continuity of the segmented sequences to which the annotations apply. This is so in particular when phonological and phonetic properties are among the data. One of the first tasks of phonetics and phonology is the segmentation of a continuous acoustic flew into phonologically and/or phonetically relevant segments to which the properties that phonological and phonetic can be assigned. Even if this kind of underlying continuity plays no role in the studies inspired by other perspectives (such as syntax and semantics) when these properties are to be aligned with phonological or phonetic ones, they too will have to be assigned to segments of the underlying continua.

It seems reasonable to think of LDRT also as a way of aligning properties from different perspectives. But an important difference is that here continuity is not an issue. The multi-level annotation concerns the different constituents of DRSs, and DRSs are discrete objects from the start, whose smallest 'segments' are their discourse referents and their atomic conditions. In this connection it is convenient to think of DRSs as built sequentially in the manner sketched in the final paragraphs of Section 4.1, with alternating discourse referent introducing steps and tests. The notational device that LDRT adds to DRT to distinguish between different levels is discrete as well. It consists in annotating each constituent of a DRS K with one or more indices, with each index identifying its own 'annotation level'. A constituent annotated with a set $\{i_1, ..., i_k\}$ of such indices is thereby marked as relevant to each of the levels of which the representing index occurs in the set. It

is easy to 'unscramble' the different levels that are identified by these index set annotations of K. For any one index i the representation provided by K at level i consists of the structure obtained by eliminating from K all constituents whose index set does not contain i. A general requirement is that the result of such a culling operation is always a well-formed DRS. For an example consider again the case mentioned in Section 3: The interpretation by the recipient B of A's utterance "I just met the piano teacher of my daughter." This interpretation, we assumed, involves the representation of an accommodation to the effect that A's daughter has a piano teacher, together with the representation of A's assertion that A met this person. In LDRT this complex can be represented as an LDRS consisting of (i) a discourse referent x annotated with the index set $\{1,2\}$, (ii) a set of conditions to the effect that x represents the unique piano teacher of A's daughter, each annotated with $\{1\}$ and (iii) a DRS consisting of presumably several discourse referents and conditions to the effect that A met the individuals represented by x on the day denoted by 'yesterday', each constituent of which is annotated with $\{2\}$. This LDRS can be unscrambled into two DRSs at levels 1 and 2: that of the presupposition accommodation (level 1), consisting of x together with all the constituents annotated with $\{1\}$, and the DRS representing the assertion, which is composed of x and all constituents annotated with $\{2\}$.

To summarize: there are two aspects to LDRT, a conceptual and an implementational one. The conceptual aspect is the explicit acknowledgement that there semantic representations are often composed of parts with distinct status. The implementational aspect consists in adding to DRS constituents annotations with index sets, which mark for each consituent to which level or levels it is making a contribution. This is a simple and yet flexible device for the representation of such multi-status representations, which can capture the binding relations between different parts of such representations whose status may be different. The device, moreover, is simple enough to be used or omitted according to need.

4.2.5 PDRT

Projective Discourse Representation Theory (PDRT, (Venhuizen, Bos, Hendriks & Brouwer 2014), (Venhuizen 2015), (Venhuizen, Bos, Hendriks & Brouwer 2018)) is an extension of DRT that makes use labels for DRS-related items, DRSs, discourse referents and DRS Conditions. In PDRT these are called *projection variables*. The name is meant to convey that they are more than mere labels, which are there just for the sake of being able to refer

to the entities labeled. Rather, projection variables serve to structure the representations to which they belong, PDRT's 'PDRSs'. A PDRS is like a DRS, except that all its constituents come indexed with a projection variable. These projection variables have a double function. On the one hand they can serve as labels to the items to which they are attached, while on the other they impose a partial ordering on those items, which captures among other things the accessibility relations between the constituents of complex DRSs (such as that the antecedent DRS K of a conditional Condition $K \Rightarrow K'$ is accessible from its consequent K', but not conversely, and that both K and K' are subordinate to the DRS to whose Condition set $K \Rightarrow K'$ belongs (and not conversely). But projection variables do more than that. They also serve to mark PDRS constituents as presuppositional as opposed to non-presuppositional; and to capture other information-theoretic distinctions, such as that between assertions and conventional implicatures (Potts 2005), (Simons et al. n.d.).

Part of the partial ordering between the projection variables of a PDRS is imposed during PDRS construction. But the order may then be further constrained in the course of further discourse processing, in which presuppositions are justified or left for accommodation. At that point, the distinction between (justified) presuppositions, assertions and implicatures will be important. For one thing, presuppositions must be recognizable as those parts of the PDRS that require justification as 'old information'; and the distinction between assertions and conventional implicatures is often important insofar as the latter are part of the non-at-issue information (and that therefore aren't possible targets for unqualified denial, as in 'No, that's not true'), while the former are not.

As regards the bits of sentence and discourse processing that PDRT does not take responsibility for, but leaves to some 'pragmatic' component of an overall theory, some care needs to be taken. Information-theoretic distinctions that are explicitly recognizable at the level of the inputs to PDRS construction – the natural language sentences and sentence-sequences, or syntactic analyses for them – may be in danger of getting lost when the only means available to the PDRS to represent it is in terms of the ordering relations between the projection variables that label the different parts of the PDRS.[38]

---

[38]In the version of PDRT I have seen, it might just be possible to distinguish between presuppositional, assertional and conventional implicature parts of a PDRS in that the projection variables i which label the conventional implicature parts are marked as asymmetrically superordinate to the label of the PDRS itself (i.e. '$1 < i$', in case 1 is the label of the entire PDRS). This would seem to be different for the projection variables that label

Indeed, at some places the proponents of PDRT suggest more elaborate ways of labeling parts of PDRSs than 'pure' projection variables can provide. One alternative is to use 'sorted variables', which carry information about, say the information-theoretic types or functions of the items they label. For instance, it would be possible to have 'sorted projection variables' of the form $<i,\text{CI}>$, to indicate that the items within a given PDRS that are labeled with this sorted variable all being to the same propositional representation (that sense in which DRSs are proposition representations) and that this representation is the representation of a conventional implicature. But variable notation in which the variables come with additional information about the status of what they label can also be used for other purposes, e.g. to distinguish between which parts of the information contained in a complex PDRS for an attitude attributing sentence or discourse represent information attributed to the attributee and parts that are solely within the responsibility of the attributor (Venhuizen et al. 2014).

To the extent that the use of such sorted variables is to be seen as part of the PDRT proposal, we should see the PDRT approach as a family of approaches, involving a range of possible 'PDRS languages' which may not only differ from each other in terms of their non-logical vocabulary – i.e. in their stock of non-logical predicates, including some set corresponding to some part of the content vocabulary of English or some other natural language – but also more fundamentally in the types of projection variables they use. This raises important questions about the relationships between PDRT and other extensions of original DRT, including the ones discussed in the earlier parts of this section. For instance, it is natural to ask how PDRT is related to LDRT (see Section 4.2.4), which uses labels to distinguish between representation levels or SDRT (see Section 4.2.2), which uses labels as devices

---

presuppositions. From what I can tell these are only marked as weakly subordinating the label of the sub-PDRS within which they are generated. That is, the only ordering information in the PDRS about the variable i labeling a presupposition would be '$j \leq i$' were j is the label of the sub-PDRS in which the presupposition originates. (Here j can either be the label 1 of the entire PDRS, when the presupposition gets triggered at the 'top-level', or the label of some proper sub-PDRS.) These conditions would then also set presuppositions and conventional implicatures tractably apart from the assertive part of the PDRS. That part would consist of all constituents that bear the label 1 of the PDRS itself, together with all constituents whose labels i are entailed to be subordinate to the PDRS label 1 (i.e. '$i \leq 1$' is a logical consequence of all the ordering constraints that are part of the PDRS together with the assumption that $\leq$ is a weak partial ordering). But one can't help feeling that this is at best a Pyrrhic victory and that when further information-theoretic distinctions more bookkeeping devices will be needed than conditions of the form $j \leq i$ and their negations can provide.

to the representations of Discourse Units (which are taken to be DRS-like representations in the most widely publicized versions of SDRT). It is not really possible to say how much of an overlap there is between PDRT and these other extensions of DRT without looking closely at individual proposals, and without consulting their proponents about details that publicized presentations do not address. But since one of our ultimate goals must be an integrated theory of interpretation and truth conditions that can deal with all the various issues that form the centers of attention in the different DRT extensions (for reasons of practical application as well as for theoretical ones), it will be important to become clearer on this point eventually.

The proponents of PDRT put considerable emphasis on the 'uni-dimensionality' of PDRSs. All the information represented at a PDRS is in some formal sense at the same level, even though the different parts may have different information-theoretic functions within the PDRS as a whole. This point is stressed in particular in connection with the handling of presuppositions. Treating presuppositions at the uni-dimensional level of PDRSs is claimed to be an advantage over the two-level treatment within DRT that goes back to Van Der Sandt (Van Der Sandt 1992), (Van Der Sandt & Geurts 1991), and that has been prominent especially in the work of Geurts (Geurts 1999).[39] But this claim seems somewhat problematic, insofar as at least the mentioned presentations of PDRT leave presupposition justification (the verification of presuppositions in the contexts in which their triggering words or constructions are used, without accommodation) to pragmatics. The importance of that decision gets somewhat underplayed because of the assumption, following Van Der Sandt, that presupposition justification and anaphora resolution are two sides of a single coin, combined with a focus in actual illustrations on anaphoric pronouns, which are resolved by fiat, or presuppositions of definite descriptions, which are treated also as picking up the discourse referents of their anaphoric antecedents. But what are we to do with presupposition triggered by the word *again*, say, or by the words *also* and *too*? For *too*- and *again*-presuppositions it is especially clear that their justification is an essential part of proper interpretation. And often – perhaps always in the case of *too* – the justification can and must be done just in terms of the discourse context. It may be disputed whether even in such cases presupposition is a semantic or a pragmatic process. But what matters for the present discussion is that something more is needed to turn PDRSs that contain presuppositions as parts into what can be considered a proper complete interpretation of the represented utterance of discourse. And that means that PDRSs, which leave

---

[39]See also for instance (Kamp 2001) and especially Ch. 4 of (Kamp 2019b).

presupposition justification to some other theory component, are more like the preliminary DRSs of Van Der Sandt and his followers than to what are regarded complete discourse representations in that tradition.

Note well, this isn't a critique of PDRT as such, only of a particular way of extolling its virtues in comparison with other approaches. In fact, it seems to me that thinking of PDRSs as preliminary representations might fit well with what must de facto be one of their currently most important function. The PDRT format of sentence and discourse representation has been chosen as the representation format in the *Groningen Meaning Bank* (GMD), a large corpus of semantic representations for natural language sentences and sentence sequences. In view of the intended uses of the GMD the representations from it should be and only can be representations of sentences and sentence sequences as *types*, and not of tokens of those sentences and sentence sequences used in some particular context. These representations can only represent information that is in the sentence forms, leaving plug-ins for all that context may have to contribute when the sentence or sequence is used in a felicitous utterance or text. From this point of view it is only right and proper that the PDRS in this corpus should be 'preliminary', at the very least in connection with presuppositions.[40]

In virtue of the GMD (and a sister corpus with PDRSs for sentences and sentence sequences of several different languages) PDRSs have already gained an importance within Computational Linguistics that is unique among DRT-based representation formats. That importance can be expected to grow over the years to come. That is a reason for taking the PDRS way of coding information-theoretic aspects of sentence and discourse content very seriously. And that, I think, is reason all the more to compare its coding options

---

[40]In the version of the Van Der Sandt approach to linguistically triggered presuppositions that I have been using in my own work for the past two decades, in which anaphora is treated as a form of presupposition (rather than the other way round), I have stressed that this way of dealing with presupposition and anaphora realigns DRT with more traditional conceptions of the distinction between semantics and pragmatics: The construction of preliminary sentence representations, with explicit representations for the presuppositions triggered by lexical and constructional triggers in the represented sentence, is part of semantics (and, more specifically, of the syntax-semantics interface). Presupposition *justification*, as distinct from presupposition *representation*, should be classified as part of pragmatics. It is only at that point that context is brought into play. And that is so for the discourse context –even when only the discourse contact is needed – as it is for non-linguistic context, such as for instance the common ground between author and interpreter, as based on earlier interactions between them or on the basis of their shared culture.

(with the help of pure projection variables or with the more expressive sorted variables mentioned above) with other DRT-based proposals, in which additional notation has been introduced to capture other aspects of meaning and use.

5 MSDRT

MSDRT ('Mental State DRT') is another one of the systems that have been developed on the basis of DRT. I have kept it to the last, and am devoting a separate section to it, for two reasons. First, it is a system in which I have myself invested much time over the past three decades, and it is an integral part of my own thoughts about various issues in language, logic and cognition more generally. Secondly, but in the same line, the discussion of aspects of MSDRT will make it easier – and in fact inescapable – to move to those more general considerations to which I would like this document to provoke some reactions, positive or negative.

5.1 Propositional Attitudes and Logical Omniscience

MSDRT goes back to work that started in the eighties. The original impulse was that DRT, whose DRSs were thought to be (or come close to) mental representations, seemed a natural framework for the semantics of sentences that attribute belief and other propositional attitudes: the truth-conditional content of a sentence like 'X believes that p' should be analyzable as a relation between the subject X and the DRS constructed from the complement 'that p'. More precisely, there is a natural way of defining an equivalence relation between DRSs which strictly entails their truth-conditional equivalence. The equivalence classes generated by this relation hold a middle ground between propositions as sets of possible worlds on the one hand and syntactic objects like sentences on the other; they provide a more fine-grained analysis of propositional identity than sets of possible worlds, while at the same time identifying truth-conditional content in a more and logically transparent form than the complement phrases of attitude attributing sentences themselves. Analyzing propositional attitudes as relations between agents and such equivalence classes might therefore be a well-motivated way out of the notorious logical omniscience problem.[41] The first publications in which the idea was

---

[41]The logical omniscience problem: There are cases where p and q are logically equivalent sentences, but where it seems intuitively true that some person X believes that p but false that X believes that q. The analysis of belief as a relation between agents and sets of possible worlds is unable to account for this. Apparently a more finely granulated

worked out that attitudinal verbs can be analyzed as relations between agents and such DRT-based entities were Asher's ((Asher 1986), (Asher 1987)). My own first publication in this connection was (Kamp 1990); in this paper too the logical omniscience problem is the central point of departure.

5.2 Complex and Multi-Sentence Attitude Attributions

The motivation for the next step towards MSDRT was a related but somewhat different one. For decades the discussion of attitude attribution sentences – especially within philosophy, where most of the discussions took place – was focused on an extremely limited repertoire of sentence forms, mostly sentences of the form 'X V-s that p', where V is an attitude verb such as *believe* or *desire* and p is a (mostly fairly simple) complement sentence. From the perspective of the attitude attributions that people actually tend to make this limitation is highly unrealistic. What we are mostly interested in when we talk about the mind of someone else is why it is that this person did or thought what she did do or think. Typically such explanations involve several interacting attitudes of different 'attitudinal modes' – beliefs, desires, intentions, doubts, ... To talk about such complex mental states one obviously needs more than a single sentence with a single attitudinal verb. In fact, when we look at naturally occurring discourses that people use to describe what they take to be going on in the minds of others (or what went on in their own minds at some earlier time, or what is going on in their own mind as they are speaking or writing) we find a wide and open-ended variety of sentence and discourse forms. Inspections of such mind-describing discourses also suggest that their producers picture the mental states they describe as complexes of propositional attitudes, which interact with each other in the various cognitive processes that we sometimes loosely refer to as 'reasoning' or 'deliberation'. The approach of MSDRT is to take this suggestion at face value, analyzing attitude attributing sentences and discourses as descriptions of mental states as such complexes.

The first explicit report of this approach can be found in (Kamp 2003), the German translation of an English manuscript written in the late nineties.[42]

account of the content of the complement sentences of *believe* and other atttudinal verbs is needed.

[42]The publication of the text in German has predictably had the unfortunate effect that it has remained inaccessible to a large part of its intended audience (all the more regrettable given the frst rate and immaculate job by the translator Ulrike Haas Spohn). A version of the original English text will soon be made available on a website with some of my unpublished and older published materials. As things stand, only informal introductions

MSDRT proceeds in two main steps. First, it proposes a formal language L_{MSD} for the description of mental states. ('MSD' is short for 'Mental State Description'). Second, it embeds L_{MSD} in some larger DRS-language, whose choice will depend on what over-all natural language fragment is to be accounted for. This embedding makes use of a predicate '*Att*', which attributes to an agent x a mental state that answers a certain description given in L_{MSD}. That is, *Att* has argument positions for (i) the agent x and (ii) for a description from L_{MSD}. (There are two further argument positions, but these are not essential to the present discussion and will not be discussed here.) The representation of an attitude attributing sentence or discourse takes the form of a DRS Condition whose main predicate is *Att*, with the attributee filling the first argument slot and the MSD description provided by the sentence or discourse that is being interpreted filling the second slot. Note that *Att*-condition can be of unbounded complexity, since there is no upper bound to the complexity of the L_{MSD} descriptions that they take as arguments.[43]

MSDRT is like DRT in that most work goes into defining its construction algorithms. A new challenge are the rules that map attitude attributing constructions from the given natural language fragment into the corresponding *Att*-conditions. One important aspect of this problem is how the MSD descriptions that occur as the second arguments of *Att*-conditions that have already been introduced into the DRS that is being constructed can serve as 'secondary discourse contexts', which guide the interpretation of the attribution made by the current sentence of clause and which the contribution from this clause or sentence will then extend – a form of incremental interpretation that mimics the way that DRT handles the incremental aspects of non-attributing discourse.[44]

---

to MSDRT can be found in published work, in particular in (Kamp 2015) and (Kamp & Bende-Farkas 2019). The term 'MSDRT' isn't used in (Kamp 2003), but was introduced only in the course of the past decade.

[43]Since the embedding of MSD in the chosen DRS language always takes this form – that of adding *Att*-conditions with descriptions from L_{MSD} as second arguments – it is largely independent of what embedding DRS language is chosen. The only requirement is that the embedding language treats verbs and their projections as descriptions of eventualities (events and states) and has the devices needed for locating events and states in time.

[44]See (Stalnaker 1988), (Kamp 1988).

5.3 Referential Noun Phrases in Attitude Attribution Sentences

An important aspect of attitude attributions, which has commanded the attention of logicians and philosophers at least since Quine, is the role of referential expressions occurring in them. (In an attributing sentence of the form 'X V-s that p' these are referential expressions occurring in p, but in less restrictive repertoires of mind-describing sentences and discourses the expressions will also be found in other syntactic environments.) The main problem that philosophers and logicians have been concerned about where it comes to these is the distinction between *de dicto* and *de re* uses. A referring expression is used *de re* when the attributor intends it as her own description of an entity about which the attributee is said to entertain one or more thoughts, irrespective of whether the attributee would be able to recognize or accept the expression as the correct description of that entity. *de dicto* uses (and the corresponding interpretations) are those which the attributor intends as entity descriptions that the attributee himself would recognize as descriptions of the entities he has in mind and that would be an integral part of the thought or thoughts that are being attributed to him and that he himself would use to express the thought of thoughts in words.

One way to explain the difference between *de dicto* and *de re*, and more generally account for the contributions that referential expressions in attitude attributing sentences and discourses make to the content of the attributed attitudes is to assume that the mental state of the attributee doesn't consist just of propositional attitudes but also contains *Entity Representations* ('ERs') for the referents of those referential expressions. As their name makes plain[45], Entity Representations have a referential function: they serve to represent entities that exist outside the mind and typically they do so by virtue of a causal relation in which the agent – the one to whose mental state the Entity Representation belongs – stands to the referent, for instance by currently observing it or on the strength of some earlier visual or other sensory perception. Furthermore, the ERs of a mental state interact with its propositional attitudes in that they can contribute their referents to the contents of those attitudes. And whenever that happens, the content will have

---

[45]In many ways Entity Representations are like the files cards familiar from the work of Perry, Heim, Recanati and others [**References**]. But there are also significant differences between ERs and the roles that file cards play in these other theories. I have adopted the name 'Entity Representation' as a way of guarding against simply assimilating ERs to file cards as they are known from the literature, and thereby overlooking the differences that would be a likely effect of such an assimilation.

the truth conditions of a singular proposition about the referent or referents thus contributed. The distinction between the *de dicto* and the *de re* use of a referential expressions in an attitude attribution can now be understood as the difference between the attribution of an ER which incorporates the descriptive information of the referential expression in the attribution (the *de dicto* use) and the attribution of an ER that does not incorporate this information.[46]

5.4 Non-linguistic Applications of MSDRT

The initial motivation for MSDRT, we noted, was to develop a viable account of the semantics of complex attitude ascriptions. But more or less from the outset the concern was in fact a more general one: to come up with a notion of mental structure that should also be defensible on other, nonlinguistic grounds and that would be helpful in clarifying aspects of cognition that do not involve the use of language; the descriptions provided by L_{MSD} should also enable us to account for issues within the theory of mind, perception and action in which language is not overtly used.[47] As the development of MSD and MSDRT has progressed, the role of ERs has become ever more important, both in the linguistic applications of MSDRT and the non-linguistic applications of MSD. One important function of ERs is that they are crucial to the coherence of mental states that change over time; the ERs are a source of continuity when the propositional attitudes to whose contents they

---

[46]In fact on this analysis of the difference between *de dicto* and *de re* it is most naturally seen as involving three-fold distinction, between (i) 'pure' *de re* uses, in which the attributor chooses her own words to refer to the referent of the relevant ER of the attributee; (ii) uses of a referential expression with the intention to refer to the referent of the ER, but in such a way that the attributee can recognize the expression as referring to the referent of her ER (and where she might herself have used the attributor's expression to describe the attitude attributed to her; such cases are still *de re* from the perspective of truth-conditional semantics: the attributed propositional content is still that of a singular proposition about the referent the ER); and (iii) 'pure' *de dicto* uses, which the attributor assumes the descriptive content of the referential expression she is using to be an integral part of the thought she is attributing; the content of this thought will *not* be a singular proposition about the description's referent, much in the spirit of Russell's Theory of Descriptions; in fact there may not even be a referent.

[47]From the start the examples that served as paradigms in development of L_{MSD} were L_{MSD} descriptions of parts of mental states that were the result of visual perception. Another area of non-linguistic application of L_{MSD} is the formation of plans and intentions-to-act and the execution of plans and intentions, though that concern is from a somewhat later date (end of the nineteen nineties; see the still unpublished (Kamp 1999)). It was also in connection with such non-linguistic applications that the need for ERs made itself felt first.

contribute their referents come and go, or change their mode or content once they have been introduced into the state.

5.5 A Communication-theoretic Approach to Meaning and Use

The linguistic applications of MSDRT, and of the development of L_{MSD} as part of it, that have been mentioned so far, concern the semantics of attitude attributions. But there is also a very different use that can be made of L_{MSD} in natural language semantics. L_{MSD} can also be used in a *communication-theoretic* approach to the use and structure of language in general. The communication-theoretic approach sees language in the first instance as a tool for communication: A speaker S has a certain thought that she wants to communicate to a hearer H. So she uses the language L that she and H share to find a sentence (or a sequence consisting of several sentences) that expresses her thought and communicates this sentence or sequence to H. H, who shares S's knowledge of L, can make use of that knowledge to reconstruct the thought she is trying to communicate to him from the words she is using. When all this works the way it should, then the result will be that H comes to share that thought.

As an informal description of what goes on in verbal communication there is nothing in this that would strike anybody as surprising or interesting. But the description can be taken as a challenge. The challenge is to make the idea work in detail, by describing precisely how thoughts are converted into language by speakers or authors and then reconverted into thoughts by their listeners or readers. That will be possible only when we have ways of identifying the contents of thoughts independently of how they are expressed in the public communication language L in which speakers communicate with each other. It is here that L_{MSD}, as a language for the description of the contents of mental states, can be usefully brought into play. First samples of how the mental state descriptions of L_{MSD} can serve to account for what goes on in the transfer of referential intentions can be found in the mentioned papers (Kamp 2015) and (Kamp & Bende-Farkas 2019), which show how proper names and epistemically specific indefinites function as go-betweens between Entity Representations in the mind of the speaker/author and co-referring Entity Representations in the mind of the hearer/reader. But this is no more than a very first beginning, What we want is a model of how *propositional contents* are put into words by the producer, and then turned back into representations of propositional content by the recipient. In those

cases where the utterance contains a referentially used expression (such as a proper name or a specific indefinite), the correlated ERs in the minds of producer and recipient will occupy the same argument positions in the producer's thought and the recipient's reconstruction of it, but that is just one small part of all that has happens in successful verbal communication, and that a communication-theoretic theory along the indicated lines should be able to deal with.

At the current juncture the main obstacle to a communication-based theory of natural language use and meaning is what happens on the side of the producer. Theories of what happens on the side of the recipient are in the advantageous position that it is reasonably clear what the *inputs* are to the processes they must describe: The inputs are the spoken or written utterances that reach the recipient. It may not be clear in which form we must assume these inputs are given – as continuous flows of acoustic material, as sequences of phonemes, or sequences of words and morphemes, or as sequences of signs on a page, to mention the main options – but these question too are questions about an empirical reality that is independent of any theoretical parti pris and can be observed and studied from different theoretical perspectives. In formal semantics to make quite specific assumptions about the inputs to the processes that are to be described. A common assumption is that inputs come in the form of syntactically parsed sentence sequences. Opinions may vary as regards what these syntactic parses are like (i.e. which theories of syntax they implement), but by and large theories of semantic interpretation can build on a solid and precise notion of the entities whose interpretation are to be described.

On the side of language production the situation is very different. There are settings in which the question what are the input to production are clear, for instance natural language driven question answering systems which translate questions put to them in natural language into some internal representation language, compute the answer to this translation with the help of their knowledge base and must generate a natural language answer from the internal representation of the answer they have computed. Here generation is finding a suitable, content-preserving natural language equivalent of the internal representation. But natural language production by human users who speak or write cannot rely on such independently given internal representations. Here we are facing the problem: What is it that the user is generating his spoken or written utterance from?

For those who believe in DRSs as mental representations a natural assump-

75

tion for the case of human language production is that the inputs to production are the very same structures that DRT posits as the outputs of interpretation, in other words: DRSs. In that case language production might be treated as the inverse of language interpretation as conceived and described in DRT – the conversion of inputs in the form of syntactic parses into DRSs. Inverting the DRS construction process into one that turns DRSs into bits of natural language will then be fairly straightforward, so long as we have a way of dealing with the fact that the conversion of natural language into DRSs is intuitively a many-one process – in any version of DRT known to me the process is many-to-one – for which there isn't a well-defined unique inverse. But so long as one is happy with a natural language generation algorithm which picks when applied to some given DRS one of the possible inputs that the interpretation algorithm converts into that DRS, building such a generation algorithm – one that turns each DRS for which it is defined into a sentence or discourse that the interpretation algorithm turns back into that DRS – is comparatively straightforward. (For an early implementation of this approach see (Asher & Wada 1988).)

One problem with production accounts based on this assumption is that the evidence for DRSs as generation inputs is that the evidence for it is very indirect: *if* it is assumed that semantic interpretation takes the form of constructing DRSs a mental representations of content and *if* it is assumed that interpretation is the converse of production, then it follows that DRSs are the inputs to production. But these are two very big *if*'s. A further problem for such an approach to production is that it seems to offer no place for the complex issues of utterance and discourse planning, the study of which has become a field in its own right. How could the insight about discourse planning be combined with a theory of production in which DRSs are converted into language that the interpreter will convert back into such DRSs? Until we think we have a reasonable strategy for dealing with this issue, the assumption that DRSs are the inputs to generation should be qualified as dubious.[48]

---

[48]Not that there are no serious proposals to deal with this problem. One study that should be mentioned here is a study of the narratives produced by speakers who are asked to retell short episodes form (animation) films in their own words (Jasinskaja & Rossdeutscher 2012). In this study it is assumed that watching the episode leads to a DRS-like representation in the observer, and that the planning needed to arrive at the words of the retelling uses this DRS as starting point. An alternative approach would be to treat the DRS that gets converted into a natural language utterance as the *result* of a planning process, perhaps in the spirit of Slobin's 'thinking for speaking' (Slobin 1987), and then the input to the conversion into natural language as final part of the generation process. But to my knowledge this second possibility hasn't been seriously pursued; and I

These doubts about DRSs as inputs to generation apply equally to the assumption that language production involves L_{MSD} descriptions as inputs. There may be good reasons for assuming that L_{MSD} descriptions are better candidates than plain DRSs as the outputs of interpretation – this is what MSDRT-based accounts of noun phrase reference might be taken to imply. But that is not the same thing as assuming that language generation proceeds from such descriptions as inputs. The transition from output of interpretation to input to generation seems hardly less fraught with hazards and undischarged theoretical commitments than it is for plain DRSs. And that being so, is that the end of MSDRT as framework for a communication-theoretic account of language? Or should we, at a minimum, lay any attempts to develop such a theory on ice until the fundamental issues raised in the last paragraph have been cleared?

Another observation that casts doubt on the possibility of a general account of language production in which either DRSs or L_{MSD} descriptions are assumed as inputs is that it isn't clear that language production always takes its departure from a well-defined input. There is a strong intuition, which I believe I share with many, that often our thoughts become clearly defined only as we are trying ot put them into words – language production isn't just conversion of well-defined contents into a different medium, it is the actual *formation* of those contents. If that is right, then a general production account along contemplated lines is misconceived from the start, and that irrespective of whether the inputs it assumes are DRSs, L_{MSD} descriptions or anything else.

These various considerations notwithstanding I believe the prospects for such a communication-theoretic account aren'r nearly as bad as they suggest. The reason, related to the point of the last paragraph, is this. It may often be true that we do not know what we are going to say or write when we start. But we do know at the end. This is true most obviously for when we write. When I have written something I will usually look it over before I make it available to someone else. (Disaster looms if I don't.) It is only if and when what I have written looks right to me – when I feel that this is what I want to say – that I am prepared to share it with others. But what then is it that I share with those to whom I have made available what I have written? My answer to that question, and the central claim I am putting forward in

---

do not know that the approach of (Jasinskaja & Rossdeutscher 2012) has been extended beyond the quite specific experimental setting that this and some related papers address.

support of a communication-theoretic approach, is that the thought I have and my audience have come to share is the interpretation that each of us gets when he interprets my writing in the way in which competent speakers of my language interpret its utterances and texts. My communication has been successful when and to the extent that the interpretations that have been constructed from what I have written by the person or people that make up my audience match the interpretation that I myself obtain from it. (Exactly what 'matching' comes to is something that needs to be carefully spelled out of course, but as a first approximation we may posit that matching must at the very least entail truth-conditional equivalence)

What goes for writing also goes, I want to suggest, for speaking. Here the case may not be quite as intuitively compelling. But (a) there certainly are cases where we know the moment we produce a spoken utterance that this isn't, or isn't quite, what we wanted to say. And often, we correct our words, sometimes even before we have got to the end of our sentence, but if not then we start all over, perhaps marking what we are doing by something like 'well, that isn't what I wanted to say. What I mean is ..'. But the general proposal I want to make is a more general one: that *always* when we speak or write a checking mechanism is at work, which subjects our own written or spoken output to an interpretation process that we also rely on when we interpret the speech or writing of others. Mostly this mechanism works in the background and we proceed as if it wasn't there. The interpretation that it outputs of the words we are speaking or writing produces no dissonance with what we want to express (whatever that might amount to in utterance-independent terms). In such cases it will seem to us that just this is what we wanted to express. But the mechanism will have been at work no less; it is just that on these occasions it has no grounds for complaints.

If this is right, then communication-theoretic success is just a matter of matching *interpretations*, the interpretations constructed by the audience and the 'self-interpretation' constructed by the producer. And a communication-theoretic account of language, in which success is defined along these lines, is now a genuine possibility. That is true in particular for the kind of account I have been pleading for, in which language meaning as grounded in language use are analyzed in terms of how thoughts are transferred from producer to recipient. Note well that this does not get us back to square one, in which the mental state that the producer is in when production starts no longer plays a role. The MSDRT-based accounts of how a speaker will choose noun phrases for the entities she wants to talk about on the basis of the ERs she has for those entities remains a valid component of such an account, just as the role

of the ERs in the mind of the recipient when he endeavors to determine what the referents are of the noun phrases she has chosen. But this no longer forces us to assume that a speaker must start with a full-fledged representation of what she wants to say – not even one that determines its truth conditions. It is only in those cases where the speaker does start from a well-defined thought represented as part of her mental state that the production side of the account should be able to tell us how this internal representation gets converted into words.

On such a conception of a communication-theoretic theory of language meaning and use we need to assume that the mental states of producer and recipient contain Entity Representations for the entities to which the producer refers and the recipient must be able to recover.[49] Among these ERs there must be in particular those that represent entities referred to by discourse-new definite noun phrases. This means that such a theory must assume Articulated Contexts – or at a minimum substantial portions of them – as parts of the mental states of the communication participants. But if Articulated Contexts are supposed to play those roles, then that can be only as parts of mental states. Is that then what a theory of reference that makes use of Articulated Contexts must inevitably come to in the end? Some observations about this questions will be discussed in the next section.

5.6   MSDRT and Articulated Contexts

In Section 2.2 it was noted that the familiarity account of definite noun phrases requires a richer notion of context than the mere Discourse Contexts of File Change Semantics and DRT. One way out of this conundrum, it was then pointed out, is the adoption of Articulated Contexts (henceforth: 'ACs'), which contain information from other sources besides the current discourse. Little was said about the details of a semantics of definites based on Articulated Contexts, but these details aren't needed for the discussion that follows below. However, there is one fundamental issue concerning the role of ACs to which we can no longer close our eyes, now that an outline has been presented of the motives for and workings of the MSDRT framework. In the communication-theoretic framework described in the last section, language interpretation is a mental process, which the processor, the recipient of a linguistic input, must carry out on the basis of the linguistic input he receives and *information that is available to him as interpreter*, i.e. information that

---

[49]Recall the cited documents (Kamp 2015), (Kamp & Bende-Farkas 2019) and (Kamp 2019*a*).

is stored as part of his mental state, and that he is able to retrieve from storage for his current interpretational needs. This applies in particular to the interpretation of definite noun phrases, including occurrences of discourse-new definite descriptions and other discourse-new definite NPs (e.g. proper names). Mostly interpreters do not have any serious problems with such occurrences because they have the information needed for the interpretation, in the form of an ER that they can recognize as representing an entity that satisfies the descriptive content of the description, or an ER that carries the used proper name as a label (Kamp 2015). If we want to re-describe such interpretation events in terms of ACs, then, it would seem, ACs should be interpreted as mental categories.

But how? To say something about this we need to say more about the components of an AC. Besides the Discourse Context component – which for present purposes we can assume is like the Discourse Contexts of the original versions of DRT, viz. a DRS – an Articulated Context has three further components, (ii) the *Encyclopedic Context*, (iii) the *Generic Context* and (iv) the *Environment Context*. The Generic Context component is a general repository of 'generic world knowledge' – strict and rough regularities that govern the ways in which our world functions. This is the least well developed part of the theory of Articulated Contexts. It is only of indirect relevance to the present discussion and I will say no more about it. But something should be said about the two remaining components, the Encyclopedic Context and the Environment Context.

As mentioned in passing in Section 2.2, the Encyclopedic Context component consists of 'entity representations'; and the same is true of the Environment Context. What precisely entity representations might be on a non-mental understanding of Articulated Contexts may be a matter of debate. But when the notion of an AC is to be understood in mental terms, then the natural way (and, it seems, the only plausible one) of identifying their entity representations is as the Entity Representations that according to MSDRT are among the constituents of mental states. In fact, the Encyclopedic Context of (the AC of) an interpreter H may be identified with the set of all ERs of H's mental state at the time of interpretation, and the Encyclopedic Context of (the AC of) a speaker S with the set of all ERs of S's mental state at the time of production.[50] The Environment Context is a more restricted set of

---

[50] The interpretation of definite noun phrases will sometimes also depend on episodic knowledge about one or more of the referents represented by ERs belonging to the Encyclopedic Context. Given the structure that is assumed for ERs, which offers room for descriptive information about the represented referents, such episodic information could

ERs. This context component is relevant only for the interpretation of 'deictic' uses of demonstrative noun phrases and pronouns as part of utterances that are made face to face. In such a situation the ERs of the Environment Context are those that represent entities in the environment in which the exchange is taking place and to which H has sensory access (for instance, and most typically, when H can see the referent. For details see (Kamp 2019a).)

By identifying the entity representations of ACs with ERs belonging to the mental states postulated by MSDRT one step has been made towards defining ACs as parts of such mental states. More is needed for such a complete reduction, having to do with the Generic Context components of ACs (and perhaps also with propositional information belonging to the Encyclopedic component, see footnote 50). But let us assume that this can be done.

With such a complete definition of ACs as parts of mental states à la MSDRT we have the possibility of recasting existing semantic accounts that make use of ACs as accounts formulated within MSDRT. To do this we need to identify the ACs referred to in the existing accounts as ACs that are parts of the mental states of interpreters at the time of interpretation and ACs that are parts of the mental states of of producers at the time of production. But should we see such a reduction as full replacements of the existing theories, which can now be dismissed as preliminary versions that have done their service and can now be discarded? That is not an easy question. It is a question that can be seen as a special case of a much more general issue, which comes up in many branches of linguistics, phonology, morphology, syntax and semantics among them. The issue is one that is distinctive of language, and for the following reason. On the one hand, language is a tool for human communication, and it can and should be studied in that capacity, by analyzing how it is and can be used, when it is used correctly and how its uses relate to the situations in which they are made. The communication-theoretic approach based on MSDRT sketched in the last section aims to study language from this perspective. But on the other hand, communication by means of language couldn't work if the users of a language didn't *share* their language – if they didn't all know the language's 'Grammar'. It is this sharing, the knowing

---

be incorporated into one or more of the entity representations whose referents are involved in this information. An alternative is that at least some of this information is given separately, e.g. in the form of DRSs with discourse referents linked to the relevant ERs. No choice between these options has been made in published and submitted work on Articulated Contexts. The choice is left open here too. How it is made does not affect what will be said in the remainder of this section. A similar qualification is arguably needed for the Environment Context component.

and applying of the same grammar rules, that makes communicative success a possibility; and not only that, it makes communicative success something like the default case, the one that communcication participants will take for granted so long as they have no reason to suspect that something has gone wrong. It is this inter-user stability of human languages, which manifests itself in like treatments of the same utterances by different users, that makes it both possible and desirable to study languages as user-independent systems, with their own autonomously fixed principles of form and content. Many linguists see it is their primary task to study languages from this angle, as autonomous symbolic systems. In particular this is the overwhelmingly dominant view held within the formal semantics community.

This methodological commitment of modern linguistics goes back at least to the Structuralism,as first formulated and advocated by Saussure and others. It has proved immensely productive and fruitful, and the successes of formal semantics are without doubt among the most important and lasting ones to which the commitment has led. Most of the work discussed in these reflections falls within this category (including DRT, so long as it is seen and practiced as a theory that studies properties of the language as user-independent system). The commitment doesn't exclude or prohibit the study of language as a tool of communication – I doubt that anybody ever thought that; at most the commitment has served as a way of what is the task of linguistics, as opposed to other scientific disciplines that have something to do with language. But views seem to differ about how the two approaches can be connected. The classical view of formal semantics, usually associated with the name of Charles Morris (and the one I was brought up with as a graduate student at UCLA) has it that the main components of a complete theory of language are syntax, semantics and pragmatics, with semantics presupposing and building on syntax and pragmatics presupposing and building on syntax and semantics. Syntax and semantics study properties that languages have qua autonomous abstract systems, whereas pragmatics is concerned with the study of language use. A crucial aspect of this way of seeing the study of language is that the syntactic and semantics properties of a language as autonomous system can be studied exhaustively before – and thus without – taking any questions into account about use. It is on this view of the relation between user-independent and use-related aspect of language that opinions may and do differ.

One reason for challenging the Morris hierarchy has to do with the role played in the determination of meaning by noun phrase reference. It is widely assumed, and consistent with Morris' hierarchical conception of syntax, se-

mantics and pragmatics, that the central task of semantics is to deliver the propositional contents of sentences: It is these that pragmatics, as the theory of language use, can freely make use of, by describing the various things that users do with the propositions expressed by the sentences they produce or receive. One problem with this picture, recognized in the very early days of formal semantics, are the contributions that are made to propositional content by referential noun phrases. The problem arises in somewhat different forms for different types of definite noun phrases, but it arises in a particularly striking form – a particularly virulent one, you might say – for those noun phrases that Kaplan refers to as 'demonstratives'.[51] Kaplan distinguishes two types of 'demonstratives', (a) indexicals, like the pronouns *I* and *you*, and (b) demonstrative phrases, which in English consist of or begin with *this* or *that*[52] What Kaplan and others have seen as uniting the two types of phrases, and justifies their classification as the two subtypes of demonstratives in Kaplan's sense, is that they both impart *singular content* to the sentences in which they occur: sentences containing occurrences of Kaplanian demonstratives express propositions that are about the referents of the demonstrative phrases they contain (and thus are, in current terminology, *singular with respect to* those referents). The problem that sentences with demonstratives present for Morris' conception is that the propositions they express depend on the referents of of their demonstratives and that these referents are determined by factors that have to do with use, and not just with form. (For the indexical *I* the referent is the one who utters the sentence containing it, for a demonstrative noun phrase the identity of its referent has to do with what the utterer is pointing at and so on.) So the propositions expressed by utterances of sentences with demonstratives depend on use as well.

Kaplan's answer to this predicament is his distinction between *character* and *content*. The character of a sentence S can be thought of as a kind of blueprint for a proposition, a structure that becomes a proposition when its open slots are filled with the referents that are determined by the different contexts in which the sentence can be used. (In Kaplan's functional implementation of the idea, this takes the form of treating characters as functions that can be applied to the contexts that determine the referents of its demonstratives and yield as values the singular propositions that result when these referents are inserted into the 'slots' in the character.) But note that combining characters and contexts is an operation that involves aspects of use and that thus

---

[51]See in particular (Kaplan 1989).

[52]This of course only covers the singular demonstratives; the determiners of English plural demonstrative phrases are *these* or *those*. But Kaplan wasn't concerned with plurals and the present remarks won't be either.

must be counted as part of pragmatics. If semantics is taken in the strict sense that Morris' view of the organization of linguistic theory entails, then it is only the relation between expressions and their characters that belongs to semantics; the transition from character to content is not. In other words, the relation between sentences with demonstratives and the propositions determined by them isn't just a matter of semantics in Morris' sense.

Kaplan emphasizes what indexicals and deictically used demonstratives[53] have in common. But the differences between indexicals and demonstrative noun phrases are just as important. One aspect of the difference has to do with the resources needed to determine the referent. The reference of a deictically used demonstrative phrase is determined via some kind of causal relationship between the referent and both speaker and interpreter. (The Environment Context components of Articulated Contexts serve to make this relationship explicit. For details see (Kamp 2019a)). For the indexicals *I* and *you* this is not so. Here the reference is just a matter of who the utterer is and to whom the utterance is addressed.

Why is the difference between indexicals and demonstrative phrases important in a discussion of MSDRT? (Kamp 2019a) starts out with a presentation of Kaplan's account of indexicals that amounts by and large to an endorsement of the notion that all we need to account for the meaning contributions that indexicals make is a user-independent notion of utterance context, as a tuple consisting of a 'speaker', an 'addressee', a 'time' and perhaps some further items. When a sentence with one or more indexicals is actually used then this use will instantiate this user-neutral concept of an utterance context, as a tuple consisting of the actual speaker, the actual addressee and so on, and this will then determine the proposition expressed by the sentence on that occasion. But the main burden of a theory of indexicals can nevertheless be adequately discharged relying only on the first, abstract notion of an utterance context. For an account of the reference of demonstrative phrases the cards are mixed differently, (Kamp 2019a) suggests. Here, the paper asserts, the analysis cannot do without a notion of Entity Representation that is perceptually linked to its referent (which must be present in the environment shared by speaker and audience, and perceptually accessible to all participants); and the only way of make sense of these Entity Representations is as constituents of the participants' mental states.

---

[53]Kaplan only considers the deictic uses of demonstrative noun phrases, those that refer to objects in the environment. Demonstrative noun phrases also have anaphoric uses, which refer to some element of the discourse context. These are important for a general theory of noun phrase anaphora but are to be set aside for the present discussion.

The final conclusion of (Kamp 2019a) is – not surprisingly after what has preceded – that a satisfactory account of demonstrative reference is in last analysis only possible within a framework like MSDRT, in which the mental states of speaker and audience play an essential part. But in the meantime I have become less certain of this conclusion. I think that in spite of the arguments made in the second part of the paper there remains a place for an abstract theory of how demonstrative noun phrases refer, in which contexts take the form of ACs, but in which there is no explicit reference to the minds of language users. Such an account may be in a position to bring out certain aspects of demonstrative reference more clearly than one in which production and interpretation are described as mental processes. And that is justification enough for retaining it, side by side with the mentalistic version that results form the interpetation of ACs as parts of mental states.

This plea for theoretical pluralism and the reason for it are a special instance of very general considerations of scientific methodology. When a scientific domain is of great complexity, a single comprehensive description of it will overshoot, and by a wide margin, the capacities of human understanding. What we human scholars and scientists need are narrowly circumscribed theories, which deal with limited sets of related phenomena and set aside everything else as much as possible. In some way or other it must be possible to fit all these narrowly focused 'local' theories together into an over-all theory within which all the phenomena of the given scientific domain can find their place. How to do that can be a major challenge, as when the task is to unify quantum mechanics with relativity theory or in constructing a unified theory of the four fundamental forces. But even where such unification efforts are successful, the unified theories retain their own individual legitimacy, as distinct pieces of a larger puzzle, each of which has its own contribution to make – in science education, as a basis for further theoretical explorations or various applications.

It is an interesting fact about our world that it should be possible for us to chart large phenomenal parts of it in this kind of piecemeal fashion, even if it isn't always easy to distinguish between the extent to which the world invites such a modular approach and our own methodological needs and predilections for it. But whatever the ultimately right apportioning may be between the structure of the world as such and our wanting to impose such a structure on it, the structure and use of language appears to be one of the large domains of related phenomena for which the modular approach is clearly appropriate. It has worked so well, it seems, that we can hardly imagine how

else the study of language could proceed, with modularization operating at many different levels.

One part of this complex modular structure to which the different branches of linguistics each make their own contributions is the theory of linguistic meaning, in which meaning is studied as the joint product of formal structure and aspects of use. This part is a complex modular structure in its own right, and it is within this modular structure that theories of reference occupy their niche. That is so in particular for the theory of definite noun phrase reference which uses Articulated Contexts as abstract structures without treating them explicitly as parts of the structures of participants' minds.

But what is the proper niche for a reference theory based on abstract Articulated Contexts? Here we must return to our remarks about the Morris conception of syntax, semantics and pragmatics. We noted that the reference theory in question must be counted as belonging to pragmatics because the context plays a non-eliminable role – it is *utterances* of sentences containing Kaplanian demonstratives that are the carriers of propositions, not the sentences as such. But on the other hand, these are among the propositions that should be made available as inputs to modules belonging to pragmatics. What that comes to in greater detail is a question which in the early days of formal semantics would have been answered very differently than it can be answered today. In the fifty years between, linguistic pragmatics has developed into an ever more complex modular structure of its own, with several different branches with its own cluster of formal and semi-formal modules – among them theories of Gricean Implicature, Information Structure, Question Answering, Rhetorical Structure and discourse coherence, Speech Act Types, or Turn Taking in Dialogue. In all those branches the dominant practice is to operate with the propositions that are expressed by the sentences used in the linguistic acts with which the theory is dealing with for granted. 'Semantics' is assumed to be responsible for how those propositions are determined by sentence form and the way the sentence is being used, and to deliver those propositions as needed.

But if that practice is legitimate, and if the propositions needed by the practice include those that are expressed in utterances of sentences with demonstrative noun phrases, then theories of use-dependent reference do not seem to fit on either side of the semantics-pragmatics divide; and that would be so in particular for the theory of reference based on abstract Articulated Contexts that has been at the center of our present discussion. Such theories seem to be holding a middle ground, or sitting akimbo across the divide, if

that is the image you prefer. But however we may want to picture or phrase this, the position of these theories within the complex network of meaning-related modules is situated in between semantics, as a theory of a purely form-driven syntax-semantics interface and pragmatics, as the multitude of theories that treat propositions as given inputs. And that is something it is important to keep in mind.

In short, if we want to hold on to the distinction between semantics and pragmatics as it tends to be understood and implemented in the theory of linguistic meaning and use today, then we cannot stick to the idea that we are dealing with a division that is strictly binary. We have to recognize that besides these two main categories there also is the in-between category of use-dependent theories of use-dependent propositional content.

5.7   Linguistic and Non-Linguistic Acts

The MSDRT-based communication-theoretic approach presupposes interactions between the acts of linguistic communication that are its direct targets and certain non-linguistic acts. Such interactions are prominent in what is said in (Kamp 2019a) about the interpretation of deictic demonstratives. As pointed out in this essay, the effect of a deictically used demonstrative noun phrase is often that it draws the addressee's attention to the referent of the noun phrase, for instance, by pointing at the referent and thereby making the addressee look at it. The addressee's perception of the referent will first lead him to introduce an ER into his mental state to represent the referent he is perceiving, and then use this ER as representation for the referent of the demonstrative noun phrase which sets the whole process in motion. The result is an ER that is anchored to its referent twice over – via the addressee's visual perception and qua referent of the demonstrative noun phrase.

Interpretations of demonstrative phrases that lead to such doubly anchored ERs are one type of case where the use of a certain type of expression involves a complex of acts some of which we would normally classify as linguistic and some as non-linguistic (as in our example the perception of the demonstrated referent). But including non-linguistic acts in an analysis that targets certain linguistic acts can also be necessary or advantageous for other reasons. For instance, the linguistic act or acts involved in extracting information from what one reads or is being told by others is often followed by using the extracted information in acts of reasoning, including perhaps the performance of linguistic response to the linguistic input received. A careful analysis of those reasoning processes may yield some insights about the form of the

information representation to which the extraction process has led, since after all it is likely that this representational form is exploited in the reasoning.

I am firmly persuaded that such 'mixed' accounts, in which linguistic processes are treated in conjunction with non-linguistic ones, are a crucially important direction that linguistics should and will pursue in the years ahead. (I doubt that within what remains of my lifetime there will be much work of this kind that I will see, but I have been wrong with similar prognoses before.) Further speculations about interactions linguistic and non-linguistic acts will come up in the course of the next and final Section 6. I conclude the present section with an all too brief description of one recent approach in which the interleaving of linguistic and non-linguistic acts is essential, albeit for different reasons than the ones I have hinted at so far.

The work I am referring to is ongoing research by Adrian Brasoveanu and Jakub Dotlačil in their forthcoming book *Computational Cognitive Modeling and Linguistic Theory* (Brasoveanu & Dotlačil forthcoming). This book uses the detailed model of mental architecture and mental processing known as ACT-R ('Adaptively Controlled Thought - Rational'; see (Anderson, Bothell, Byrne, Douglass & Lebiere 2004)) to analyze the processes involved in tasks performed by the subjects of certain psycho-linguistic experiments. The last and most complex experiments discussed in the book concern the on-line processing of sentences in self-paced reading. The authors analyze executions of the tasks in question as sequences of processing steps in which steps of syntactic and semantic processing alternate with non-linguistic actions – actions by the subject's motor system (tapping the space bar for the next word to appear on the screen) and perceptions (locating and identifying the next word when it appears on the screen). The computational models that the authors develop of these sequences of processing steps are detailed and explicit enough to enable it to make precise predictions about how much time is needed to execute each of the steps. (How much time will of course vary, as is expected for typical self-paced reading experiments, for the successive words that make up the sentences from the modeled experiments.) These predictions from the computational model can then be matched against the data that have been obtained from the corresponding psycho-linguistic experiments.

A crucial feature of the models developed in (Brasoveanu & Dotlačil forthcoming) is that they are specified in the form of computer programs, which can be run as computer simulations of the modeled task performances, with as output predictions of how much time is needed for the successive steps

that experimental subjects will need for the sentences they are made to go through in the modeled experiments. To my knowledge the attention to the quantifiable details of linguistic performance that we find in this work is novel to linguistics, and the demonstration that the approach can actually be made to work a major achievement. And it is radically different from everything else that has been discussed or mentioned in the present reflections up to this point.

The distance between the models of specific linguistic performance tasks developed in (Brasoveanu & Dotla*č*il forthcoming) and the theories that gave rise to the speculations in the first three paragraphs of this section seems to be a very big one. The abstract reconstruction proposed in (Kamp 2019*a*) of the interpretation of deictically used demonstrative noun phrases has nothing to say about how much time will be taken up by the different steps presumed in the reconstruction, and it is plain that much more details are required before there can be any hope of making meaningful estimates of the amounts of time needed to perform these steps (for instance, details about where the referent is located in relation to the addressee). And accounts that combine acts of interpretation with subsequent acts of subsequent reasoning with the acquired information seem to be even further from the level of detail needed to make precise quantified predictions. But common between the Brasoveanu & Dotlacil models and the speculations in the first half of this section is, I repeat, the development of theories of processes in which linguistic and non-linguistic acts are intertwined. And – this too is a repeat, but one well worth making – my guess is that such mixed theories are going to play an increasingly important part in future studies of use and meaning.

6   Back to Logic

Towards the end of Section 1 I said a few words about how much my understanding of the nature and substance of logic changed during the decade that started with the work on my dissertation at UCLA. Let me, now that I have said a fair amount about the things that have preoccupied me in the decades after that (starting in the late seventies), add some more to those first remarks.

The main change I described in Section 1 was a move away from an absolutist view of logic, in which the central quest is for *the true logic*, towards a more diversified perspective, according to which different phenomena come with their own logics and where the logician's task is to discover or develop

'custom made' logics for the different phenomena. Determining the 'logic' for a given set of phenomena is not unlike 'modeling' these phenomena – that is: describing them at a certain level of abstraction. Such descriptions yield corresponding logics when they include the explicit formulation of special logical languages for representing the phenomena in question, with an explicit syntax and an explicit model theory.

The phenomena that have given rise to such logics in work of the past 75 years or so typically involve, in some way or other, the representation and/or manipulation of information. It may be a matter for debate whether this is so in all cases, but that depends in part on how the term 'information' is construed. In what follows I will assume that all logics are about information representation and processing. Or, to use a slogan that is strongly associated with the *New Amsterdam School of Logic*: The subject of Logic is the representation and manipulation of information.[54]

6.1 Away from the Primacy of Formal Deduction

One tendency that has grown in parallel with the shift to custom-made logics is a lessened interest in what had been the central concern of formal logic since antiquity, viz. the identification of formal principles of valid inference. In the early days of the rise of special purpose logics, this continued to be a major concern, witnessed by the innumerable completeness theorems for modal logics (including tense logics, deontic logics, epistemic logics, interrogative logics and so on). The standard recipe for doing logical work in those early days, in the wake of Kripke's model-theoretic treatments of modalities, was: define a logical language to describe certain modal phenomena by formally specifying its syntax and its model-theoretic semantics. The model-theoretic semantics provides a semantic definition of logical consequence in the familiar way: the formula ψ is a *logical consequence of* the set of formulas

---

[54]What I am referring to as the 'New Amsterdam School of Logic' is that multitude of developments in the logical approach to the study of mathematics, computer science, language and cognition that I associate in the first instance with the University of Amsterdam's ILLC (Institute for Logic, Language and Computation). My greatest personal debt for how my views about the nature and scope of logic have changed in the course of my career is to Johan van Benthem, who has put his imprimatur on the 'New Amsterdam School' more than anyone else. The 'New' in 'New Amsterdam School of Logic' is meant to distinguish it from what has been referred to as the 'Amsterdam School' in the past (and perhaps also today): the group of those who are best known for the development of Intuitionistic Logic as an alternative to Classical Logic – first and foremost: Brouwer, Heyting, Beth – as well as other classical contributions to logic, such as Beth's work on definition and his development of the method of Semantic Tableaux.

Φ iff every possible truth evaluation (in a model or in a model at an index; for indices see below) which yields truth for all formulas in Φ also yields truth for ψ. A completeness result for this semantic notion of logical consequence then consists in defining a system of formal proofs in which it is possible to derive ψ from Φ if and only if ψ is a logical consequence of Φ. (For the most part the proof systems considered in this work were systems of axiomatic deduction, in which a conclusion is defined by successive applications of steps that either take the form of the instantiation of one of the system's *axioms* or applying one of the system's formal *inference rules* to one or more already established lines of the proof. But other types of proof systems – especially Natural Deduction systems, in which there are rules, but no axioms – were considered as well.)

As time went by, there was a growing sense that more often than not such completeness results do not have much to tell us that is of conceptual interest and that has caused the community to lose much of its initial interest in work of this kind. There have been at least two factors that I believe have played a major part in this loss of interest. The first is that it is almost always quite easy to prove that the given consequence relation is 'axiomatizable' – i.e. that there must be a proof system, axiomatic or other, that matches it – even before any particular proof system is exhibited for which completeness can be demonstrated. (In more technical terms: It can be shown without much difficulty that the consequence relation – the set of pairs $<Φ,\psi>$ such that ψ is a logical consequence of Φ – is recursively enumerable, and that without showing of some particular proof system that it matches this consequence relation.) Once axiomatizability has been established, something might still be learned from the explicit formulation of some particular matching proof system. For instance, it might be of interest to see that some particular set of axioms and rules suffice for completeness, especially when the axioms and rules enjoy some intuitive plausibility. But still, often the surplus value of finding some particular complete proof system is quite limited.

A second reason, which applies in particular to logics that serve as semantic representation formalisms for fragments of natural languages, is that these logics aren't 'axiomatizable' at all. This is true for instance for languages designed as representation formalisms for fragments that include plurals. The consequence relation for such languages is almost inevitably second order – this is because they must have the expressive power of quantifying over sets of individuals and not only quantification over individuals – and as a rule it isn't hard to demonstrate that they have this property. For familiar reasons

this excludes the possibility of a completeness theorem in the strict sense.[55] That doesn't mean, of course, that there is no room for partial deduction systems, which do not cover the entire set of valid consequence pairs, but that might be argued to capture the full repertoire of valid consequence pairs that are *cognitively relevant*, in the sense that they are within the capacities of competent speakers of the natural language fragment. But to my knowledge no results of this latter kind have been established either, for one thing because there are no good characterizations of which inferences are cognitively relevant in the indicated sense and which aren't. And one may doubt that such characterizations are possible at all.

Non-axiomatizability of the consequence relation for logical languages that are designed as representation languages for fragments of natural languages may also arise for another reason. Quantification in natural language is very often over domains that we know must be finite, but where we have no knowledge of how large the domain is, and where our understanding of the inferential properties of those quantifications do not include any specific finite upper bound on the domain size. A good case can be made, it would seem, for the claim that the logic of such a representation language is like that of the 'theory of finite models': the class of models consists exclusively of models in which the domain or domains in question are finite, though with no finite upper bound to their size. It is a long and well-known fact that the consequence relations generated by such model classes are not axiomatizable. Such consequence relations are 'Π1' (i.e. the complements of recursively enumerable relations).

When domain finiteness is the source of non-axiomatizability, it may be possible, however, to attach a more operative sense to the distinction between valid inferences simpliciter and valid inferences that are cognitively relevant. Perhaps the cognitively relevant valid inferences in this case are those that hold not only in models in which the domain or domains in question is/are finite, but also in models in which the domain(s) is/are infinite. If this is right, then the set of cognitively relevant valid inferences could well be axiomatizable although the set of valid inferences is not.

---

[55]For one argument to this effect see (Kamp & Reyle 2011).

6.2 The Role of Context: Contextual Dependence of Evaluation and Contextual Dependence of Interpretation

One aspect of the 'logics of language' – i.e. of the logics that have been developed for the purpose of logically transparent representation of contents expressible in natural languages – is the role of context. In the majority of custom-made logics of the 75 years formulas are not simply evaluated in models simpliciter; instead evaluation is in a model *at an index*. (Typical examples of indices or index components are possible worlds, times and those elements which determine the references of the classical paradigms of *indexical expressions*, such as the pronouns *I* and *you* or the adverb *now*.) As more logics have been developed for dealing with more meaning aspects of natural languages, the repertoire of indices on which truth evaluation may depend has grown too.

Index-dependent evaluation is one of the forms in which context dependence can manifest itself, but it isn't the only one. Context dependence is arguably an even more important factor in the representation relation itself: the algorithms that compute the logical forms/semantic representations for expressions of the given natural language fragment also use contexts as inputs, and not just the syntactic forms of the expressions for which the logical forms are being computed. This dimension of context dependence has been prominent in much of what these reflections have been about: how context affects semantic representation has been discussed in Section 2 as one of the core concerns of Dynamic Semantics and in particular of systems based on DRT.

Context dependence is a topic that seems to offer scope for more formal work. Over the years much information has been collected about the ways in which the various uses of natural language expressions rely on the context of use for the content of the expressions used. Mapping natural language utterances to semantic representations in the manner of, for instance, DRT is a way of capturing the contextual contributions by integrating them into the form of the resulting representation. (What context-dependence remains will manifest itself as index dependence of the truth evaluations for the resulting representations.) What to my knowledge is missing from the logical literature is a systematic investigation of the forms that context dependence can take, and in particular of context dependence of the first, representation determining sort. A systematic investigation of this sort will require a

formal characterization of possible contextual contributions. As part of that we will need a formalization of such information at the 'meta-level' at which context-dependent representation constructions are formulated. Also, more will have to be said about the kinds of effects the contextual information thus formally described can take on the representations constructed. (Work on Articulated Contexts can be thought of as a first step on the way toward such an analysis of context dependence.)

One of the results of such an investigation would be a classification of representation formalisms for natural language fragments in terms of the kind and amount of context-dependence is involved in going from the natural language expressions to their representations. We may expect that this classification will induce a certain partial order among those systems, determined by some measure of the amounts of context dependence involved. As things stand I have no clear idea what this partial order might be like, and it is something that I do not dare speculate about.

6.3 Context Dependence in the Language of Mathematics

That context dependence is a feature of language as we use it has become a truism. The challenge is to figure out all the different ways in which context-dependence manifests itself, not to show that there is any. But is context dependence truly ubiquitous? Are there no corners in the garden of language that are fully 'context-free' – corners where the weeds of context have been fully eradicated? The question is important as a universal question about human languages and about human cognition, as that which makes human language possible.

The question is also important for the history and philosophy of science. By the end of the first quarter of the last century it had become plain to many that formal logic was indispensable to the foundations of mathematics. The use in mathematics of natural language, with its multiple ambiguities, had led to a mess that needed formal logic to be sorted out. And once the mess had been sorted out with its means, natural language should be shunned henceforth, lest one would land back in the same kind of mess before long.

But of course that didn't stop the use of natural language in mathematics, not even in discussions of its foundations. Perhaps there is no need to see this as

inconsistent with the view just expressed: so long as the mathematical statements that really matter are given in the language of formal logic, the use of natural language by mathematicians discussing the formalized statements can be seen as a kind of heuristics, which facilitates the communication between them, without affecting the independently established content of those statements. But it somehow remains puzzling that the need for these heuristics should be so strong: if mathematicians are smart enough to express their statements in formal terms – and there is an implicit consensus that all or most of them are a whole lot smarter than that – then why should they need the frills of natural language when communicating? We can see the makings here of a kind of informal paradox

An important antidote to this paradox – the view that natural language is fundamentally unsuitable to the needs of mathematical discourse and the impossibility to do without it when engaging in such discourse in practice – was the development of natural language semantics. Perhaps the most famous quote from Montague's work is that according to him there is no important theoretical difference between natural languages and the artificial languages of formal logic.[56] Whether this claim is true is arguably no longer even a matter of debate: The more we learn about natural languages the more prominent seem the differences that we can see between them and the classical formal languages that I believe were foremost in Montague's mind when he made this statement (the classical Predicate Calculus and the Lambda Calculus (and Montague's own version of it, his Higher Order Intensional Logic)). But even today the profound insight behind the statement is, I'd like to think, more important than all that speaks against it. Appreciating the insight requires of course looking at the substance of the paper, which consists of a fragment of English for which Montague develops a strictly defined syntax, and building on that, a mathematically exact model-theoretic semantics. The fragment is somewhat stilted, but it has the methodologically important property that it covers the full first order Predicate Calculus. So every mathematical statement that can be formulated in predicate logic can also be formulated within this fragment (and in fact using sentences from the fragment that are entirely free of ambiguity).

The moral of this and Montague's other work on natural language semantics for the topic of this section is this: if we want to, we can use natural language in the formulation of mathematics without compromising on any of the salient virtues we associate with languages of formal logic. But the

---

[56]The relevant passage is the opening sentence of (Montague 1970).

work doesn't fully resolve the 'paradox' I spoke of above. For when mathematicians do not limit themselves to the use of formulas, but employ natural language in addition or instead, the kind of natural language they use is typically not of the neat, ambiguity-free sort that Montague's fragments exemplify. The voluminous amount of work on natural language semantics of the past six decades for which Montague's work was the initial example and impetus has done some to resolve what remains of this paradox in that it has given us an ever better understanding not only of the ambiguities and context-dependencies that are found in language as we know and use it, but also of how the contexts that we need to resolve the ambiguities are available in the situations in which the utterances displaying them present themselves to us. But nevertheless, the question why the use that we make of natural language in the practice of mathematics can be so remarkably effective unfolds into a range of more specific puzzles, to which we still lack satisfactory answers.

Over the past two decades these puzzles have drawn the attention of mathematicians. One motivation has been the following. A question of permanent importance in mathematics is the correctness of proofs. But proof checking is a delicate matter, and more so, obviously, as proofs get longer and more complex. Therefore automated proof checking, done by computers with with their indefatigable and unfailing attention to boring but crucial formal detail, would seem natural assistants to which proof checking tasks might be delegated. The first powerful automated proof checkers go back to the sixties.Probably the best-known of these is De Bruijn's Automath; see e.g. (Nederpelt, Geuvers & de Vrijer 1994). But the problem with Automath – and I believe the same is true for other automated proof checkers – is the form of its inputs. The inputs it requires are compilations of mathematical proofs in the form of unabridged formal deductions in some proof system of formal logic. That is hardly ever the form in which mathematicians conceive and present their proofs. So to use Automath in actual practice, as a tool for checking proofs for which checking is felt to be important, the given proofs first have to be translated into such a fully explicit form. For interesting proofs that tends to take a lot of time, and it also turns out to be not all that easy and to require sometimes non-trivial knowledge of the mathematics in question.

This part of the checking problem – the casting of given proofs into a form that can be fed into automated proof checkers like Automath – is something that we would like to leave to computers too. But writing algorithms that translate proofs as mathematicians like and want them into such forms has

proved to be a surprisingly hard problem, and much more recalcitrant to automation than the design of the automatic proof checkers themselves. My own familiarity with efforts towards a solution to this 'preparatory' part of the automated proof-checking problem is largely limited to those which have thought to use DRT as an interface: Proofs in the forms in which they are found in the mathematical literature – in particular in textbooks that the mathematical community has come to recognize as models of clarity and 'informal rigor' – are first translated into DRS-like representations, and a further algorithm is then to translate those representations into the inputs that Automath (or some other similar proof checker) accepts.[57]

That these projects chose DRT as an approach to natural language semantics that they expected to suit their needs is not surprising, given the alternatives that linguistic semantics has to offer. Many of the problems that a translation from textbook proofs to inputs to classical proof checkers must be able to solve have to do with sentence-transcending anaphora. But often the anaphoric connections are much more complex than those familiar from the DRT literature. To give a flavor of just some of these consider the following example, taken from a recent article by Thomas Hales, in which he makes a case for a new concerted effort to develop a *Controlled Natural Language* ('CNL') for mathematics (Hales 2019). This CNL should make it possible to formulate proofs of mathematical theorems in ways that mathematicians would accept as capturing the ideas behind those proofs as they understand them and that can nevertheless be automatically converted into the inputs that are accepted by existing automated proof checkers (such as automath).

6.3.1   An Example of a Proof Text

As an example of such a formulation Hales quotes a formulation proposed by the Naproche project for the proof that the structure $\mathcal{R}$ of the real numbers is *Archimedean*: If x, $y \in \mathcal{R}$ and $x > 0$, then there is a natural number n such that $y \leq n.y$.

Here are the quoted Naproche formulations of theorem and proof:

Theorem 1 If $x \in \mathcal{R}$ and $y \in \mathcal{R}$ and $x > 0$ then there is a positive integer n such that $n.x > y$.

---

[57]Examples are the work by Zinn, in particular in his Erlangen doctoral dissertation, the Naproche project of the University of Bonn. Related to these approaches is also the Cambridge University dissertation and following work of Ganesalingam; (Zinn 2003), (Kühlwein, Cramer, Koepke & Schröder 2008), (Ganesalingam 2013).

Proof. Define $A = \{n.x \mid n \text{ is a positive integer}\}$. Assume the contrary. Then y is an upper bound of A. Take a least upper bound α of A. $\alpha - x < \alpha$ and $\alpha - x$ is not an upper bound of A. Take an element z of A such that

(15b) $\qquad z \not\leq \alpha - x.$

Take a positive integer m such that $z = m.x$. Then $\alpha - x < m.x$ (by 15b). $\alpha = (\alpha - x) + x < (m.x) + x = (m + 1).x$. $(m + 1).x$ is an element of A. Contradiction. Indeed α is an upper bound of A.[58]

Let us have a look at this proof as a piece of text for which we want to find a correct, logically transparent representation, of the kind envisioned by DRT, one that consists of a number of inferentially connected propositional representations whose inferential connections can be verified (and thereby 'checked') by a suitably equipped deduction system. Note well that if we succeed with this, that doesn't mean that this text representation can serve as input to one of the existing automated proof checkers. But I believe that if we do not find a general method for converting proofs like this one into such logically transparent representations, then there can be little hope of developing an algorithm that translates such proofs into suitable inputs for existing automated proof checkers either.

I won't try to provide an exhaustive analysis from the DRT-inspired perspective of what is needed to derive such a logically transparent representation for this particular proof here, let alone a general algorithm for deriving such representations from such proofs in general. But the following remarks should give a clue of how DRT can contribute to the general project of going from theorems and proofs as mathematicians like to see them to the inputs that automated proof checkers want.

The following remarks will make use of aspects of DRT that were not available at the time when the mentioned projects got under way, or at any rate when these aspects were not prominent and probably not visible to those who undertook the projects. But these aspects have been prominent enough in what has been said in earlier parts of these Reflections and I will make freely use of some of that.

---

[58]This last sentence sounds a little peculiar to me. True, it has just been established – immediately before "Contradiction." – that $\alpha < (m.x) + x$, which contradicts the assumption that α is an upper bound of A. Perhaps the sentence is just meant to remind us of this.

First, there is the over-all *rhetorical structure* of theorems and proofs. One of the important insights, we noted in Section 4.2.2, is that no text has been properly understood unless enough rhetorical relations have been established between its Discourse Units. In this regard mathematical texts are no exception, and this is true in particular of text bits consisting of a theorem and its proof. For instance, it is (obvious but) crucial that we recognize the proof as a proof of the theorem, something that is conventionally indicated by the fact that the proof part of the text immediately follows the theorem part. (In cases where that is not so, e.g. when the proof of the theorem has been interrupted by the statement and proof of a Lemma, it is common to say explicitly when the proof of the theorem is resumed, or one finally gets to it, that what follows is the proof of the theorem, by referring to the theorem explicitly.) But then there is also the internal 'rhetorical' structure of the proof. And here, the structure of proofs differs considerably from other text types, and especially from those that have been the principal focus in work on rhetorical structure in linguistics, including that of SDRT. The discourse relations that have been most prominent in that work, among them Narration, Elaboration and Explanation, may have some application to proof texts too, but these are not the relations that are most directly relevant. The most important 'rhetorical' relations for the understanding of a proof have to do with what is supposed to follow from what – which sentences and clauses play the part of premises, and which play the part of what is claimed to follow from them, at which stage in the development of the proof. (The temporal development of proofs is a crucial aspect of them, for the roles of the Discourse Units of the text – its sentences and clauses – typically change as the proof progresses; the same DU that serves as what is claimed to follow from other DUs at one stage will function as one of the 'premises' from which something else is claimed to follow in its turn at the next stage.)

To cut a longer story short, the 'rhetorical' structure of proofs can best be understood as the one that has found its most explicit and best known formulation in systems of Natural Deduction.[59] I will assume in what follows that it is possible to formulate an algorithm that convert proofs into repre-

---

[59]One should also mention here the development of logic out of the rhetorical structure of debates, a modern version of which was developed in detail by Lorenzen (Lorenzen & Lorenz 1978) and some of his students. Here the rhetorical motivation and foundations have been much more explicit. Nevertheless, I am not following up on this connection between rhetoric and logic because it seems to have played not much of a role (indeed, if any) in the efforts to create converters or interfaces between mathematical texts and the inputs to proof checkers.

sentations that have the form of natural deductions, in which the lines are representations of the DUs of the proof text – I assume that these representations are DRSs or DRS-like representations; for more about this below. How the Natural Deduction can be reconstructed from clues in the text is to a considerable extent a matter of the conventions observed in writing proof texts. Part of these is the use of certain *discourse particles*, like *then, therefore, hence, so* or *thus*. I assume that sentences governed by one of these are to be considered as making claims that follow from what precedes them in the text and thus that their representations can be deduced from the representations of that preceding material. (Exactly how the algorithm is to determine in general which part of the preceding text can serve as premise material for the deduction is something I do not understand all the ins and outs of, but I do not expect serious difficulties on this account.)

Part of the material in a proof text from which a given DU is claimed to follow takes the form of sentences in the imperative mood. Examples in the proof above are 'Define $A = \{n.x \mid n$ is a positive integer$\}$', 'Assume the contrary' and 'Take an element z of A such that ... '. Here we touch on another aspect of discourse structure that cannot be found (to my knowledge) in the publicly accessible DRT-literature to date. This is the distinction between different types of *Speech Acts* (Searle 1969). Typically, sentences in the imperative mood are used to perform different Speech Act Types than sentences in the indicative mood. The latter are primarily used to make assertions, whereas imperative mood are used to give advice or instructions (and occasionally also to give commands). Imperative mood sentences are found in proofs because proofs are often conceived of as recipes for how to get from the premises of an argument to its conclusion. As part of these recipes proofs often 'recommend' the introduction of a 'name' for an explicitly defined concept. Thus the sentence 'Define $A = \{n.x \mid n$ is a positive integer$\}$' in the proof above introduces the name A for the set denoted by the definiens '$\{n.x \mid n$ is a positive integer$\}$'. This is a kind of contribution that DRT is well-equipped to deal with. The representations of such sentences can be assumed to take the form of DRSs in which the 'name' of the defined notion – A, in our example – takes the form of a discourse referent that is placed in the Universe of the DRS representing the sentence while the definition itself is represented in the form of one or more conditions in the condition set of this DRS (as well as additional discourse referents in its universe, if and when the individual case requires). Sentences following the definition in the text can then build on this DRS, in particular through anaphoric reference to the new discourse referent it introduces (here A). (Also, as standardly the case in DRT, the DRS itself may have to be merged with that for the relevant

100

preceding part of the text.)

Sentences beginning with *Take*, as in 'Take an element z of A such that ... ', can be represented in much the same way. In this case it is the symbol that occurs as direct object of *take* that plays the part of a new discourse referent (here: z) that is inserted into the universe of the DRS representing the imperative sentence, whereas the constraint imposed on the choice that is expressed by the rest of the sentence (here the relative clause beginning with *such that*) is represented by conditions of this DRS. Note that in building such representations names like A and variables like z are both treated as discourse referents. This is a good illustration of the principle that discourse referents can serve as representations of particular entities and as variables, which range over some set of entities. And it also illustrates another important point: The introduction of variables in mathematics has often been hailed as one of the most important steps in its development, as a device that vastly simplifies the formulation of mathematics propositions and arguments and thereby has liberated the mathematician's mind from a lot of unnecessary ballast imposed by the particularities of natural language grammar. But a large part of why variables are so exceptionally useful is that they can be *bound across sentences*: A variable introduced in one sentence can then be re-used in subsequent sentences, while it is clear how these later uses of it are logically connected with its earlier use. What the logical connections are between the different occurrences of the variable is a function of how the different sentences or clauses containing those occurrences are logically connected with each other. The logical connections between the different sentences of a proof are a perfect example of this. When a variable is introduced in a sentence that is part of the premise material of a claim that some proposition C follows from that material, then it will end up universally bound in the representation of that claim.[60] When the variable is introduced

---

[60]I have not been very precise about the notion 'proposition' in this document. In some places where I have used the term, as in the discussion of singular propositions in Section 5.3, the notion that has been in the back of my mind was the construal most widely assumed within formal semantics, viz. that according to which propositions are sets of possible worlds. But that notion is notoriously inadequate in relation to the use of the term 'propositions' in mathematics (as that term is used informally by both mathematicians and non-mathematicians). One consequence of the logical omniscience that is built into possible worlds semantics is that when the proposition expressed by a sentence is identified with the set of possible worlds in which the sentence is true, then far too many mathematical sentences come out as expressing the same propositions. (For instance, any two sentences formalized in ZF Set Theory that are both provable in the theory or both refutable will express the same proposition.) So when we are talking about propositions in the context of mathematics sets of possible worlds cannot be what we want. But what

by the sentence that expresses C, then it is bound existentially.[61]

It should also be noted in this connection that the enormous advantages of using variables in mathematics depend in large part on the practice of their trans-sentential binding. This aspect of the use of variables is easily obscured by the treatment of variables in formal logic since Frege, where the scope of variable binding is always limited to single formulas. The tension between this constraint on the scope of variable binders in the standard formulations of predicate logic and the role of variables in logical deductions has played an important part in the 'formatting' of deduction systems, where variables can be 'flagged' (Quine), or complemented by 'parameters', a kind of halfway house between variables and constants (e.g. in certain definitions of the method of proof via semantic tableaux). My own impression is that the discourse referents of DRT are closer analogues of variables as these are used in 'informal mathematics' than the variables of the standard formulations of predicate logic. But substantial support for this impression would have to come from a more detailed and historically better informed study than I am able to provide. (For a formally detailed approach to this problem, combined with careful philosophical reflections see (Fine 1985). But this study doesn't discuss the roles of discourse referents.)

The third imperative mood sentence occurring in our sample proof text was 'Assume the contrary.' Sentences beginning with *Assume* or *Suppose* can also be used to introduce new discourse referents (e.g. 'Assume that y is one of the members of the set ... .'). But more often such sentences introduce propositions - propositions that are to count as part of the premises in the next part of the text, which is 'governed' by the assumption or supposition. This is true in particular for the sentence 'Assume the contrary.' But here we encounter another problem familiar from discourse semantics: the interpretation of definite descriptions. In the case at hand the description that needs to be interpreted is *the contrary*. This is an example of a so-called *bridging description*: the head noun *contrary* of the description is a relational noun;

---

else? I don't know and I believe nobody quite knows. In the present section I nevertheless go on using the term 'proposition', but now more informally, according to which the proposition expressed by a sentence is identified by the semantic representation of the sentence, but without having to be identical with that representation.

[61]In this respect the 'follows' claims made in proof texts are just like the conditional DRS conditions – conditions of the form $K \Rightarrow K'$, where K and K' are DRSs – in which the discourse referents in the Universe of K are universally and those in the universe of K' existentially bound. I also recall that in early presentations of DRT it was observed that conditions of the form $K \Rightarrow K'$ can and should not only be used to express conditional sentences, but also bits of text of the form 'Suppose P. Then C'.

the question it raises is: 'The contrary of what?'. Intuitively the answer is clear: The referent is the contrary of the theorem. But how is the algorithm to decide that it is this proposition whose contrary is the referent of the description?[62]

The representation of the sentence 'Assume the contrary.' involves a fair amount of unfolding. Wanted is a representation of the negation of the representation of the anaphoric antecedent of the bridging description *the contrary* – in other words, a representation of the negation of the representation of the statement of the theorem. Complexities of this kind are not discussed in the versions of DRT that were available to the mentioned projects; the treatment of nominal anaphora in those versions is limited to pronouns, and even for those little is said about the details of anaphora *resolution* (i.e. on how to *find* the intended anaphoric antecedent in the actual construction of sentence and text DRSs). But the complications that arise for the interpretation of bridging descriptions are still comparatively simple in comparison to others forms of anaphora, with which mathematical texts are also rife. Examples are phrases like *and analogously/likewise for the remaining cases*, *The other case is proved the same way.*, *By a similar calculation the second term also evaluates to 0.* and so on. Expressions like *analogously*, *likewise*, *similar* and the likes are of course also frequently used in non-mathematical texts. (I am aware of being one of the worst sinners, as amply illustrated by the present non-mathematical text.) The use of such expressions in non-mathematical discourse is for the most part not crucial to the semantic representation of the discourse or text. For the semantic representation it suffices to register their contributions in the form of their existing some kind of similarity between the two or more items in question. This leaves open what the similarity consists in, but it seems legitimate to consider the answer to that question part of the *evaluation* of the representation, i.e. of the process that leads to the interpreter's decision whether or not to accept the text or discourse as 'right'.

But occurrences of phrases of the mentioned kind in proof texts cannot be

---

[62]If I am right, bridging descriptions have a preference for finding their antecedents in the immediately preceding sentence. With *the contrary* this would mean: the proposition expressed by that sentence. This may explain my own feeling that in the proof text above the placement of the sentence 'Assume the contrary.' is a little odd. My impression is that the reader first looks at the immediately preceding sentence, the one beginning with *Define.* But that sentence is clearly unsuitable for the purpose. What would it mean to form the contrary of a definition? So one zeroes in on what seems the next best option, viz. the statement of the theorem, which is also the next earlier statement. But how should we design a DRS construction algorithm that is going to make the right choice in cases like this one?

handled in this manner. In particular, if the semantic representation of the text is to provide a suitable basis for the construction of proper inputs to proof checkers, then it must make the analogies or similarities that these phrases speak of fully explicit. Typically this will have to take the form of finding the part of the text to which the 'remaining cases' are claimed to be analogous or similar – the 'anaphoric antecedent' of the phrase – and then identifying the analogy or similarity by seeing how the representation of this anaphoric antecedent can be applied to the 'remaining' case or cases. In many cases this analogy or similarity can be identified by abstracting from this representation with respect to those constituents that have to be replaced by constituents from the 'remaining' case or cases to obtain the representation of the part of the proof that the phrase in question refers to.

6.3.2 The Challenge posed by Phrases and Sentences that invoke Analogies

To make this a little more concrete[63] let us suppose that the phrase in question is the sentence

(15)   The remaining case is proved analogously.

From a formal semantics perspective this example is easier to deal with than the ones mentioned above, since those involve various forms of ellipsis or abbreviation that makes it more difficult to deal with the compositional aspects of their meanings. (15) is a regular subject predicate sentence whose VP is the phrase *is proved analogously* and whose subject is *the remaining case*. The first observation to be made about this sentence is that, although it is in the indicative mood, it should be understood in the spirit of an instruction – the spirit in which proof texts are to be understood generally and which explains why so many of the sentences found in them are in the imperative mood. So the contribution that the sentence should make to the representation of the text should take the form of the representation of another proof structure, which is constructed on the basis of the part to which the sentence points.

---

[63]This paragraph should be replaced by an actual example involving a proof from some actual proof text (from some mathematical textbook, say) with an actual occurrence of a phrase of the kind being discussed here and the actual construction of the DRT-based representation of this text, with a detailed treatment of the phrase in question as part of that construction. I do not have the time right now to do this.

The next three observations are about the subject DP *the remaining case.*
The first observation concerns the interpretation of the widely and loosely
used noun *case.* What sort of case is being talked about here? The answer
is given by the verbal head *prove* of the VP to which the DP *the remaining
case* is the subject: Things that can be proved are propositions. So the in-
terpretation of *the remaining case* in (15) must identify some proposition as
its referent.

Secondly, there is the contribution that is made by *remaining* to the DP *the
remaining case.* The adjective *remaining* is like *other* in that (i) both are
2-place adjectival predicates and (ii) when used prenominally, their second
arguments are implicit and have to be recovered from context. (The first
argument of the predicate is the referent of the phrase in which it occurs
as modifier, here the phrase *the remaining case.*) The implicit argument of
remaining in *the remaining case* must be a thing (or collection of things; but
in the rest of this discussion I will suppress the collection possibility) that
belongs to the same ontological category as the referent of the DP in which
it is a modifier; so in the case at hand it must also be a proposition.

Thirdly, the referent of the definite description *the remaining case* will also
have to be recovered by the representation. Furthermore, both this recov-
ery and the recovery of the second argument of *remaining* will have to be
'anaphoric', in the sense of being based on the preceding part of the text.
Exactly how these recoveries would have to proceed cannot be discussed in
the absence of an actual text in which (15) would occur. But typically a
sentence like (15) will occur at a point where the proof has already been
divided into a number of sub-cases. Let us assume that this information is
easily accessible in the representation that has already been constructed at
this point, and that it is also readily recognizable which of the cases have
been proved already and which case is remaining, and that this information
can be used for both recoveries. Typically, we may expect, recovery of the
referent of *the remaining case* and recovery of the missing second argument
of *remaining* will go hand in hand. But without an actual text in which (15)
occurs it won't be possible to say anything more precise.

To summarize these remarks about the subject of (15): *the remaining case*
requires two identifications, of its referent and of the second argument of
remaining. Both the recovered entities must be propositions and typically
these recoveries go hand in hand relying to a large extent on the same parts
of the (representation of) the preceding text.

We now turn to the VP of (15). As noted above, the constituent *remaining* of *the remaining case* forces its interpretation to identify *remaining*'s implicit argument. That renders the DP *the remaining case* a kind of bridging description, like the DP *the contrary* discussed above. But note that (15) also contains another bridging constituent, viz. the modifier *analogously* of its VP. *analogous* is like *remaining* and *other* a 2-place predicate. That is equally true of its adverbial form *analogously*. When this form is used on its own as modifier in a VP, then its second argument is also implicit and thus must be reconstructed from context.[64] And here the same constraint is operative that we observed for the reconstruction of the second argument of *remaining* in *the remaining case*: the reconstructed second argument of *analogously* must be of the same ontological category as the entity described by the verb *prove*, viz. a proof. In other words, the VP of (15) must be construed as referring to a proof that is analogous to the one (or the ones) that the interpretation must recover.

A further aspect of the interpretation of (15) is that the reconstructions of the second arguments of *remaining* and *analogously* must be correlated: the proof reconstructed as second argument of *analogously* must be a proof of the case reconstructed as second argument of *remaining*. I expect that this kind of 'deictic convergence' is a common interpretational requirement for sentences and phrases of the sort we are discussing. But as things stand, I do not know how general such deictic convergence is, or what would be the right way to state it without overgeneralizing.

Let us suppose that our representation algorithm can identify the earlier case and the corresponding proof. The representation of the proof text that has been established at the point where (15) gets interpreted will contain representations of both the case – the representation of a proposition – and the proof – a proof representation of the kind alluded to when we spoke about the rhetorical features of proof texts. Likewise the referent of the DP *the remaining case* will be identified as the representation of a proposition. The remaining task for the interpretation of (15) is to construct a representation for the proof that the sentence describes: a proof representation that stands to the representation of the referent of *the remaining case* in the same way as the proof representation recovered as the second argument of *analogously* stands to the representation of the recovered second argument of *remaining*.

---

[64]*analogously* differs from *analogous* in that it can be as the head of more complex verb modifiers, as in *proved analogously to the proof we just went through*, where the DP *the proof we just went through* fills the predicate's second argument slot.

106

The problem we are facing here is a subspecies of the general problem of analogy as it is understood in linguistics, psychology and cognitive science: what does it mean for A to stand to B in the same way as C stands to D? As I suggested when discussing the semantics of *analogous* and *analogously*, the everyday use of these words might be dealt with in natural language semantics by adding conditions that represent 'analogy' relations between the relevant arguments. (If analogy is understood along the lines above, then these conditions should take the form of predications involving a 4-place predicate and discourse referents for the four arguments A, B, C, D.) What the analogy then precisely consist in, I suggested, may be assumed to belong to another module involved in the processing of verbally supplied incoming information.[65]

Whatever the communication-theoretic status may be of analogies in the use of language outside mathematics, the reference to analogies in proof texts obviously have to be resolved at the level of content representation. (For one thing, if analogies like the one we have identified in connection with (15), then a semantic representation that leaves the analogy unresolved couldn't possible provide the correct input for a proof checker.) What we need for instance for the interpretation of the contribution to a mathematical text that can be made by (15) is the construction of some particular representation of a proof of the proposition denoted by *the remaining case*. But how can such a representation be constructed from the three different pieces identified in the last section – the representations of the two propositions and the representation of the proof of one of them? The basic strategy seems clear intuitively. First, we have to compare the representations of the two propositions. These representations should correspond in the sense that one can be mapped to a propositional representation that is logically equivalent to the other by a substitution of discourse referents. More specifically, suppose that in the representation K_{p1} for the proposition that has been recovered as the second argument of *remaining* (i) there are discourse referents $\alpha_1, ...,$

---

[65]This should be stated a little more carefully. It has been observed many decades ago that the understanding non-literal uses of words and word combinations – and this applies in particular to those non-literal uses that have been identified in the literature as 'metaphorical' – requires the perception of an analogy between the non-literal use and a corresponding literal one. In such cases an interpretation which merely registers that there is some kind of analogy between the given use and one or more literal uses (perhaps without even identifying the literal one(s)) wouldn't be good enough; it would have the status of some kind of *underspecified representation*, which would first have to be completed by articulating what the analogy (and what the missing arguments of it are). Perhaps the status of such unresolved analogies is not unlike that of certain unidentified rhetorical relations. There would appear to be ample scope for more work in this area.

α_n that are either free in K_{p1} or belong to its universe, and likewise (ii) that there are discourse referents β_1, ..., β_n occurring in the representation K_{p2} of the proposition denoted by *the remaining case* that are either free in K_{p2} or occurring in its universe, so that (iii) substituting the β_i for the α_i in K_{p1} produces a proof representation K'_{p1} that is a proof of $p2$. If that is so, then K'_{p1} can be used as representation of the missing proof.

I do not know whether constructing the missing proof representation will always be quite as streamlined as just sketched. It may well be that sometimes more changes are needed to the recovered proof than the mere replacements of discourse referents. But I suspect that to the extent that such more difficult cases occur, this is because some of the missing details can be filled in by the competent reader on the basis of knowledge about the relevant branch of mathematics.

6.3.3 Representing Proofs with the help of a Mathematical Data Base

With this we reach the last point I want to make about the interpretation of mathematical texts. By far the biggest challenge for a discourse semantics of mathematical texts is the availability of the knowledge that enables mathematicians to read the theorems and proofs they need, want or like to read and understand. In toto current mathematical knowledge is vast, and vastly beyond the capacities of even the most knowledgable members of the mathematical community. Mathematical knowledge has a hierarchical structure, with knowledge that is assumed to be shared by all mathematicians at the bottom and then, at higher echelons, knowledge that belongs to more specialized areas of mathematics. The more specialized theorems and their proofs get – the more 'esoteric', as some might put it – the more likely that their proofs will rely on theorems from higher echelons in this hierarchy as lemmas. A proof checking regime must be organized in such a way that when a proof uses a theorem as lemma, which will involve mentioning the theorem by its name, then it can take the theorem's application as proved and use it as a proved premise in the semantic representation of the proof. For the proof checker that is part of the regime this application will then have the status of an axiom, a statement that doesn't need further checking. In what remains of this section I assume that our proof representations make use of a precompiled mathematical data base MDB, in which all theorems can be found that are referred to in proof texts. Furthermore I will assume that the theorems occurring in this data base come with proofs and that those proofs are also given in the form in which mathematicians prefer to present

proofs, and that our DRT-based construction algorithm can convert those presentations, just as I have been assuming for the proof texts discussed so far.

For the DRT-based representations of proof texts on which this section has been focusing this means that theorems referred to in proof texts can be represented as extra premises, just as that is done for lemmata that are stated and proved as part of the proof text itself. But the cases of a lemma and a mentioned theorem are of course different in that a lemma will have its own proof as part of the given proof text, and applying the construction algorithm to this proof of the lemma will yield a representation of it that will automatically become part of the representation of the proof text as a whole. For a theorem that is 'called' in the proof text there typically won't be a proof as part of the text itself; rather, the theorem can be treated as having been proved, as documented in the MDB. So the application of the theorem to the case described in the proof text can be treated as a premise without further need of justification – 'as an *axiom*', as I put it above) The only thing that needs checking when a known theorem is mentioned is that the theorem is applied correctly: the objects to which it is being applied must all be of the sorts to which the theorem is restricted – natural numbers, real numbers, complex numbers, sets of those, functions of various types (e.g. functions from complex numbers to complex numbers, functions from real numbers to real numbers in the interval [0,1], functions from sets to cardinals, members of a group or of a Boolean algebra and so on, and so on). These preconditions for the application of the theorem can be represented as conditions which predicate the relevant sorts of the terms that represent the objects in the semantic representation of the proof text. It is then a further matter of design of the proof checking regime whether proofs of these preconditions should be explicitly represented as part of the DRT-based representation we are talking about, or whether finding such proofs can and should be left to the proof checker or to some other special purpose module operating in between the DRT-based representation and the proof checker.

The construction of the MDB is a task for mathematics, not for linguistics, and there is next to nothing that I venture to say about it here. Just a couple of remarks. First, I already mentioned that the knowledge it documents should be organized in a hierarchy the bottom of which contains knowledge shared by all members of the mathematical community, while its higher echelons are for more specialized knowledge – the more specialized the higher the echelon. This hierarchy is roughly reflected in mathematical practice insofar as authors of proofs will quote theorems by their names, without explaining

109

their content or providing proofs for them, only when these belong to a level of the hierarchy that they consider shared knowledge among their targeted audience. But apart from such a socially motivated hierarchy, MDB will also manifest a more formal hierarchical structure in that its proofs too may call other theorems that are documented elsewhere in it. The proofs of theorems called in one proof may themselves contain calls to other theorems, and so on, but not forever: After a finite number of steps we should hit bottom, with proofs in which there is no reference to results elsewhere in the MDB. In other words, the hierarchy established by reference of theorems in proofs must be well-founded.

This formally grounded hierarchy will be much more finely granulated than the socially grounded one. But the two should dovetail in that no proof intended for an audience whose knowledge is not supposed to go beyond a given level of the social hierarchy should call theorems from a higher level without explaining their content and, at a minimum, a reference to where a proof can be found.

Since we are assuming that the theorems in the MDB and their proofs are presented in a form to which the DRT-based construction algorithm can be applied, the algorithm should be able to find the theorem on the basis of its reference in a proof for which it has to construct a representation and then convert the statement of the theorem into a representation that can be included in the representation that is being constructed for the proof. (Alternatively, this representation might be explicitly included in the MDB.) Needed for this to work is that the expressions which proof texts use to refer to the theorems enable the algorithm to find the theorems they refer to in the MDB. Since there are quite a few theorems in mathematics that go by more than one name, and any of the different names for a given theorem can be used to refer to it in a proof text, the theorem should appear with all of its names in the MDB. (But that is surely one of the lesser headaches in putting a comprehensive MDB together.)[66]

One thing that the MDB should supply, then, is a set of theorems, each labeled with the name or names that mathematical texts use to refer to it.

---

[66]As a preparation to writing Section 6.3 I should have properly reviewed what resources have been put together within mathematics over the past years that could serve as MDB (of the form and serving in the way as I have roughly described them in this section), or that could be converted into such a data base without too much additional work. This is one thing that I still have to undertake and where help from those with expert knowledge in this area would be most welcome.

But that isn't all the MDB should contain. Mathematics texts also refer to mathematical structures and to particular mathematical objects. More precisely, definitions can be of particular structures, like the structure of the real numbers or 3-dimensional Euclidean space, classes of structures, such as groups, lattices, and subclasses thereof, such as symmetric groups, Boolean lattices and so on, and particular numbers, like the real number 0, the number e or the S3 group (the group of permutations of a set of 3 elements).

A definition always comes with a name for its Definiendum. This makes it possible to think of the definitions occurring in an MDB as *labeled Entity Representations*: Entity Representations that are equipped with one or more conditions to the effect that some name N is a name for them. (It is not uncommon for mathematical structures to go by more than one name, e.g. 'Boolean Lattice' and 'Boolean Algebra'.[67]) The difference with the Entity Representations that have been prominent in existing applications are the kinds of entities represented. In an MDB these will all be mathematical entities, and thus abstract. But in addition only some of them will be objects; most will be structures and classes of structures. This entails that the Entity Representations of these various kinds stand in various relations to each other, in particular the subsumption relations of 'is a'-hierarchies. Subsumption makes it possible to use the definition of of a super-concept in the definitions of it subconcepts, which 'call' the superconcept, just as proofs can 'call' theorems on which they rely.

In the same spirit we can also think of the theorems in the MDB as labeled Entity Representations. In this case the represented entities are yet of a different sort. Theorems are propositions, propositions that are always about some mathematical entities of some sorts, just as the propositional attitudes of the mental states of MSDRT are often about objects. And just as the content representations of propositional attitude constituents of the mental states of MSDRT make aboutness explicit by incorporating Entity representations that also belong to the state, so the theorems documented in the MDB can make what they are about by mentioning the labels of the definitions of the structures and objects they are about. So in this respect the structure of an MDB is not unlike that of the mental states assumed in MSDRT. And those wary of any kind of psychological connotation may prefer to think of an MDB as part of an Articulated Context, which corresponds roughly to a combination of its Encyclopedic and its Generic Context components.

---

[67]Recall Section 5.6. For labeled Entity Representations see (Kamp 2015).

Trying to cast MDBs as parts of ACs, however, may be deemed a rather artificial and unilluminating exercise. It is arguably more fruitful to see MDBs as pointers to a variety of different forms that contextual information may take. ACs may have their merits as context structures for the interpretation of definite noun phrases in a broad spectrum of discourses and texts. But there is no reason to assume this structure is universal. The study of context in mathematics can help us to rethink this issue.

An MDB suited to the purpose I have described – that of supporting the semantic representation of proof text that among other things can serve as basis for the construction of inputs to automated proof checkers – must contain a lot of information that anybody with a background in mathematics would consider too trivial to draw attention to, for instance the familiar properties of plus and times for the natural numbers, or for the integers, for the rationals, for the reals or for the complex numbers. These properties are typically taken for granted when mathematicians write proofs but they are no less relevant to the verification of those proofs. Our sample proof contains several examples of this, as where it says 'Take a least upper bound α of A'. This instruction is coherent only when there exists at least one least upper bound of A. That is guaranteed by the least upper bound theorem for the reals, according to which every non-empty set with an upper bound has a least upper bound. Another illustration is the series of inequalities and equalities 'α - $x < m.x$. $\alpha = (\alpha$ - $x) + x < (m.x) + x = (m + 1).x$'. Every subsequence of this string which consists of a relation sign ('$<$' or '$=$') flanked by the term preceding and the term following it is an instance of a little theorem, to the effect that the numbers denoted by the two terms stand in the relation the sign denotes. Most mathematical readers would only be irritated by any kind of reference to these theorems, or other overt justification for the instances of them that the proof uses. So it is right for our sample proof to omit them. But for a proof checking regime of the kind we are discussing it is precisely these 'trivial' theorems that present a special problem. Since the proof will typically lack reference of any kind to these theorems, how is the construction algorithm going to recognize these instances *as* instances of the corresponding theorems? In some such cases the wording of the proof may provide a useful clue. For example, the term 'least upper bound' may be enough for the representation constructor to find the relevant information in the MDB (in the form of the Least Upper Bound Theorem for the reals) and once it has found the theorem, build the representation of the relevant instance of it into the representation of the proof. But in other cases there may be no such obvious retrieval cues. Consider again the inequalities and equalities repeated above: How is the proof for each of those to be repre-

sented in the proof representation and how is the MDB going to help with that? For formulas of this particular type – equalities and inequalities between terms denoting real numbers – it shouldn't be too difficult to find a solution. But I have no clear idea what a general solution to the problem of unmentioned theorems might be like.[68]

This is an aspect of the interaction between the proof representation constructor and the MDB that the linguist is not equipped to deal with and that should therefore be left to the mathematicians. In fact, it is connected with a more general problem that is also much 'deeper'. As emphasized with particular force in (Ganesalingam 2013), an important part the effectiveness of the language of mathematics as mathematicians use it is its peculiar mixture of natural language and mathematical notation. Established forms of mathematical notation, Ganesalingam urges upon us, are indispensable to the conduct of mathematics because they incorporate, in conveniently condensed form, an often very considerable amount of mathematical knowledge and assumptions. The facts about the structures of natural numbers, integers, rationals, reals or complex numbers not only ground the general properties of addition and multiplication for these different structures. They also justify the use of various notational devices like the use of the sum symbol Σ when it is applied to finite or denumerably infinite sets of numbers, from any of the number structures just mentioned. (This use of Σ is justified by the fact that for each of the number structures mentioned the binary operation of addition is commutative and associative.) Likewise, the notation '$x + i.y$', with x and y real numbers, for arbitrary complex numbers and the possibility of applying the arithmetical operations to these terms which treat them as sums with the familiar properties of addition as on operation on the reals, is another example of this. Yet other examples are the use of exponentiation as an operation on the reals, the rules for calculating with the trigonometric functions *sin*, *cos* and *tng*, or – even more dramatically in my perception – the notational devices that have been developed for matrix and tensor calculus or those used in the theory of differentiation and integration (including in particular notation of partial derivatives and integration over multiple variables). In all these cases understanding the notation presupposes familiarity with the underlying theories, consisting not just of one or more theorems, but a combination of the relevant structures (such as the mentioned number

---

[68]Note that it wouldn't do for a proof checking regime to treat the applications of such 'trivial' theorems as not requiring checking since the theorems are so trivial that we can trust their applications as beyond serious doubt. Especially in chains of equalities and inequalities it is not uncommon for proofs to contain errors, precisely because the underlying principles are considered so trivial and their applications so obvious.

structures, with a bunch of theorems that identify the properties of those structures and can be proved on the basis of those definitions).

What should an algorithm for the representation of mathematical texts do with the bits of special notation that it encounters in those texts? How is it supposed to unpack these bits, so that their contribution to the proofs in which they occur can be made fully explicit? This is perhaps the most difficult task for a representation algorithm of the sort discussed in this section. Again, this seems to be a problem for which there may be no uniform solution and again it is one that should be left to the mathematicians. And if that is so, then – again – the design of such a representation algorithm will have to be a cooperative venture between linguists and mathematicians, with the latter doing most of the really hard part. (So it isn't just the design of the MDB that should be left to them, but also the lion's share of the design of the representation algorithm that makes use of the MDB.)

6.3.4   Conclusion to Section 6.3

The motivation for this section is the role of DRT in projects from between one and two decades ago which made use of DRT in their effort to account for the semantics of the kinds of mathematical texts that mathematicians like to read and write. DRT, I suggested, should have looked like a natural choice for those projects because of its commitment to representation and its trans-sentential dimension: trans-sentential anaphora is an important aspect of mathematical texts, which cannot be modeled easily by other approaches toward natural language semantics that were available at the time. I noted in this connection that the management of variables in mathematical texts more closely resembles the behavior of discourse referents in DRT than that of the variables of classical formal logic (i.e. like the variable of the predicate calculus). By much the same token, DRT's treatment of anaphoric definite noun phrases whose anaphoric antecedents occur in a sentence preceding the one containing the noun phrase can be of help in the semantic analysis of mathematical texts (although this help was limited at the time when DRT was adopted by those projects, since at that time the only anaphoric noun phrases for which a treatment could be found in the publicly available versions of DRT were pronouns). And that clearly isn't enough. In the discussion of the sample proof text chosen from the Naproche project we spent a good deal of attention on the 'bridging description' *the contrary*. That discussion was premised on existing work in DRT: (Kamp 2015), (Kamp 2019*b*). But this work is of more recent date and part of it isn't even published.

114

We then observed that proof texts also exhibit other forms of anaphora – like those associated with subject and verb phrase of the sentence 'The remaining case is proved analogously.' and that these are forms that so far have not been addressed in any published DRT work. A careful and more broadly cast analysis of the semantics and pragmatics of such other forms of anaphora is indispensable, we noted, for a satisfactory account of logically transparent representation construction of proof texts. But not only that. The study of these forms is equally important for the study of natural language discourse semantics in general.

I also suggested that the study of proof texts can teach us a good deal about the rhetorical structure of texts as well as about the contributions made to that structure by sentences of different grammatical moods. But the most important lessons that I expect from the study of mathematics texts is the structure and content of the contextual information that they presuppose. The extent and structure of context in the interpretation of written and spoken language is still in its infancy, and the notion of an Articulated Context in DRT is only a first shot in a general direction. To me personally the study of mathematical texts appears as a fruitful challenge especially as a route towards rectifying a notion of context that I have at times been in danger of having broader validity than it does.

To conclude: There are two complementary perspectives from which to study the semantics of language in mathematics. One is that of the mathematician, who may want a systematic analysis of mathematical prose for purposes that are internal to mathematics itself, for instance in the context of building automated proof checking regimes. The other perspective is that of the semanticist who looks at mathematical texts as one genre among many, but who believes he can learn from a closer study of this particular genre, precisely this is prose whose content has the precision indispensable to communication of the logically complex contents that mathematics is about, and that uses natural language for this purpose without compromising its commitments. There remain good reasons for both linguists and mathematicians to engage in this study jointly in their complementary capacities.

6.4 Situation Semantics

There is one form of context dependence that I want to mention here for two reasons. In the first place it is a form that was very influential at the time when it was introduced (the late seventies and early eighties). But the approach in question has also been an important influence on my own under-

standing of logic, and thus of the various perspectives on what is part of logic that I am trying to come to grips with in this section. What I am referring to is the Situation Semantics of Barwise and Perry (Barwise & Perry 1983). There are many aspects of Situation Semantics – or better: of Situation Theory, of which Situation Semantics is only one part, if perhaps the most prominent one – whose importance for a general connection of logic is as undeniable now as it was in the eighties and nineties, for instance Situation Theory's treatment of the logical paradoxes, its use of non-well-founded Set Theory (and more generally of its theory of non-well-founded situations) or Situation Theory's Channel Theory. I will have nothing to say about any of those aspects here; and if any of the following remarks may come across as unfairly critical, it should be kept in mind that none of these aspects are their target.

The aspect of Situation Semantics that I have just referred to as a 'form of context dependence' is best explained in terms of Barwise and Perry's use of the term *Austinian proposition*. According to B&P the propositional content of the utterance of a declarative sentence is determined in part by the situation – some particular, often narrowly circumscribed part of the actual world, or perhaps also of some possible or imagined world – that the utterance aims to be about. What that situation is, is typically not made explicit by the uttered sentence and will be clear to the interpreter (if it is) because of the context in which the utterance is made. The Austinian proposition expressed by the utterance is the content it has in virtue of being about the targeted situation. This doesn't tell us what an Austinian proposition exactly is – to make this explicit is Situation Semantics' central task and there can be no question here of filling all the necessary details. But at least this much can be said: the targeted situation – the *described situation* in B&P's terminology – determines the truth conditions of the utterance: the utterance counts as true insofar as it is *a true description of* this situation.

But what exactly does it mean for an uttered sentence to be true *of* the situation that its utterance targets? Without further explanation that isn't clear. In fact, there are two kinds of questions here that need to be resolved. The first has to do with the *partiality* of situations. A situation may leave many elementary questions unanswered about the individuals that are part of it, and it will tell us nothing about individuals that are not part of it in the first place. For instance, Bill and Mary may both be part of a situation s, but the situation may contain no information about whether Bill loves Mary or whether Mary loves Bill. And if Susan is not part of the situation, then it won't tell us anything about her, and so in particular nothing about her

116

relations to Mary or Bill. This (obviously) entails that a simple utterance such as 'Mary loves Bill' or 'Susan is Canadian' will be neither true nor false in s, so 'ϕ is true in s' will be a partial relation, which is undecided for some combinations of ϕ and s.

The second question has to do with how many different situations may be involved in the semantic analysis of a single utterance. For many simple categorical utterances the idea of a particular situation they target and of the Austinian propositions they express on account of that are plausible. But there are also many utterances we make for which this is not so clear. For instance, aren't there conditionals where the antecedent targets one situation and the consequent another (somehow related) situation? And is there then in addition a situation targeted by the conditional as a whole? Also, there are, according to Situation Semantics, certain linguistic constructions that denote situations, in much the same way that proper names and other noun phrases denote individuals of various kinds. Situation Semantics' most prominent instance of this was mentioned in Section 4.6: the infinitival complements of naked infinitive perception sentences, such as *Mary cycle past* in 'John saw Mary cycle past' (14) of Section 4.6), are supposed to denote 'scenes', a special type of situation; and the perception verbs of such sentences are to be analyzed as relations between the denoted scene and the perceiver.

In view of all this the situation-semantic analysis of an utterance may involve several situations, the situation targeted by the utterance and various situations denoted or otherwise involved in the interpretation of constituents of the uttered sentence. For many complex sentences this means that the general assumptions and methods of Situation Semantics allow for a range of possible analyses. And partiality will play a large part in all or many of them.

The reason for including Situation Semantics in these reflections is its partiality. So let us for the time being focus on the situations targeted by utterances. To see what form or forms situation-theoretic partiality can take we need to (i) have a precise characterization of situations and (ii) decide on a language to which the situation-semantic analysis, with its inherent partiality, is to be applied. The nature and properties of situations are matters that have been dealt with extensively and explicitly in Situation Semantics, but for our present purposes the following will suffice. Each situation, we will assume, determines a set of *infons*, where an *infon* is a bit of information that says of some particular individual that it does, or does not have, a certain property, or of two or more individuals that they stand, or do not

117

stand, in a certain relation. The information carried by the situation is the total information given by all the infons the situation determines. It will be convenient to identify situations with the sets of infons they determine and that is what I will assume. A situation s will be *consistent* so long as for no n-place predicate and tuple $<a_1,...,a_n>$ s contains an infon to the effect that Q holds for $<a_1,...,a_n>$ and an infon to the effect that Q doesn't hold for this tuple of individuals.

The choice of a language for which a situation-semantic account of truth and falsity is to be made explicit, will depend on variety of considerations, the intended application (to some particular fragment of some natural language), or a more abstract investigation of how Situation Semantics can or should work, using a streamlined formal language to show how the approach works in a more abstract setting. For purposes of this second kind, a natural choice is first order Predicate Logic, given its standing as the logical formalism par excellence, of which everybody with an interest in logic and semantics knows the syntax, and also its standard bivalent semantic model theory, with which a situation-semantic treatment can be easily compared. We too will focus on first order Predicate Logic to start with.

Since an infon can give information about the presence or absence of a property or relation, this is therefore also true of the situations that contain infons as their constituents. Moreover, situations will in general, we noted, leave some such information undecided. To do justice to this three way division – true, false, undecided – a situation semantics for a given language has to be *two-sided*, i.e. it has to articulate both what it is for a sentence to be true of a situation (or for a situation to *support* the sentence, as the official terminology of Situation Semantics has it) and what it is for the sentence to be false of a situation (or for the situation to support its negation), with the possibility that the situation may neither support a sentence nor support its negation. (In practice this means that a situation semantics for our chosen language has to take the form of a simultaneous recursion of the two notions 'true of' and 'false of'.) There is a straightforward analysis for negation in such a set-up – for any sentence ϕ, $\neg\,\phi$ is true/false of situation s iff ϕ is false/true of s – and also for conjunction: $\phi\,\&\,\psi$ is true of s iff both ϕ and ψ are true of s, and $\phi\,\&\,\psi$ is false of s if at least one of ϕ and ψ is false of s. As far as the sentential part of Predicate Logic is concerned, we can leave it at that, and define the remaining connectives, disjunction, material implication and the biconditional, in terms of $\neg$ and $\&$. Let us assume that these three remaining connectives are dealt with in this way.

That still leaves the quantifiers. To state when $(\exists x)\phi$ or $(\forall x)\phi$ is true/false of s we have to specify the quantification domain of the quantifier. But what do the quantifiers range over in relation to situation s? What domain does s have to offer? One possible answer to this is that the *Domain of s*, DOM(s), is the set of all and only those individuals that are constituents of infons belonging to s. Let us adopt this answer at least for now and take this Domain to be the quantification domain of the quantifiers of our Predicate Logic language. We can then stipulate, as part of our situation semantics for Predicate Logic, that $(\forall x)\phi$ is true of s iff for every individual a in DOM(s) ϕ is true of s when a is assigned to x and that $(\forall x)\phi$ is false of s iff there is an individual a in DOM(s) such that ϕ is false of s when a is assigned to x (and likewise, mutatis mutandis, for $(\exists x)\phi$). This completes our two-sided truth definition for Predicate Logic.

Suppose a set $\mathcal{S}$ of possible situations is given. Then we can define a relation of logical consequence between sets of sentences (the premises) and sentences (the conclusion), in familiar ways. Notoriously, there is a complication here that arises for partial truth definitions but not for the bivalent truth definitions of classical logic: There are different ways of formulating what is one and the same consequence relation for classical logic, but that give different consequence relations with partial truth definitions. Here we only consider the most common formulation: $\Gamma \vDash \phi$ ('ϕ is a logical consequence of Γ') iff for any situation s in $\mathcal{S}$, if every γ in Γ is true of s, then so is ϕ.

The logic of this combination of partial truth definition and definition of logical consequence is reasonably well understood. Provided $\mathcal{S}$ is rich enough[69], this logic is the one known as Kleene*.[70] This is a logic that is considerably weaker than classical logic; for instance, it has no logical theorems (i.e. $\emptyset \vDash \phi$ for no ϕ, not even for a formula like $\phi \rightarrow \phi$). A number of different sound and complete proof (in different proof-theoretical formats) systems are known for Kleene*.

---

[69]For instance, when for some infinite set of individuals a_1, a_2, a_3,... and each n-place predicate Q of the language there is an infon to the effect that Q holds for some tuple $<a_{i_1}, ..., a_{i_n}>$ and an infon to the effect that Q does not hold for $<a_{i_1}, ..., a_{i_n}>$ and $\mathcal{S}$ consists of all sets of these infons.

[70] In unpublished work from the early eighties (Kamp 1983) a formal semantics is developed for naked infinitive perception sentences, which uses the partial truth definition described above to define the 'true of' and 'false of' relations between the naked infinitive complements of such sentences and the situations denoted by these complements. As expected the logical inference relation between naked infinitive perception sentences that results from this proposal is then governed by the Kleene* logic as well.

But there is much more to Situation Semantics. One way to get a glimpse of further issues is to explore possible connections between Situation Semantics and DRT.

6.4.1   Situation Semantics and DRT

An opportunity to look at possible relationships between Situation Semantics and DRT arose in the early eighties when a small group of people spent the academic year '81-'82 at the Stanford Center for Advanced Study in the Behavioral Sciences, with some of us representing Situation Semantics (among them in particular Barwise and Perry themselves) and one (me) representing DRT. Looking back to that year (which in all sorts of ways was wonderful and intellectually enormously enriching), I have often felt that the opportunity for comparing or relating Situation Semantics and DRT wasn't optimally exploited. I think the reason was that we all thought of Situation Semantics and DRT as competitors in some way – as competing new accounts of the role that partiality plays in semantics. That perception stood in the way of arriving at a synthesis between the two, which I believe could have been a useful and illuminating contribution on top of what Situation Semantics and DRT had to offer on their own and one that also could have thrown new light on each of them.

What substance is there to the thought that Situation Semantics and DRT are competing approaches to partiality? That partiality is an essential feature of Situation Semantics is clear: As we have seen, a situation can support a statement and it can support its negation, but it can also fail to support either. But what does DRT have to say about partiality? That is not at all obvious. For one thing the standard semantics for DRS languages is just like the model-theoretic semantics for classical logic; there is no kind of partiality there. That doesn't mean that a partial semantics for DRT cannot be given as well, but there is nothing in DRT as such that either encourages or favors such a semantics.[71]

On the other hand, however, there is a connection with partial model structures that played an important role in early discussions of DRT: Every verifying embedding f of a Simple DRS K (a DRS all of whose Conditions are

---

[71] In current work I have formulated a partial semantics for DRS languages with vague predicates, along the lines of the Supervaluation Semantics proposals of (Fine 1975) and (Kamp 1975). There isn't much that sets this partial semantics for DRS languages apart from the corresponding semantics for languages of first order Predicate Logic (though there are some details that require special attention).

atomic, that is, are predications of the form 'Q($x_1,...,x_n$)' where Q is some n-place predicate of the language) in a model M carves out from any M a *partial sub-model M'* defined as follows:

(16) a. The Universe of M' is the Range of f (i.e. the set of those elements from the Universe of M that are values that f for one or more discourse referents);

   b. For every n-place predicate Q of the given DRS language L the extension of Q in M' $[[Q]]_{M'}$ is the set of all n-tuples $<a_1,..,a_n>$ from the Universe of M such that for some drefs $<x_1,..,x_n>$ the Condition Q($x_1,...,x_n$) belongs to the Condition Set of K and for i = 1,..,n $a_i = f(x_i)$.[72]

It is easy to see (i) that f verifies K in M' and (ii) that M' is the smallest sub-model of M in which K is verified by f, and a minimal sub-model of M in which K is true. Intuitively, M' is the 'image' of K under f, and K can be seen as a description, modulo f, of M' in DRT terms.

The models that verifying embeddings of simple DRSs carve out of the models in which they verify them can be seen to bear some resemblance to the situations of Situation Semantics. But formally they are just models for the language in the sense of classical model theory, which determine 'complete theories' of the language – sets of sentences such that for each sentence S of the language exactly one of S and not-S belongs to the set – as the sets of sentences true in them. In this regard M' is no different from M, from which M' is 'carved out'. We come a little closer to the situations of Situation Semantics by considering *literals-DRSs*. A *literals-DRS* is a DRS whose Conditions are all *DRS-literals*. A *DRS-literal* (or, more simply, *literal*) is a DRS Condition that is either (i) atomic or (ii) a *Simple Negation Condition*, a Negation Condition of the form ¬ γ, with γ an atomic Condition.[73]

---

[72]Note that a partial sub-model M' thus carved out from larger model M is in general not a sub-model of M according to the standard model-theoretic definition of 'sub-model'. This is because the extension of a predicate Q in M', while always a subset of Q's extension in M, need not be identical with the restriction of Q's extension in M to M''s Universe, as the definition requires; the extension of Q in M' may be a proper subset of this restriction.

[73]Negation in DRT is standardly represented by Conditions of the form ¬K, where K can be any DRS. When this DRS K is of the form $<\emptyset,γ>$ where γ is an atomic condition, then it is convenient to use the simplified notation ¬ γ.

Let K be a literals-DRS and f a verifying embedding of K in the model M. Then f carves out of M a *two-sided partial model* M'', defined as in (17):

(17) a. The Universe of M'' is the Range of f (i.e. the set of those elements from the Universe of M that are values that f for one or more discourse referent) (see above);

b. For every n-place predicate Q of the given DRS language L the *positive extension of* Q *in* M'', $[[Q]]^+_{M''}$, is the set of all n-tuples $<a_1,..,a_n>$ from the Universe of M such that for some drefs $<x_1,..,x_n>$ the atomic Condition $Q(x_1,...,x_n)$ belongs to the Condition Set of K and for i = 1,..,n $a_i = f(x_i)$.

c. For every n-place predicate Q of the given DRS language L the *negative extension of* Q *in* M'', $[[Q]]^-_{M''}$, is the set of all n-tuples $<a_1,..,a_n>$ from the Universe of M such that for some drefs $<x_1,..,x_n>$ the Simple Negative Condition $\neg\, Q(x_1,...,x_n)$ belongs to the Condition Set of K and for i = 1,..,n $a_i = f(x_i)$.

The partial models carved out by literals-DRSs resemble the situations of Situation Semantics closely. (We can think of the elements of the positive and negative extensions for the predicates of the language as infons.) And once more the partial model M' that a verifying embedding f of a DRS K carves out from some larger model M is the smallest partial sub-model of M in which f verifies K. But, again, this too isn't something that is specific to DRT: The translations of literals-DRSs into formulas of Predicate Logic with its standard syntax carve out partial models from larger models just as the literals-DRSs of which they are the translations.

But there is also another connection between DRT and Situation Semantics. It is most clearly illustrated by the conception behind DRT's treatment of conditionals. As noted repeatedly at earlier points in these reflections, one of DRT's original selling points was its treatment of donkey sentences, among them sentences of the form 'If a farmer owns a donkey he beats it', which I give a label here for reference below.

(18) If a farmer owns a donkey he beats it

The guiding idea behind the DRT analysis of such sentences is that their antecedent describes a type of 'state of affairs' or 'situation' about which the consequent of the conditional then has more to say. And – this was the core to DRT's account of donkey anaphora – in phrasing this further information the conditional's consequent can exploit the way in which the antecedent has

specified the situation type about which the consequent makes its further comment. A paradigm of this exploitation by the consequent of the wording in the antecedent is the use of pronouns in the consequent that are anaphoric to NPs occurring in the antecedent, and in particular to indefinite NPs: the indefinite NP introduces an entity into the situation described by the antecedent and the consequent can use a pronoun to refer to that entity in the comment it makes.

In the light of this informal background, the analysis that DRT offers of donkey sentences of the *if..., then* form, such sentences can be conceptualized as universal quantifications over situations, with the conditional's antecedent specifying what kind of situations are at stake, and its consequent stating some property or properties of the such situations. (And that in general, not only in those cases where pronominal anaphora puts our noses to it.)

Here we seem to see a clear point of contact between DRT and Situation Semantics: At least form an informal point of view DRT's analysis of conditionals involves situations, as presented by the antecedent of the conditional and described in further detail by the consequent. But we should be careful not to overstate the point. After all, the standard semantics of DRT, of which the treatment of conditionals I just described is an integral part, is a standard bivalent one, in which each proper DRS is either true in any model M or false in it, and in which embedding functions either verify or do not verify DRS Conditions and improper DRSs in models, or don't, again without the possibility of a *tertium*.[74] This shows for one thing that it is possible for a semantic analysis to 'involve partiality' in the sense of involving situations without having to involve partiality in the sense of a partial notion of truth.

On the other hand it is also possible at least in principle to embed the DRT account of conditionals just sketched within a situation semantics for DRT. Suppose that such an account, in which DRSs and DRS Conditions are evaluated for truth and falsity in partial models, is in place and that

---

[74]It should also be noted that anaphora resolution, as for the pronouns *he* and *it* in the donkey sentence 'If a farmer owns a donkey he beats it' repeated in the one but last paragraph, is something that must be executed at the level of the interpretation of antecedent and consequent as descriptions of *situation types*, and not at that of the situations instantiating them. On the usual interpretation of donkey sentences like the one above as quantifying over possibly multiple situations of Pedro owning a donkey, the comments made by the consequent of the sentence on any one of these situations, will be about the donkey that is part of that situation, but that is guaranteed by the connections between the types that the anaphoric links establish. I'll come back to this point below.

the DRS Conditional (19) is being evaluated in partial model M under the embedding f which is defined for the free drefs of (19) but for no others.[75]

(19) $K_1 \Rightarrow K_2$

Then we could stipulate:

(20) a.   *f verifies* (19) in M iff for every extension $g \supseteq_{U_{K_1}} f$ that verifies K_1 in M there is an extension $h \supseteq_{U_{K_2}} g$ that verifies K_2 in M.

   b.   *f falsifies* (19) in M iff there is some extension $g \supseteq_{U_{K_1}} f$ that verifies K_1 in M every extension $h \supseteq_{U_{K_2}} g$ falsifies K_2 in M.

When M is partial then (20.a) and (20.b) may leave a truth value gap in the sense that f may neither verify nor falsify (19) in M. (For an example suppose that K_2 contains the condition Q(x), that Q is not mentioned in K_1 and that both the positive and the negative extension of Q in M are empty.)

Suppose now that K_1 and K_2 are both literals-DRSs. Then we can also reformulate (20) in terms of situations. Suppose that g is any embedding function in M such that $g \supseteq_{U_{K_1}} f$ and that g verifies K_1 in M. Then g will carve out a partial model M' from M and this model M' determines in its turn a situation $s_{1,g}$.[76] Suppose that h is an embedding in M defined on U_{K_2} that extends g – so $h \supseteq_{U_{K_2}} g$ – such that h verifies K_2 in M. Then h too carves out a partial model out of M, which determines a corresponding situation $s_{2,h}$ in the same way as g and K_1 determine $s_{1,g}$. Then $s_{2,h}$ will be an extension of $s_{1,g}$ (formally, $s_{2,h}$ will be a superset of $s_{1,g}$). Under these conditions we can state (20.a) also as in (21). (A similar reformulation is possible for (20.b).)[77]

---

[75]Recall: the free drefs of a conditional DRS Condition such as (19) are (i) the free drefs of K_1 together with (ii) the free drefs of K_2 that do not occur in the Universe of K_1.

[76]The infons of $s_{1,g}$ are those corresponding to the Conditions of K_1: the 'positive infons' are given by atomic Conditions Q($x_1,...,x_n$), together with the argument tuples $<g(x_1),..,g(x_n)>$ and the 'negative infons' by the Simple Negation Conditions $\neg$ Q($x_1,...,x_n$), together with the argument tuples $<g(x_1),..,g(x_n)>$.

[77]The notions involved that have not yet been defined ought to be guessable from their names. But for all events and purposes, here are their formal definitions: (1) A situation s is *included in* the partial model M iff (i) all individuals occurring in s belong to U_M; (ii) for every positive infon in s, to the effect that the n-place predicate Q holds of the argument tuple $<a_1,...,a_n>$, $<a_1,...,a_n>$ belongs to the positive extension of Q in M and (iii) for every negative infon in s, to the effect that the n-place predicate Q does not hold of the argument tuple $<a_1,...,a_n>$, $<a_1,...,a_n>$ belongs to the negative extension of Q in

(21) *f verifies* (19) in M iff for every situation s included in M and embedding function $g \supseteq_{U_{K_1}} f$ in the universe of s that verifies K_1 in s and where s is a minimal such situation, there is a situation s' included in M which extends s and which is such that some extension $h \supseteq_{U_{K_2}} g$ in the universe of s' which verifies K_2 in s'.

It is not hard to see, moreover, that the references to embedding functions in (21) can be omitted without changing the content provided we define '*s verifies K*' as: 'there is a function f from the Universe of K into the Universe of s which verifies the Conditions of K in s' (and adopt an analogous definition for '*s falsifies K*', for the falsification condition corresponding to (20)). With this definition for '*s verifies K*' (21) can be turned equivalently into (22).

(22) *f verifies* (19) in M iff for every minimal situation s included in M verifying K_1 there is an extension s' of s that is included in M and that verifies K_2 in s'.

Note that the formulations in (21) and (22) differ from our informal description of the semantics of conditionals according to DRT in that according to these formulations the contribution made by the consequent of the conditional is described as pertaining to certain *extensions* of the situations posited by the antecedent of the conditional, rather than as a comment on those situations themselves. I'll come back to this point. (There is also the restriction to minimal situations verifying the antecedent of the conditional in (21). This restriction is needed to exclude situations that contain in addition to the information required for the verification of K_1 also information that is spurious for the verification of K_1 while blocking the possibility of extending the situation to one verifying K_2.)

As a final comment on this discussion of possible treatments of conditionals in DRT: We started out with the DRT treatment of English conditionals like (18). But then I switched to the conditional DRS Condition given in (19).

M. (2) A function g from the Universe of a literals DRS K into the universe of a situation *s verifies K in s* iff for each atomic condition $Q(x_1,...,x_n)$ of K, s has an infon to the effect that Q holds of $<g(x_1),..,g(x_n)>$ and for each Simple Negation Condition $\neg Q(x_1,...,x_n)$ of K, s has an infon to the effect that Q does not hold of $<g(x_1),..,g(x_n)>$. (3) s is a *minimal support situation for K under g* iff (i) (2) holds for s, g and K and (ii) there is no situation s' properly included in s (i.e. no s' that is a proper subset of s) such that (2) holds for s', g and K. Also '*s extends s*' for situations s and s' simply means that s' is a subset of s.

I made this switch without comment, hoping that no one would balk at it without explicit acknowledgment of it. But if a justification is needed, then, on pain of repeating the obvious, it is this. DRT is intended as a 'logical form formalism': it provides semantic analyses for linguistic constructions, and for natural language fragments containing those, via a construction algorithm that turns the relevant bits of natural language into a DRS, and it is then the model-theoretic semantics for those DRSs that turns the conversion into a genuine semantics for the represented bits of language. That is why it is both legitimate and necessary to focus on the DRT representations of the relevant bits of language – English conditionals, in the present context – rather than on the bits of language themselves. Note also that issues of anaphora resolution of the kind exemplified by donkey anaphora, can and have to be resolved at the level of DRS construction; they are no longer an issue when we consider conditional DRS Conditions like (19). Note well that any application of Situation Semantics to a sentence like (19) will have to deal with such anaphora problems. If DRT is not used as a logical form formalism as part of such an application, then these problems will have to be dealt with with in some other way.

So much for now about conditionals in DRT, with our without Situation Semantics. Note that a like-spirited analysis of conditionals is also possible within Situation Semantics.[78] Earlier, in Section 6.4, we saw a partial truth definition for Predicate Logic. That definition could be recast without difficulty as one in situation-theoretic terms, using situations where the given formulation uses partial models. That definition treats conditional formulas as material conditionals. That isn't wrong, but the general ideas of Situation Semantics suggest other, more 'intensional' interpretations of conditionals. For convenience, and in the light of the observation of the last paragraph, we will stick with conditionals in the DRT format of (19).

The formulation of the verification conditions for (19) in (22) (together with the corresponding falsification conditions) seems to me a plausible proposal for an account of the truth conditional content of conditionals from a situation-semantic perspective. But we should not forget that the formulation presupposes that the antecedent and consequent K_1 and K_2 are literals-DRSs. If the formulation is to be extended to conditionals with more complex antecedents and/or consequents, then more work is needed, and as things stand I have no clear idea how any of this might go.[79]

---

[78][**Check for proposals along this line in the SS literature**.]

[79]This isn't a reason for dismissing the formulation we have given as unfinished and

126

Suppose we accept (22) and its negative counterpart as specifications of the verification and falsification by situations of conditionals of this special form. From the perspective of Situation Semantics mentioned above, there is then a further question about the general truth conditional framework within which these specification should be integrated. A central feature of that framework, it was claimed in Section 6.4, is that each utterance comes with an independently identified targeted situation that is one of the constituents of the Austinian proposition expressed by the utterance. Suppose that (19) is the semantic representation of the utterance of a conditional sentence of which (19) has been computed as semantic representation and that s_0 is the situation targeted by this utterance. What then is the relation between s_0 and the situations involved in the semantic analysis of the conditional (represented by) (19)?

Here is a plausible answer to that question. Our situation-semantic partial version of the truth conditions of (19) assumed that (19) would be evaluated within a partial model M. The situations referred to in (22) are all situations 'included' in this partial model. But nothing was so far said about where M is supposed to come from. What determines M – in preference to any other partial model – for the utterance of a conditional sentence represented by (19)? Situation Semantics, with its view that part of making an utterance is targeting some particular situation which the words of the utterance then describe, has an obvious answer: M just is that targeted situation. (Recall the observation we have made use of repeatedly in he discussion above, that partial models and situations are isomorphic structures.)

Let us then, in what I take to be the spirit of Situation Semantics, reinterpret our analysis (22) of conditional represented in the form (19) as follows:

An utterance u of a conditional represented in the form (19) targets a situation s_0 (corresponding to the model M spoken of in (22)). The verification conditions for u are then given by the following modification of (22), where

to be shelved until a more general version of it is available. Work in Philosophical Logic from the past half century has made it increasingly clear that conditionals with logically complex antecedents or consequents pose problems that are often quite different from those that arise in connection with simple conditionals, and that when complex conditionals are considered capable of meaningful semantic analysis at all, they should be considered separate chapters of their own. Even if (22) is usable only in connection with conditionals of the special kind for which it is formulated, even as a formulation of the truth conditions for conditionals of this special form it may have its merits.

the inclusion relation between situations is now plain set-theoretic inclusion. (A similar modification is of course required for the falsification of (19) in s_0.)

(23) *f verifies* (19) in s_0 iff for every minimal situation s included in s_0 verifying K_1 there is an extension s' of s that is included in s_0 and that verifies K_2 in s'.

How good an analysis of utterances of conditional sentences is this? First, a point that counts in favor of an analysis along these general lines. As we have been interpreting conditionals in this section they involve some kind of universal quantification over cases; in (22) and (23) these are the values of the situation variable s. And quantification involves – invariably or almost invariably – a Domain restriction, which may be wholly or partly made explicit in the sentence itself, but also may have to be, wholly or partly, reconstructed from context. To focus more clearly on this quantification dimension of conditionals, consider the following sentences.

(24) a. If/When/Whenever Pedro buys a donkey he doesn't like, he sells it a week later to someone he has found on eBay.

b. If/When/Whenever Pedro doesn't like a donkey he has bought, he doesn't show it to anyone.

c. If/When/Whenever Pedro bought a donkey he didn't like, he sold it a week later to someone he had found on eBay.

There are several aspects of these sentences that will be relevant for what I want to say about them. The first has to do with the conjunction governing the subordinate clause. The three versions of each of the sentences in (24.a) – beginning with *If*, *When* and *Whenever*, respectively – are not equivalent, and the same is true for the corresponding three in (24.b). The *whenever*-sentences unambiguously express universal quantification. The *if*- and the *when*-sentences are ambiguous between the sense of a quantified conditional (on which they are for practical purposes equivalent to the *whenever*-sentences) sentences), but also an episodic reading.[80] In what follows I will focus on the *whenever*-sentences, with their unambiguously quantificational

---

[80]In the episodic readings of the *when*-sentences the subordinate clause beginning with *when* can be understood as fixing the time at which the main clause is true (i.e. not long after Pedro's purchase of some donkey). For the episodic reading of the *if*-sentences episode in question must have been fixed by context; the sentence then says of that episode that it was as described by the *if*-clause, then it was also as described by the consequent.

meaning.

The sentences in (24.a) share with those we have discussed up to now that they are all in the present tense. For the *whenever*-sentence in (24.a) this entails that an utterance of it makes a statement about situations verifying the antecedent and the consequent that are temporally situated within some interval surrounding the utterance time. But unless a particular period is mentioned explicitly the duration of that period tends to be very underspecified. For the sentences in (24.b) this is different. Utterances of past tense conditionals are about some period preceding the utterance time. Sometimes this period is mentioned explicitly in the sentence, as in (25). But it is also common that the past period that is intended must be retrieved from the context in which the utterance is made.

(25) Last summer, whenever the weather was good, Mary went for a swim in the afternoon.

If the propositional content of utterances of conditionals is to be analyzed as in (23), then one thing that we should expect from the 'described situation' s_0 is that it specifies the period of time within which the situations s and s' should be temporally included. Suppose for instance that an utterance of the *whenever*-sentence of (24.b) – henceforth '(24.b)(*whenever*)' – targets a situation according to which the past tense period has been from the time Pedro started using eBay until the time he sold his farm. What should we assume the situation s_0 to be which imposes this constraint on the quantification expressed? Everything that happened in the course of that period, or only those things that had something to do with Pedro? Or only those things that had to with his dealings with donkeys? Or ... ?

This is a weakness of the situation-semantic account as I have presented it. If situations are sets of infons, what situation thus constituted would be the one that comes with an utterance of (24.b)(*whenever*)? Intuitively what the interpreter of such an utterance needs is the period during which the quantification expressed by *whenever* is supposed to quantify over – nothing less, but also nothing more.[81]

---

[81]A situation-theoretic treatment of sentences like those in (24) and (25) also raises another matter that I have kept under wraps up to now. Things change and many of the things we say are about how they change. The sentences in (24) are examples of this. They are about Pedro having bought one or the other donkey. That is is a state of affairs that holds at some times – the times after he has bought the donkey in question and before he has sold it again – and not at others (those before he bought the donkey an those after he

This points at a much more general problem. If the notion of a 'situation', as we have been using it up to now, is to serve all the different purposes thus far mentioned, then it looks like there is a large variety of different kinds of entities that must be included within this category. And the question is then whether by allowing for this wide a variety of types of entities to qualify as kinds of situations, the notion of a situation isn't hollowed out to a point where it ceases to be of explanatory use.

This is also the right point to mention another difference between Situation Semantics and DRT, which is especially salient in relation to sentences like (24) and (25). DRT has made use of times and eventualities (events and states), as entities that can be represented by discourse referents, since its very beginnings (Kamp 2017). This applies to all sentences, and it applies in particular to conditionals like, for instance, (24.b)(*whenever*). According to such a treatment, both the antecedent and consequent of the sentence are represented as descriptions of events and the analysis of the conditional as a whole is that for every event of the kind described by the antecedent, and within the period of time (or set of occasions determined in some other way) there is an event of the type described by the consequent, that stands in a certain relation to the first event (in this case the relation of following the first event within a week). That is, in this instance, and in fact more or less across the board, DRT appeals to eventualities – events as here, or states in the representation of other sentences – where Situation Semantics as I have presented it appeals to situations.[82]

---

sold it). The infons we have been taking about so far do not seem to provide room for this kind of dependence on time. If they are to do this, then there must be, at a minimum, a way of distinguishing between an infon to the effect that the pair $<a,b>$, where a is Pedro and b some particular donkey, belongs to the extension of *own* at time t and an infon to the effect that the pair $<a,b>$ belongs to the extension of *own* at time t'. A simple way to create this possibility is to treat those predicates whose extensions can vary with time as having an additional argument for a time. For instance, *own* will be treated as a 3-place predicate with one argument for the owner, one for the entity owned and on for the time of owning. (I do not advocate this as a particularly good solution. But it will do just about well enough in the context of the present discussion.)

[82]As this informal description of the analysis of (24.b)(*whenever*) makes clear, it and the other sentences in (24) also reveal another tension with the analysis in (22). In these conditionals antecedent and consequent clearly describe distinct events which occur at temporally separate moments. It seems a little counterintuitive to analyze the consequents of these conditionals as about situations that are extensions of the situations described by the antecedents. The extension would have to be mereological sum of the event described by the antecedent (or, more correctly, what corresponds to it situation-wise) and the event (or its situation-theoretic counterpart) described by the consequent. Formally there may be no serious problem with this: Many versions of mereology are very liberal with regard

This adds a further complication to the comparison between Situation Semantics and DRT: In many cases where they appear to offer competing analyses for the same sentences or utterances, the difference may in the end be more one of terminology than of substance. This question must be asked in particular in relation to the use that has been made in DRT of states. States were originally introduced into DRT on a par with events, as the 'eventualities' described by state verbs, like *own* or *like*, just as events were assumed as the entity described by event verbs, like *buy* or *find*, but also as part of dealing with aspectual differences between, for example, the Passé Simple and the Imparfait in French or that between the Simple Pasts and Past Progressives of English. But then, a need for something like states manifested itself also in other contexts, for instance in dealing with the combination of tense and negation and – more directly relevant to the examples we have been looking at – sentences involving quantification. For instance, the DRT analysis I personally favor of (24.b)(*whenever*) assumes that the quantification over events that the sentence expresses serves as the specification of a 'quantification state' – one which holds over the time of its duration by virtue of the quantification being true over that period of time (Kamp 2019*b*). An utterance of the sentence is then true if a state of this type holds throughout the period that the interpreter must retrieve from its context. Likewise, the DRT analysis of (25) is that of a quantification over states (of the weather being good) and that claims of such states that they were accompanied by an event of Mary having a swim. And the claim made by the utterance is that a state of the type described by this quantification held for the period of the last summer before the utterance was made.

A possible objection to this kind of analysis is that it seems to beg the question what such quantificational states are. There is a good argument that they are more than just the periods over which they hold. For quantificational states can, just as other states and events enter into causal relations, whereas times cannot. But even so it may be felt that a more illuminating ontological clarification is needed, and that as far as this is concerned, quantificational states (and perhaps 'negation states' as well) are not in much better shape than might recharged against some of the situations that figure in certain applications of Situation Semantics.

---

to the formation of sums and without serious danger of thereby running into contradiction. But I want to signal nevertheless that using sums over what intuitively different, temporally and/or spatially separate events seems to me an awkwardness that one might want to avoid. This is a bigger point, I think, than I make it here. But I will say no more about it here.

Before concluding this juxtaposition of Situation Semantics and DRT I should mention one further context in which the controversy between situations and events has come up. In fact, it came up at an early point in time. I mentioned Barwise's paper 'Scenes and other Situations' (Barwise 1981) as presenting one of what many of us took to be one of the strongest arguments to show that situations were needed. The constructions in question were the naked infinitive perception sentences first mentioned in Section 4.6 and about which more was also said earlier in the present section. What hasn't yet been pointed out is that not long after the appearance of (Barwise 1981) the argument was said to be unconvincing because the complements of naked infinitive perception sentences can also be analyzed as denoting events. That would give an equally or even superior way of explaining those properties of such sentences for which Barwise invoked situations, with the additional advantage that events are needed in semantic analysis in any case (Higginbotham 1983). In fact, I – among many others – became convinced that events are plausible and therefore preferable denotations for the the infinitival complements of these sentences, and now feel I should have taken a conscious step to that effect earlier.[83] The question whether a given type of expression should be analyzed as denoting situations – as opposed to, for instance, events – is a question that can be raised not only in relation to the infinitival complements of naked infinitive perception sentences. Perhaps this kind of use of situations can be dispensed with in favor of other types of entities, that can be shown to be needed in other contexts too.

It should be emphasized that isn't necessarily an argument against Situation Semantics as such. Its case for the partiality of truth and its logical

_____

[83]One point in favor of the event analysis of naked infinitive perception sentences is that these sentences appear to be felicitous only when the infinitival verb is an event verb. Consider for instance the sentences in (26).

(26)   a.   She saw him know the answer.

       b.   She saw him be sick.

       c.   She saw him sick.

(26.a) seems irredeemably awkward, and should probably qualify as ungrammatical. And the sentences (26.b) and (26.c), if not ungrammatical, cannot be interpreted in the way in which a naked infinitive perception sentence should. (26.b) can be understood only in a sense which involves coercion, in the same way that progressives of state verbs or verb phrases need coercion, viz. as saying that she saw him throw up, or something to that effect. (26.c) seems to be possible as a case of depictive secondary predication: She saw him while he was sick. On this reading the sentence is true so long as the subject saw the direct object, and that he was in fact sick when this happened, but without his being sick playing any part in the event of her perception.

implications need not be affected by these criticisms. I will return to this in the conclusions to the discussion of Situation Semantics in Section 6.4. But before we get to this I want to return to a point that was abandoned, as it were, in midstream.

6.4.2 Situations without partiality of Verification and Truth

This section still needs careful checking with the works and authors mentioned.

In (20.a) I stated the verification conditions for DRS Conditions for DRS Conditions of the form $K_1 \Rightarrow K_2$. As intended these were just one half of a two-sided definition of verification and falsification in a partial model M. But it is also possible to assume that M is a classical bivalent model and in that case (20.a) should be read as giving the full statement of a bivalent definition of verification in M, and with that of a definition of truth in M for Conditions without free drefs. That wasn't the interpretation of (20.a) we wanted and from that point onwards the focus has been on two-sided definitions. That included in particular the interpretation of situation-theoretic formulation of the verification clause in (25) (and the corresponding truth clause for conditional Conditions without free drefs); this clause too was just one half of a pair of complementary verification and falsification clauses.

But (25) too can be read in a different way. Whether or not we think of the verification relations between situations and sentences as partial, in the sense that it is possible for a situation s to neither verify nor falsify a sentence ϕ, or non-partial, in that s either verifies ϕ or else s falsifies ϕ and therewith verifies $\neg\phi$, it will be the case that either s verifies ϕ or s doesn't verify ϕ. That makes it possible to take (25) also as giving exhaustive verification conditions for the sentences in question. And if we do that, we get a bivalent account for the truth conditions of these sentences, even if 'hidden within it' there are occurrences of a verification relation that is partial. The same point applies to accounts of naked infinitive perception sentences that are situation-semantic in that their infinitival complements are taken to be situations that verify the infinitival clauses that denote them.

This is also the sense in which I have read the situation-semantic analyses of donkey sentences and related constructions in the work of Elbourne. (See in particular (Elbourne 2005).) Elbourne accounts for the truth conditions of the sentences he is concerned with by making a thorough-going use of verification by situations, down to the atomic constituents of those sentences.

That leads among other things to analyses of conditionals that are superficially like the one given by (25) (although his analyses are considerably more complex, due to the main point of the account, which is to explain when anaphoric connections between pronouns and putative antecedents are possible). On this reading of (Elbourne 2005) the formal properties of his theory would seem to be that of a many sorted first order theory formulated within Predicate Logic, presumably with a first order Set Theory like ZF, or a substantial part of that as axiomatic core, and with situations as one of the sorts. Among the axioms of the theory are further needed those that capture the general properties and relations of and between situations, as well as between situations and entities of other sorts, and as part of the latter axioms that determine the verification relation between situations and the relevant expressions of the object language. The question what the 'logic' is of such a situation semantics can then be understood in two different ways:

(i) As the 'logic' imposed by this theory on the logical forms that the theory associates with sentences of the object language: for which sentences ϕ and ψ of the object language to which the theory assigns the logical forms $\mathrm{LF}(\phi)$ and $\mathrm{LF}(\psi)$ is the conditional $\mathrm{LF}(\phi) \to \mathrm{LF}(\psi)$ a theorem of the theory? (The relation between a sentence ϕ and its logical form $\mathrm{LF}(\phi)$ must of course be derivable within the theory too.)

(ii) As the 'logic' of the verification relation between situations and object language sentences. Here the logic generated by the theory has to do with what what implications are provable within the theory that are of the form $(\forall s)(\mathrm{Verif}(s,\phi) \to \mathrm{Verif}(s,\psi))$, where again ϕ and ψ are sentences of the object language and 'Verif' is the relation between situations and object language sentence for which the theory provides an axiomatic characterization.

And to iterate a point made towards the end of the last section: It may be that the logic in the sense of (ii) may be partial, in the sense that $(\forall s)(\mathrm{Verif}(s,\phi) \to \mathrm{Verif}(s,\neg\phi))$ is not a theorem for all ϕ (with $\neg\phi$ the negation of ϕ), and yet that the logic in the sense of (i) is bivalent nevertheless – in the sense that for any object language sentence ϕ, and with $\neg\phi$ the negation of ϕ, $\mathrm{LF}(\phi) \lor \mathrm{LF}(\neg\phi)$ is a theorem. (Though this possibility may be blocked for other reasons, e.g. because the object language is expressive enough to allow formalization of its own syntax. See the Section 6.5.)

6.4.3 Conclusions to Section 6.4

This has been a discussion of Situation Semantics that may be qualified as partial in two senses (neither of which is the technical sense in which the term has been used): in that only a very small part of Situation Semantics has been discussed, and probably also in the sense of 'biased', especially where possible connections and comparisons between Situation Semantics and DRT were concerned, even if I have made a genuine effort to be even-handed. But I believe that enough has been said to permit us to draw the following general morals.

• The main tenor of this discussion has been that two motivations went into Situation Semantics: (a) that situations are the denotations of certain constituents of certain natural language utterances; (b) the intuition that, sometimes or always, the question whether what we say is true or false is to be understood as based on limited amounts of information, so that a three-way distinction is needed, between (i) true on the basis of this information, (ii) false on this basis and (iii) neither. While these two considerations are not independent, there is a certain degree of independence between them and it is therefore important to distinguish between them.

• On the one hand, as argued in the last section it is possible to maintain that certain sentence constituents denote situations, and that this denotation relation is grounded in a partial notion of verification of the denoting constituent by the situation it denotes, while yet the evaluation of sentences containing such constituents is bivalent and classical.

• On the other hand it is also possible to adopt a partial truth definition for a language whose semantics does not involve situations as denotations. We haven't discussed this case explicitly and fully. But the verification and falsification conditions for conditional DRT Conditions in (20) can be seen as one pair of clauses of such a definition of verification and falsification, with truth and falsity definable in terms of those.

• It is of course also possible to combine both motivations into a semantic theory in which a partial definition of verification and truth is given for a language some expressions of which have situations for their denotations. In principle there several dimensions of possible variation here and I don't think we even near an account of what the most important options are.[84]

---

[84]This very richness can also become a problem. This has to do in part with a dimension of variation that has not so far been mentioned. It arises in connection with the analysis

• Despite these various difficulties, one undeniable and lasting virtue of Situation Semantics has been to draw our attention to an aspect of semantics that had not been recognized before: The question what it is for a claim or statement to be true on the strength of a limed quantity of information, and what it is for the claim or statement to be refuted by that information quantity, with the obvious contingency that neither is the case, is an important question for logic; and it is that independently of what applications the idea may have in a plausible theory of truth for a fragment of natural language.

---

of particular expressions and the concepts they express, first and foremost with those expressions that have been traditionally thought of as standing for the main operations of formal logic. One recurring difficulty is of the following kind. Suppose we want to specify what it is for a situation s to verify a formula of the form $O\phi$, or $\phi O\psi$, where O is a 1-place or 2-place logical operator. Are we to assume that the evidence for this is to be found within s or can information outside s be relevant as well. One example is that where O is the universal quantifier: is $(\forall x)\phi$ verified by s when it is true for all individuals in s to satisfy ϕ? Or are individuals that are not part of s relevant as well? And if so, which? This is problem that one encounters, in one form or another, as Robin Cooper and I came to realize when in the early nineties we set out on an exploration of possible and logically, conceptually and/or linguistically interesting accounts that might be given of the standard operators of classical Predicate Logic (Cooper & Kamp 1991). We began with negation, assuming that this would be the easiest case and thus the best place to start. But even settling on a plausible situation-theoretic account of negation proved to be remarkably difficult, and a struggle with different formal options without a good way we could see for choosing between them. To the formally inclined logician this may sound like a paradise, but for someone whose concerns are linguistic or cognitive such an 'embarras du choix' is not desirable. A novel semantic framework should lead us to analyses of the notions at issue – here the semantics of the logical operators expressed by 'not', 'or', 'if .. then', 'there is/are' and 'for all' – that once we have been led to them we can recognize as correct; what we do not want is a formalism that presents us with options between which we have to make a choice, but where we cannot see conclusive independently motivated reasons for the choice we should make. This was one reason why Robin and I never got beyond the study of negation. (Another reason was that because of the different options we encountered in our investigation of negation even that one notion took us a good deal longer than we had expected.) We had no reason to expect that things would be simpler for the other logical operators, so the project came to look much more daunting than we had expected it to be initially. Moreover, the semantic analysis of more than one operator should come with an investigation of the ways in which the two or more operators for which separate analyses have been obtained would semantically and logically interact. When one has competing analyses for the various operators between which one feels one cannot make a well-motivated choice, the interaction problem quickly becomes very hard to manage, since the problem has to be considered for each combination of operator analyses that are still on the books. I don't think Robin and I ever decided consciously that there was no point in pursuing these matters further. But the enterprise seemed to be headed in a direction where it would be likely to collapse under its own weight.

• The partiality advocated by Situation Semantics is not the only reason why truth definitions for natural languages and fragments thereof may have to be partial. Another source, mentioned in passing in the last section, are the semantic paradoxes, a central concern for philosophical semantics ever since the semantics became a subject in its own right, but one that most linguists and cognitive scientists as well as some philosophers have a tendency to ignore, if they are aware of it at all. And let me mention here also one other source of partiality, viz. vagueness.

One approach to the logic of vagueness, mentioned in footnote 71, is the so-called Supervaluation approach. At the bottom of that approach is a partial definition of truth and falsity for sentences with vague predicates in models in which those predicates have partial extensions, consisting of a positive and a negative extension that may not meet. Crucial to the Supervaluation approach is furthermore a relation of partial sub-model according to which the partial model M' is a partial sub-model of the partial model M iff the positive and negative extension of each predicate in M' are included in the positive and negative extension of the predicate in M and moreover the Universe of M' is identical with the Universe of M. It is this last condition – the constancy of the Universe – that is to a large extent responsible for the way in which Supervaluation Theory works. In Situation Semantics a corresponding relation plays a somewhat comparable role. This is the relation that holds between two situations s' and s iff s is an extension of s. Since as we have seen there is a close correspondence between situations and partial models (at least when situations are characterized in the way we have assumed), then this extension relation corresponds fairly closely to the partial sub-model relation just defined. But there is one important difference: When thought of as a relation between partial models, the extension relation between situations is not one that holds between s and s' only if the sets of individuals of s is identical with that of s but also when the latter is a subset of the former. That turns out to make an important difference for logical relations defined in terms of these relations. (To my knowledge, what logics are possible when the semantics makes use of the weaker relation, which only requires inclusion for the sets of individuals, is not very well understood.)

• As a final word to this Section: Even within the general setting of Situation Semantics there are many options that can and should be pursued further and that, as far as can be seen at this point, may lead to different logics when 'logics' are identified in terms what are the valid inferences. When Situation Semantics, as one source of partiality is combined with other, conceptually and formally quite different sources, the range of possible options becomes

much larger. Much work remains here, for one thing because so far the different sources of partiality have been the concern of different communities with not much contact between them.

6.5 Natural Language and the Paradoxes

There have been only two references to the term 'paradox' so far in this document, where it is used in the specific sense in which something about is to be said about paradox in this section: in Section 4.2.3 on Lambda DRT and, just now, in the concluding part of Section 6.4.

At least since Tarski's 'The Concept of Truth in Formalized Languages' logicians have been aware that you can't have a truth predicate T, obeying the principle that every sentence S is provably equivalent to the sentence that says that S is true (the sentence 'T($\ulcorner S_0 \urcorner$)'), in a language that has the resources for coding its own syntax (in a way that allows for the construction of a sentence S_0 such that $\vdash S_0 \leftrightarrow \neg T('S_0')$). With the intuitively true principle that every sentence S is equivalent to the claim that S is true this would lead to the logical contraction 'T($\ulcorner S_0 \urcorner$) $\leftrightarrow$ $\neg$T($\ulcorner S_0 \urcorner$).

The conclusion that Tarski drew from this result is that if a language has the resources for this kind of encoding, then a truth predicate T – a predicate of sentences, which is true of the code $\ulcorner S \urcorner$ of a sentence S iff S is true, in the strong sense that the formulas '$\ulcorner S \urcorner \leftrightarrow$ S' are provable for all sentences S of the language – must be excluded from the language. The concept of truth should be banned from the language and exiled to the Metalanguage.

(Montague 1963) shows that this problem is not confined to the concept of truth, but that it extends to other familiar concepts such as knowledge, necessity or belief: When a knowledge, necessity or belief predicate is part of a language capable of encoding its own syntax, in the specific sense that it permits the derivation of the Diagonalization Lemma of (Montague 1963), then contradictions similar to 'T($\ulcorner S_0 \urcorner$) $\leftrightarrow$ $\neg$T($\ulcorner S_0 \urcorner$)' result for these predicates as well. More accurately, contradiction ensue when it is assumed these obey certain general laws that seem intuitively right for these notions (e.g. principles corresponding to the axioms and rules of the modal logic S5). Such predicates too, it may seem a plausible attitude to take, should be subjected to Tarskian surgery and moved to the meta-language.

Montague suggested a different cure. (Montague 1963) proposes that knowledge, necessity, belief and possibly some other notions – notions also governed

by logical principles that entail inconsistency in languages validating the Diagonalization Lemma – should not be treated as predicates of sentences but as sentence operators, as in the standard systems for modal logics, for which this difficulty does not arise.[85] That too is sound advice, but, like Tarski's proposal, it helps only to the extent that there is a way to subject the language that is threatened with inconsistency to the excisions that need to be made.

And for natural languages that is a real problem. First, most (if perhaps not all) human languages have the equipment for coding their own syntax through arithmetization; and English certainly is one of those languages. English has names for all the numbers, even if these names will get quite baroque when the numbers get big, and also words for addition and multiplication, viz. *plus* and *times* and otherwise all the linguistic equipment for making complex number-theoretic statements (even if some of those statements will be quite awkward, recall Section 6.3). And it is part of our understanding of this part of our vocabulary that all these terms have their standard number-theoretic meanings. That is, to put things in model-theoretic terms: the intended models for such a language L are ω-*models*: models that contain the natural numbers (or some structure isomorphic to them) as part of their Universe, that have a predicate 'N' the extension of which consists of those numbers, and in which the numerals and 'plus' and 'times' have their standard extensions. So statements made in L about numbers (as elements in the extension of 'N') will get their standard interpretation, and therewith their 'true' truth value, in any such model. So if logical consequence for L is defined in terms of its ω-models, then every true sentence of arithmetic expressible in L will count as a logical theorem. On these assumptions, everything that is needed for arithmetization of the syntax of L is available in L (even if the sentences required will be quite 'unspeakable' from a practical point of view). English and similar languages are thus among the victims targeted by the semantic paradoxes; and unless something is undertaken towards containing the disease, the victims will remain very sick.

So something will have to be done. But the logician (or the philosopher or the linguist) cannot just change human languages by legislation, for the languages aren't theirs. They are the common property of the communities that use them, and those communities won't let themselves be told what they can

---

[85]In addition, I suppose, one might have sentence predicates for these notions in a metalanguage that does not only talk about truth but also about the propositional attitudes of speakers and so on.

say and what not (not at least when they are being told that they cannot use 'true' or 'necessary' or 'known' any more, since if they aren't careful they may get tied up in paradox and contradiction).[86] The best that it seems one can do is:

(i) to spot as well as we can where in the language the dangerous predicates can do their damage and where they can't, and

(ii) to circumscribe the large areas (i.e. large ranges of sentences and utterances) where the predicates are harmless and where they can therefore be used with all the logical properties that they intuitively seem to have.

In the course of the last half century substantial progress has been made with both (i) and (ii). I am thinking here in particular of the work of Kripke (Kripke 1975) and that of Herzberger (Herzberger 1982), Gupta and Belnap (Gupta 1982), (Gupta & Belnap 1993) on languages with partial truth predicates. In the models for a language L with such a partial truth predicate T, the sentences of L will be among the individuals in the Universe and the T predicate will have a positive and a negative extension, which between them do not exhaust the total set of sentences, and leave the paradoxical sentences – those sentences S for which $T(\ulcorner S \urcorner) \leftrightarrow S$ can be established – out. The methods used by the two approaches are different in that Kripke provides an inductive definition which builds the positive and negative extensions of T up from below, starting with empty extensions computing the truth values of the sentences of L in the given model M on the basis of this assumption, putting the truth sentences into the extension of T in M and the false ones into the negative extension of M and then repeating the procedure with these new

---

[86]Note how drastic both the Tarskian and Montagovian recipes are especially in the case of the predicate 'true' and its complement 'false'. We are using these predicates all the time and naturally, saying things like 'At most three of the sentences that you find on this page are true', 'Not all of the sentences spoken by the president are false', and so on, in which we quantify over sentences while predicating truth or falsity of them. And perhaps some of those sentences have the predicate 'true' in them, and perhaps in iterated ways. Having the word 'true' as a truth predicate in the metalanguage isn't going to do us much good when we want to say such things (unlike the language we speak is to be reconstructed as a hierarchy of meta-languages, but that a notoriously dubious project). And replacing the truth predicate by a truth operator isn't going to help either. That reduction leads to a sentence operator T such that for arbitrary sentences S of the language TS is equivalent to S. De facto that comes to nothing and it gives us no way of making quantified statements of the kind mentioned above. Replacing the truth predicate by a truth operator is, you might say, a radical version of the 'Redundancy Theory of Truth' with all the drawbacks of such a theory.

extensions for T, and continuing in this way until the procedure stabilizes. The Herzerger-Gupta-Belnap procedure approaches the partial extension of T 'from above'. It starts with some bivalent extension for T, compute the truth values of sentences on the basis of that and adopt the result of that as revision of this extension. The revision can be expected to be different, since computing the truth value of a paradoxical sentence S on the basis of the truth value given for it will lead to the opposite result (thus, if S is true according to what the extension of T tells us, the computation will lead to the result of its being false and if the extension of T says that S is false, then the computation will lead to the answer 'true'). Truly paradoxical sentences will go back and forth between true and false in this way, and in the limit will be eliminated as being of inherently unstable truth value. The result is once again a positive and a negative extension for T, with the stably true sentences making up the positive extension and the stably false ones the negative extension.

In the light of what was said above about the degree to which predicates like 'true' and 'false' are entrenched in the languages we speak, the approaches of Kripke and HGB strike me as plausible cures against the semantic paradoxes. Adopting the position that 'true' and certain other 'semantic' sentence predicates are partial doesn't seem too much of a concession when we consider that partiality of predicates is something that, as we have seen, needs to be acknowledged for other reasons as well. And there is an additional consideration in favor of such a partial extension approach, to which the next set on will be devoted.

6.5.1 Semantic Paradoxes of Pragmatic Origin

One way in which a language can fall victim to semantic paradox is though codification of its own syntax. But as Kripke observes in (Kripke 1975) natural languages have other ways of referring to their own sentences than via arithmetization or some other means of structure-preserving codification. In a discussion that is as amusing as it is to the point Kripke gives examples from (or inspired by, but little matters) the Watergate hearings in the summer of 1974. Here is one example in this spirit:

(27) Haldeman: Everything Nixon says is true.

Nixon: What Haldeman is saying is false.

On the assumption that the only things by Nixon and Haldeman are the sentences on display, the Nixon-Haldeman team has problem justifying themselves. If we suppose that what Haldeman says is true, then what Nixon says

141

is true. So everything that Haldeman says is false, including the statement of his on display. So what Haldeman says cannot be true.

So, since the supposition that Haldeman's statement is true leads to the conclusion that it must be false, we can conclude – categorically – that what Haldeman is saying is false. But if that is so, then something that Nixon is saying is false; and since by assumption the displayed statement by Nixon is the only thing that Nixon is saying, th false statement by Nixon can only be this one. But if this statement is false, then what Haldeman is saying must be true.

In this argument we have been referring to the statements by Haldeman and Nixon by description. If the sentences had been written on a piece of paper or a blackboard, we could also have referred to them by using a demonstrative phrase, like 'this statement' or 'that statement', while pointing at the display of the statements we want to refer to on the page or board as we are going through the argument. And there are also more specifically graphic conventions for referring to them, by using labels as in (28).

(28) (1): (2) is true.

      (2): (1) is false.

When the two utterances are displayed in this form, there is not even a need to assume that no other statements are in play. The point is that the labels '(1)' and '(2)' each play a kind of double role, as identifiers of the statements displayed, for use by us who are looking at the pair and want to draw our conclusions about them, and also as referring expressions that are used in the statements themselves to refer to the other statement; in this way the statements can talk about each other.[87]

The definite descriptions used in (27), the demonstratives mentioned in the paragraphs immediately preceding (28) and the labels (a kind of ad hoc proper names) used in (28) are all ways of referring to particular sentences

---

[87]There is, by obvious analogy, an even more condensed way of presenting paradox, as in (29).

(29)   (1): (1) is false.

Here the same labeling convention as in (28) is used in a succinct version of the sentence 'This very sentence is false', which also is true iff it is false. One is easily tempted to think that there is something fishy about (29). But when we see it as on a par with (28), the feeling that one is being hoodwinked largely disappears.

or statements that become possible when instances of the sentences or statements referred to are made available for reference in the ways that that is needed for these different types of referring expressions. But that is part of language too, language is there for being used, and using language is to produce tokens of expressions that can be heard or looked at.

It is because of this that liar-like paradoxes can come about so easily and be produced by people who have never heard of Gödelization or even syntax. And it is one reason why the problem is so obviously in need of being addressed. Both Kripke's own method and the HGB method strike me as good ways of dealing with it.

An interesting feature of Watergate-type paradoxes is that when combinations of statements by different speakers are jointly paradoxical, this is as a rule not only a property of the collection of utterances and not of any of the utterances individually, but its paradoxality also depends on what else has been said, or, more often, on what has not been been said. (Recall the assumption made in our example above that the two statements considered were the only ones made (at the given time) by Haldeman and Nixon, respectively.) This means that when the extensions of the truth predicate T in a model M are determined – from below, à la Kripke or from above à la HGB – the result will depend on what has been said and has not been said in a kind of global way. But in order to be say a little more about this, we must be more specific about utterances and sentence forms.

Part of the model-theoretic strategy of both the Kripke and the HGB approach is that the sentences of the given language L are elements of the Universes of the models for which the extensions of T. We now modify this assumption as follows. The Universes of our models now contain *statements*: particular utterances involving a sentence form (the sentence uttered) and an identifier (or 'label') that makes it possible for other statements to refer to the statement that is labeled by it. As we saw above, the identifiers that people use to refer to statements can take various forms. To simplify things let us assume that each statement comes with an index serving as its label. Specifically, a particular statement using the sentence S will be given as <i,S>, where S is the sentence uttered and i the label of the particular statement in which S us used. We assume that no two statements have the same label, though some statements may involve the same sentence.
Furthermore, given this way of identifying statements it is possible to form sentences like those in (30.a,b,c) and statements like those in (30.d). (Of course, the sentences that can be formed can be much more complex than

these, see below.)

(30) a. T(1)

    b. ¬T(2)

    c. T(2) & T(3) → ¬T(1)

    d. <1,¬T(2)>

As an example of how the Universe of the model determines the ultimate extensions of T consider once more the pair of statements in (28), now in its new guise (31).

(31) a. <1,T(2)>

    b. <2,¬T(1)>

When both (31.a) and (31.b) belong to the Universe of M, then neither of them will end up in either the positive or negative extension of T in M if we proceed à la Kripke and both will be filtered out of the positive and the negative extension of M when we proceed in the manner of HGB,and irrespective of whether either or both are placed within the extensions of T at the outset.

For a contrast, consider the two pairs of statements in (32.a.b) and (32.c.d).[88]

(32) a. <1,T(2)>

---

[88]You may think of these two pairs as exhibiting two forms of political discussion. The friendly debate in (32.a.b) consists of two utterances that consist in nothing more than that the speakers are in agreement. What they are in agreement about is just that they are in agreement. (32.c.d) epitomizes the hostile debate in which the participants charge each other with being wrong. Here too there is a kind of agreement, to the effect the other is wrong, and that is all that is being asserted in toto.
(32.c,d) can be seen as a more abstract rendering of a dialogue that was prominent in early discussions of Situation Semantics (see Section 6.4), in which the two dialogue participants both say: "I am right. You are wrong.", with th first sentence left out, without serious loss of content. (If you want, you may think of the dog of the Barwises and the dog of the Perrys barking at each other.) But on a more serious note: One important concern of Situation Semantics, about which nothing has been said in Section 6.4, is to give a situation-theoretic analysis of utterance pairs like this one. The conclusion Situation Semantics drew from such examples is that situations need not always be *well-founded*. Non-well-founded situations cannot be modeled as sets of infons, but require a much more sophisticated ontology,based for instance on non-well-founded set theory. This is one of the interesting dimensions of Situation Semantics and Situation Theory about which I said in passing in Section 6.4 that they have not been included in these reflections.

b.   <2,T(1)>

c.   <3,¬T(4)>

d.   <4,¬T(3)>

The pairs in (32) reveal an interesting difference between the Kripke and the HGB strategy. If we apply the Kripke strategy, starting from empty positive and negative extensions for T, then none of the statements in (32) will end up in either positive or negative extension of T. What we end up with on the HGB strategy depends on the starting extension of T. If both (32.a) and (32.b) belong to the starting extension of T, then they remain in the extension of T after each revision, and so end up in the positive extension of T. Analogously, when neither (32.a) nor (32.b) belongs to the starting extension of T, then they both remain outside on each revision and so both end upon the negative extension. And when one of the two is in the starting extension and the other is not, then they will alternate on each revision. (For instance, when (32.a) belongs to the starting extension, but (32.b) does not, then after one revision (32.b) will be in the extension of T and (32.a) out of it.) So in this case both will end up in the truth value gap of T.

The case of (32.c,d) is different. If both (32.c) and (32.d) are in the starting extension, then after one revision they are both out of it, after the next revision they are both back in and so on. So both end up in the truth value gap. Same if the two statements both start outside the extension of T. And when at the outset one of them belongs to the extension of T and the other does not, then this will remain so after every revision. So the first statement ends up in the positive extensional of T and the other in the negative extension. (Rightly so, for when the discourse participants both accuse the other of being wrong, only one of them can be right, even if nothing of substance is being said.)

It might be argued that in contexts like these the HGB has a certain advantage in that it enables us to formalize a certain attitude towards the evaluation of certain statement combinations, an attitude of taking people at their face value to start with: We start from the assumption that people speak truthfully. That is, we start from the assumption that what they say is true, i.e. that their statements belong to the positive extension of T. (That of course is somewhat extreme position. It may be that some statements are so obviously and irredeemably false that we should put them into the negative extension of T and that is then where they will remain on any revision.) When an HGB starting model reflects this 'presumption of truthfulness' as-

sumption, then initially true statements will remain in the extension of T so long as nothing forces them out of it. So the pair (32.a,b) will end up in the positive extension of T (as we saw above); and that seems fine; the two may not have said anything of substance, but their agreement that what they have been doing is agree with each other can be acknowledged as true, little as that may advance the debate. And that application of the truthfulness presumption to the speakers of (32.c) and (32.d) will have the effect that both statements end up in the truth value gap may seem to serve them right as well.

We can approximate the Watergate-type scenarios more closely by allowing for certain ways of classifying statements. The most directly relevant classifications are in terms of authorship. We can introduce such classifications formally by adding to our language statement predicates like 'statement made by Haldeman', 'statement made by Nixon' and so on. Let us add just these two: 'H' for 'statement made by Haldeman' and 'N' for 'statement made by Nixon'. Then we can also form statements like those in (33).

(33) a. $<5,(\exists x)(\ N(x)\ \&\ (\forall y)(N(y) \rightarrow y = x)\ \&\ T(x))>$[89]

b. $<6\ ,(\forall x)(H(x) \rightarrow \neg\ T(x))>$

Let M be a model whose Universe contains both of the statements in (33) and in which the (bivalent) extension of H = $\{<6\ ,(\forall x)(H(x) \rightarrow \neg\ T(x))>\}$ and the extension of N = $\{<5,(\exists x)(\ N(x)\ \&\ (\forall y)(N(y) \rightarrow y = x)\ \&\ T(x))>\}$. Then we can deal with the extensions of the T predicate in the same ways as above. But note that because of the quantifiers in (**??**) the entire Universe of M is now involved in the computations on which the inductive steps of the Kripke method and the revisions of the HGB method are based. (Computing the details of those revisions can be tricky as well as laborious, as those who have played around with these methods will have experienced.)

We could further refine our analysis of examples like those in amplify our analysis of examples like those in (31) - (33) to account for yet other ways in which people can speak about their own statements and those made by others. but I think this suffices to bring out my primary concern here: The role that played in paradoxical statement combinations by pragmatics, in the general sense of language use. Those are instances of paradox, you might say, for which the language itself is not to blame, but only those who make uses of

---

[89]I adopt this as the formalizations of 'What Nixon says is true', using a Russellian rendering of the definite description *what Nixon says*.

it that they should not have, if paradoxality is a sin that ought to be avoided. In this regard the paradoxes we have been talking about in this section are different from those that are found in languages with truth predicates on account of their expressive power, which enables them to encode their own syntax. From a methodological point of view this distinction is crucial. When formal semanticists provide their model-theoretic analyses of certain fragments of natural languages, there is generally an implicit assumption that those analyses will permit up-scaling to analyses for larger fragments and eventually to semantics for the language as a whole. But if semantic predicates like *true* are part of the language as a whole, there will come a point where this goes awry unless appropriate measures are taken. To repeat, the methods of Kripke and HGB mentioned above strike me as good candidates for restoring consistency to accounts that at that point would become self-contradictory. There may be other ways to do this too, perhaps proposals have been made to this effect that I haven't seen and that are preferable. But it is important that we have at least one general method for this that we can trust. With such a method in the background, we can carry on doing model-theoretic semantics the way we have been, knowing that when push comes to shove, the method can be applied to put us back on the rails.

6.5.2   Conclusions to Section 6.5

Let me repeat the main concerns of this section. First, the threat caused by semantic predicates like truth should be taken very seriously, and especially by those engaged in the semantics of natural language. Second the emphasis has been on three main points:

(a) the intrinsic paradoxes to which languages are open with sufficient expressive power, a point first made by Tarski in relation to truth, and extended by Montague to a slate of other notions whose presumed axiomatic properties are not as strong as that represented by the Tarski 'T-equivalence', but strong enough to derive a contradiction in languages with this power.

(b) Besides this source of semantic paradoxes, such paradoxes can also arise for what might be called pragmatic reasons, when speakers can refer to each others' statements and in that way produce the effect of incoherent semantic dependencies.

(c) It is important to distinguish between these two sources of paradoxality because they have very different impacts. The problem posed by paradoxical

speech situations may be a problem for the language users involved in them. But the language cannot be blamed for them. (In fact, such paradoxcal situations can arise even for very simple languages, for instance a language of first order Predicate Logic with as its only vocabulary the truth predicate T, a bunch of individual constants as names for certain sentences of the language.[90] The expressive power of such languages can be very limited, far too limited for self-arithmetization. This is surely true of the language of Predicate Logic just described. Paradoxical sentence or statement combinations of this sort need not be a worry for the semanticist; they are just an intriguing topic of investigation. But the intrinsic paradoxality of Tarski and Montague should be a worry for all those whose preoccupation is with the expressive languages which have the means of self-arithmetization. And that includes, whether we like it or not, pretty much all the languages spoken by humans around the globe.

(d) Let me throw in a question here that was not mentioned so far. Throughout the discussions of Section 6.5 I have assumed that 'intrinsic' and 'pragmatic' semantic paradoxes are separate phenomena, in that particular paradoxes arise either for the one or for the other reason. But this is something that I am actually not sure of. First, there may, for all I know still other ways in which semantic paradoxes arise, and second, even when that ids not so, are there interesting problems that can arise through interactions between th two sources we have considered.

I concluded Section 6.5.1 with the urgent recommendation that we avail ourselves of a general method that can be used to render our semantic analyses immune against the semantic paradoxes. And to repeat, the methods of Kripke and HGB seem to be good for this, but I feel under informed at this point about all that is known about this general issue; perhaps there are objections against either of these two methods and perhaps there are better ways of dealing with the issue than I know.

As a final point, and one that reaches beyond Section 6.5, both Kripke and HGB treat truth as a partial predicate. Quite a bit has been established about the logical properties of these two methods. But a wider question, which to my knowledge has not so far been systematically investigated, is what happens when these sources of partiality combine and interact with

---

[90]We have seen in Section 6.5.1 that in models for such a language in which these constants do denote the sentences they are meant to denote, the truth predicate cannot have a classical extension if there are sentences (with occurrences of those constants) that form paradoxical loops.

other sources. One of those is the notion of truth or falsity in virtue of a limited amount of information, as studied in Situation Semantics. Another one, referred to only obliquely in this document, as vagueness. There may be others. Each may have its own impact on logic, even if we focus just on relations of logical consequence. But what can we say about languages in which all these forms of partiality manifest themselves? Does anybody have an idea, even if it is just a rough picture?

6.6 Logic's large and varied Battlefield: Where the Discrete meets the Continuous

I have kept the topics of this and remaining sections to the last because I am aware with particular acuity that my knowledge of the matters I want to touch upon doesn't really justify my saying anything much about them at all. That I include them at all, and in spite of my deficient knowledge, is because I see them as the main challenges to the more traditional approaches to formal semantics with which I am less unfamiliar and that have been the almost exclusive concern of these reflections up to this point. The challenges emerge on a number of different sides and I believe each has to be met in its own way. But there is one feature that they all have in common and that sets all them apart from everything we have discussed so far.

This common feature is the introduction of *continuous* as opposed to *discrete* concepts and methods. In some cases these methods are meant to deal with problems that were never within the horizon of the practitioners of the old discrete methods. But the new methods are also often applied to the very same problems that were previously tackled with the old methods. Here too the new methods often achieve remarkable results, especially when speed and broad coverage are of the essence to the results one wants. With some this has led to the opinion that the new methods should replace the old ones across the board. But not everyone shares this view. There are also many tasks, this other side believes, where the best way to proceed will prove to be a combination of the new methods with the old ones, and that one of the major challenges for the near future is to find ways in which these very different methods can be made to work efficiently together.

There are various ways in which the opposition between the continuous and the discrete can be conceived and several are important to the few things I want to say in this section. In general the distinction has to do with continuous and discrete domains. As I intend this difference here, a continuous

149

domain is one which comes with a notion of *distance* such that for any non-zero distance ϵ and any point x in the domain, there are points y distinct from x such that the distance between x and y is non-zero but less than ϵ. Admittedly, this is not a very good definition insofar as it begs the question what the non-zero distances ϵ are supposed to be. Rather than making too much of a fuss about this[91], let me take an easy way out, which will do well enough for the concerns of this section (and in fact, for all the practical contexts in which these concerns will become tangible). As simple way out I will assume the structure $\mathcal{R}$ as basic continuous domain, and define other domains as continuous in the sense described above assuming that ϵ can be any positive number (no matter how small) from $\mathcal{R}$.[92]

---

[91]Here is one way the fuss could be made (but by all means, skip this footnote): At a minimum we can introduce distance as a comparative notion, a 4-place relation CD which holds between x, y, z and u (formally: $CD(x,y,z,u)$) iff x is least as close to y as z it is to u. CD is to have the following properties (i) for all x,y,z,u, if $CD(x,y,z,u)$, then $CD(y,x,z,u)$ and $CD(x,y,u,z)$; (ii) for all x,y $CD(x,y,x,y)$; (iii) for all x,y,z,u,v,w, if $CD(x,y,z,u)$ and $CD(z,u,v,w)$, then $CD(x,y,v,w)$; (iv) for all x,y,z $CD(x,x,y,z)$; (v) for all x,y,z,u, $CD(x,y,z,u)$ or $CD(z,u,x,y)$. Of these, (i) entails that CD can be understood as a binary relation between unordered pairs of individuals. (ii) says that this relation is reflexive and (iii) that it is transitive. (iv) says that pairs of the $\{x,x\}$ are minimal according to the relation CD: the distance between x and x is less than or equal to that between any y and z whatever. (v) says that the comparison expressed by CD is exhaustive. Because of (ii) and (iii) the relation $\equiv$ between pairs (x,y) and (z,u) defined by '$\equiv(x,y,z,u)$ iff $(CD(x,y,z,u)$ & $CD(z,u,x,y))$ is an equivalence relation and thus that we can form the corresponding equivalence classes $[(x,y)]$. These equivalence relations are also ordered by the relation between (x,y) and (z,u) expressed by $CD(x,y,z,u)$. (iv) says that what ever x is taken, $[(x,x)]$ is the smallest element of this ordering (since $CD(x,x,y,y)$ for all x and y). Furthermore the ordering between the equivalence classes is total in virtue of (v).
In a model M in which CD satisfied these conditions, a -2-place function d can be defined by: (vi) for any a, b in U_M, $[d]_M((a,b)) = [(a,b)]_{\equiv}$. That is we can also add the 2-place function constant d to our language with the stipulation that in every model M its extension is defined as in (vi).
To make sure that the ordering between unordered pairs is dense, and thus that between any two distances $[(x,y)]$ and $[(z,u)]$ there is a third, we need a further postulate, to be formulated with the help of the notion that the distance between x and y is strictly smaller than that between z and u. This relation can be defined by the conjunction '$(CD(x,y,z,u)$ & $\neg\, CD(z,u,x,y))$'. Abbreviating it as '$\prec(x,y,z,u)$' the additional postulate we need is: (vi) when for any x,y,z,u $\prec(x,y,z,u)$, then there are v, w such that $\prec(x,y,v,w)$ and $\prec(v,w,z,u)$. (vi) renders the distance structure densely ordered. For the difference between density and continuity see the next footnote. Continuity, in that sense in which it is used in mathematics and in which it is different from density, can be defined for the structures involving CD and d as well.
[92]Strictly speaking this does not fit the mathematical definition of 'continuity' as opposed to that of 'density'. On the definition just given the structure of the rationals, with absolute difference as distance function, qualifies as a continuous domain, even though according to standard terminology this structure is densely but not continuously ordered

Logic has traditionally been concerned with structures that are discrete through and through. They are discrete in the first place in that syntactic structures are discrete. (This is true both for the languages of formal logic and for natural languages according to the concepts of generative syntax.) Although all standard examples of logical languages have infinitely many well-formed formulas, with syntactic structures that establish their well-formedness, and this is also assumed by most current syntactic theories for all or most natural languages, the syntactic structures assumed for either kind of language are discrete in that each well-formed expression has immediate constituents, with nothing between those and the syntactic complex immediately composed of them. Furthermore, on the semantic side well-formed sentences are assumed to have one of two possible truth values, True and False, or 1 and 0; or, formulated in terms of Model Theory: every sentence has in any given model (or in any given model at every given index) on these two truth values.Traditionally, logic has been about preservation truth and in more modern times it has been about the truth values that sentences get in different models, and also about various other issues, many of which have to do with the truth values that sentence take in different models. But (cutting things a little short), all these problems have to do with values from a discretely structured set taken by the discretely structured formulas from some artificial or natural language. As noted in the introductory part of Section 6, logic is no longer just, or even primarily, about the question how valid inferences can be established via formal deductions.

Even in this heartland of traditional logic continuity made its first inroads as much as half a century ago or more. It was recognized close to a century ago that bivalent could be generalized to *many-valued logo* in which the two truth values True and False are replaced by some set of 3 or more truth values, with many-valued truth tables for the connectives and corresponding adjustments for the truth conditions of quantified formulas. That still remains within the realm of the discrete. But thinking about more than two truth values prepared the found for a further generalization, towards a truth value space of probabilities. In *probability logic* formulas get probabilities – real numbers with the interval [0,1] – as 'truth values', with the values 1 and 0 corresponding to the old values '(definitely) true' and '(definitely) false'. This introduction of continuity, in the form of the real number interval [0,1]

(since it isn't closed under limits of converging sequences, unlike the real themselves, which are closed under such limits). The distinction isn't irrelevant to all matters raised in the section, but for most it can be kept in the background.

as truth value space, gives rise to a range of new questions about logical validity, two of which I will mention here.

Two related results about propositional logic are associated with the work of Adams.[93] For both the topic are valid arguments of classical propositional logic. Suppose that $\Gamma \vDash \phi$ is such an argument from a given language L of propositional logic. Let a probability model for L be a probability function P from the sentences of L to probability values. That is, P is a function from the sentences of L into the real interval [0,1]) satisfying the principles:

(i) for any ψ, $P(\neg\psi) = 1 - P(\psi)$, and

(ii) for any ψ, χ, if $\vDash \neg(\psi\ \&\ \chi)$, then $P(\psi \lor \chi) = P(\psi) + P(\chi)$.

Then it is possible to say something about how the probability of ϕ depends on the probabilities of the premises γ in Γ in view of the fact that $\vDash \phi$ is valid. One would hope that something can be said about the minimal values that a probability model could assign to ϕ given the values it assigns to the premises in Γ, and indeed that turns out to be possible. In fact, a rather crude, but nevertheless useful lower bound can be stated just on the strength of the premise set Γ alone. (That is, this lower bound will hold for any ϕ that follows from Γ in classical propositional logic, on the assumption that Γ is finite.) The bound is given in (34).

(34) $P(\phi)\ \geq\ \sum_{\gamma\in\Gamma}P(\gamma) - (|\Gamma| - 1)$

In general this isn't a terribly good lower bound.[94] For one thing it doesn't tell us anything of interest unless the probabilities of the premises are fairly high, and this becomes ever more important as the number of premises $|\Gamma|$ goes up. The expression to the right of $\geq$ in (34) quite easily is ≤ 0, in which it imposes no constraint on $P(\phi)$ at all. Even when $|\Gamma| = 2$ this happens already when each of the two premises gets probability $\leq 1/2$, when $|\Gamma| = 3$, when each premise gets a value $\leq 2/3$ and so on. But with high premise probabilities the bound can be informative. In particular, it is easy to derive from it the following theorem:

---

[93]See (Adams & Levine 1975), (Adams 1998). In trying to write this section I have much profited from the Stanford Encyclopedia of Philosophy article on Logic and Probability (Demey, Kooi & Sack 2019).

[94]Better bounds can be obtained when the conclusion ϕ is taken account as well (and not just the premise set Γ) in the definition of their definition. See for instance (Demey et al. 2019).

(35)   Theorem (Adams)

Suppose that the argument $\Gamma \vDash \phi$ of the propositional language L is valid. Let ϵ be some real number in [0,1]. Then there is a number δ such that for any probability model P for L if $P(\gamma) \geq 1 - \delta$ for all $\gamma \in \Gamma$, then $P(\phi) \geq 1 - \epsilon$.

In other words, when an argument of L is classically valid, it is also valid in the sense stated in the Theorem: if the probability values approximate 1, so does the probability of the conclusion. The converse clearly holds as well.[95] So, the statement of the Theorem is another necessary and sufficient condition for classical propositional validity. [96]

My reason for drawing this much attention to this connection between logic and probability has to do with a strong hunch about the distinct roles that are played in human cognition by probability on the one hand and by logical reasoning as an activity aimed at preservation of truth on the other. In both cognition and epistemology there has been considerable emphasis over the past decades on our beliefs admitting of different degrees and that the degrees of confidence with which we hold the different things about which we have an opinion is crucial to the roles our different beliefs play especially in our practical reasoning – in our planning, decision making and the actions that ensue from those activities. The stronger/weaker our convictions, the more/less impact they have on the practical conclusions at which we arrive. To do justice to this aspect of our doxastic states a theory of practical reasoning must be able to capture the confidence degrees of our beliefs as well as their content, e.g. by representing each belief as a pair of such a degree and a propositional content. The technical term that has been adopted for

---

[95]Suppose we take ϵ to be 0, then whatever a corresponding number δ may be that exists according to the Theorem, a probability model P in which all the premises are given the value 1 will satisfy the precondition. So any P that assigns 1 to all the premises γ will also assign 1 to the conclusion ϕ.

[96]To my knowledge there are no good analogous to these results for predicate logic. (The problems start with the definition of probability models for predicate logic, which impose intuitively correct constraints on the probabilities of quantified formulas.) of course the mentioned results have a certain relevance for predicate logic too, but only in the fairly trivial sense that when $\Gamma \vDash \phi$ is an argument of predicate logic whose validity can be established on the basis of propositional calculus principles only, then probability assignments satisfying the mentions constraints will satisfy the same bounds (such as, in particular, that of (35)) as in the propositional case (such as, in particular, the one of (35)). But this does not extend in any way that I can see to arguments whose validity rests essentially on principles of quantification theory. (**I suspect that there exist results of the kind I am looking for here. If anyone knows, please let me know too.**)

such 'graded beliefs' in the literature is *credence*.

Reasoning with credences is of necessity a more complex matter than reasoning with beliefs as traditionally conceived. As traditionally conceived, all beliefs come, you might say, with the same 'absolute' degree of confidence (which need not be and isn't explicitly articulated) and the principles of valid reasoning are defined purely interns of traditional discrete syntactic structure. In credence reasoning probabilities play an essential roles but I do not really understand very well what form or forms this role can take. (Only in a few special cases, do I think I have an idea of how things work.) But the hunch I spoke of above entails that the role of probability in probabilistic reasoning is only one part of the problem. Centrally important as probabilistic reasoning may be in human cognition, reasoning of the kind traditionally studied with the tools of deductive logic also plays its part. What theories of deductive logic have to say about reasoning isn't just an early, simplifying stage of the theory of human reasoning, of which we have at last come to realize that it should be replaced by a theory that takes degree of belief into account as well. Human reasoning is modular in that it often combines different modes of reasoning. Some of our reasoning really *is* in the degrees-free mode, but it alternates with reasoning in which degrees play an essential part, and often the two modes work hand in hand, for instance when we are trying to make up our minds about what to do when we have only uncertain information about the situation in which we will have to act.

In fact, it seems that degree-free reasoning can be a useful part of a larger, more complex reasoning process in which probabilities do play a role for two distinct reasons. First, it may be helpful because of the lower bounds with which it may provide us for the probability of the conclusion of a deductively valid argument $\Gamma \vDash \phi$ given certain lower bounds to the probabilities we feel en4titled to assign to the premises in Γ, in the sense of Adams's work and (34) above. And second, there are many situations in which we reason from premises that we take for granted, and where the question of their degree simply doesn't arise for us.[97]

---

[97]In (Kamp 2020) it is proposed that we need to distinguish between those contents that an agent treats as beliefs – contents that can be challenged and are up for debate – and contents that the agent takes for granted, as unquestionably true, and which together make up the agent's *doxastic horizon*. (Kamp 2020) implements this idea within the SMDRT framework (see Section 5). Here it may be added that the beliefs of (Kamp 2020) can be represented as credences. The information that is part of the doxastic horizon shouldn't. It has degree 1 by default, simply in virtue of being where it is.

If this is right, then the challenge for logical theory posed by degree-related reasoning is not just a matter of understanding the role or roles that probabilities may play in probabilistic reasoning. The challenge is also to understand how modes of degree-based reasoning alternate and dovetail with the kind of degree-free reasoning that has been the almost exclusive focus of logical theory over more than two millennia.

6.6.1   Changing one's Mind

Changing one's mind has not so far been part of our story. Of course, we have been speaking, implicitly at least, of how minds change. That happens with any cognitive process that a mind engages in, including any form of reasoning. For instance, if I logically deduce a proposition p from things that I have believed for a long time, and that I have never found any reason to doubt or question, and the conclusion p is nevertheless that I did not realize until now, even though it does logically follow from those things that I have been taking for granted, then this will be a genuine change of mind, with perhaps significant consequences for my further behavior, even if there is a sense in which my knowledge hasn't increased, since it is entailed by things I already held to unquestioningly true.

But such changes are not what we normally understand by changing one's mind. Someone changes his mind when he gives up something that he believed to be the case so far and replaces it with some other, incompatible belief.[98] When we think of beliefs as credences changing one's mind can be thought of as involving a credence and changing its confidence degree while retaining its content. And surely many changes take some such form as this. Something we were firmly convinced of until now we come to see as not at all obvious; we keep the content, but lower our confidence degree from 1 or close to 1 to 1/2 or to some value close to that. Or, lees dramatically, we go from some moderate degree of negative belief, of around 1/3, say, to a moderately positive degree, around 2/3.

There are also maximally dramatic changes along this dimension, in which the degree switches from 1 to 0 (or from 0 to 1, which comes to the same thing in a way, since changing you credence in p from 1 to 0 is the same thing as changing your credence in $\neg p$ from 0 to 1). Such maximal changes can also be understood as changes of degree-less beliefs, from p to $\neg p$; and changes

---

[98]Or when someone changes his judgment or opinion of something, form condemning it, say, to the view that it isn't really all that bad after all. But such cases will play no further part in our considerations here.

of this sort have been a subject of investigation for many years, stating well before the continuous began to make it impact on the formal study of logic and reasoning. The subject has been known since those early days as *Belief Revision*. Classical Belief Revision theory[99] studies policies for revising formal theories – theories on some (often scientific) topic or other – when new information comes to light that conflicts with the theory as one has it. It is assumed, moreover, that this new information overrides the evidence supporting the theory up to that point, and that revision of the theory is required to render it compatible with the new information, which can then be added to it after the revision has been made. The main concern of this work is to develop rational policies for going about such revisions, based on the logical structure of the theory that needs revising.

Belief Revision Theory can be seen as a chapter of traditional logic, in which object and methods are thoroughly discrete in the sense explained at the beginning of Section 6.6. Credences, with degrees other than 0 and 1, have no place in it. But Belief Revision too has its continuous counterpart (or 'revision', as some people seem to see it). In fact, this alternative has also been around for a long time, and for much longer, though perhaps not so much in connection within the formal logic community. What I have in mind is what is usually referred to as 'Bayesianism', a somewhat loose term, which includes more or less everything that builds on Bayes' 'Theorem' according to which conditional probabilities of two propositions A and B are related according to the rule:

(36) $P(A|B)\cdot P(B) = P(A \,\&\, B) = P(B|A)\cdot P(A)$

The basic principle underlying this is that the probability of a conjunction A & B is the probability of one of the conjuncts multiplied by the conditional probability of the other conjunct on it; that is, you might say, what conditional probability *means*, or an immediate consequence of what it means. It is the commutativity of conjunction that allows us to apply this conceptual identity in either direction, with B as the 'independent variable and A the 'dependent variable' or the other way round.

What makes (36) so important is its use in probability adjustment. This is the foundation of all of *Bayesian learning*, learning from new information by adjusting your credences. The basic intuition here – again this is part of the very concept of conditional probability – is that when new information, i.e.

---

[99]See for instance (Alchourron, Gärdenfors & Makinson 1985), (Gärdenfors 1988), (Hansson 1999).

a new proposition E – becomes available, in the sense that there is no doubt about its truth, then the probability of H should be revised to what had been its probability conditional on E so far.

But what does this revision amount to in practice? If the prior conditional probability $P(H|E)$ of some 'hypothesis' H on some piece of evidence E is known (where H and E are both understood as propositions), then we can simply take this conditional probability as the new unconditional probability of H. But very often this conditional probability isn't known. However, in some such situations there will be knowledge about the converse probability, $P(E|H)$. And in such situations Bayes' 'Theorem' (36) will often enable us to say something concrete about how to revise our probability for H. The simplest and most intuitive examples of this that I am familiar with are cases where we have a number of competing hypotheses $H_1,..., H_n$ each of which bears on the evidence E, and where this bearing of H_i on E is given in how probable E is given H_i. That is, for each H_i we have the conditional probability of E given H_i, $P(E|H_i)$. Suppose further (i) that the hypotheses $H_1,..., H_n$ are mutually exclusive, i.e. that for $i \neq j$, $P(H_i \ \& \ H_j) = 0$, and (ii) that $H_1,..., H_n$ cover all the possible explanations for E, i.e. $P(E|(H_1 \vee ... \vee H_n)) = P(E|(H_1) + ... + P(E|(H_n) = 1$. Then Bayesian revision is possible upon finding that E is true, using Bayes' 'Theorem' (36).

For simplicity let us assume that $n = 2$, i.e. that there are just two such hypotheses H_1 and H_2. Then on the basis of (36) we can derive:

(37) $P(H_1|E) = P(E|H_1) \cdot \dfrac{P_{(H_1)}}{P_{(E)}} = P(E|H_1) \cdot \dfrac{P_{(H_1)}}{P_{(E|H_1)} \cdot P_{(H_1)} + P_{(E|H_2)} \cdot P_{(H_2)}}$

To get a sense of what this formula says, note one immediate implication of (36): If and only E is stronger evidence for H_1 than it is for H_2 – if it strengthens the case for H_1 more than it strengthens the case for H_2, formally: $P(E|H_1) > P(E|H_2)$ – then the revised probability of H will be higher than its 'prior probability' (the probability that it had prior to revision).

These are among the most elementary examples of Bayesian revision, but they display the core of its many far more sophisticated applications in, mostly, the natural sciences. Let me in passing mention one other kind of application, which involves continuity in the more familiar sense of the calculus, that of functions from the reals into the reals or from the complex numbers into the complex numbers, functions that are continuous in the familiar sense of the calculus that small distances between their arguments guarantee small distances between their values. (Many of the functions used in applications

of probability theory and statistics are continuous, and many have additional properties, which make it possible to make use in those applications of the sophisticated mathematical toolkits that exploit those properties.)

The kind of application I have in mind is one where there are good grounds for assuming that there is some particular magnitude – say, the average amount of time within a large given population that it takes from infection with a certain virus to the first manifestations of the disease it causes –, but that it isn't known exactly what that period of time is, and wants to find this out by observing individual cases. All that one feels confident to assume before the inquiry starts is that the value one is after (the average incubation time of the disease) has some natural distribution, given by some particular Gaussian distribution, say. For each possible value that the true magnitude x can take within the range of this distribution, the value of particular observations of incubation times y may be expected to be distributed around the given possible value x (and here too the distribution can often be assumed to be a Gaussian). Suppose now that a certain incubation time y is observed. (I am assuming that in certain cases this period can be determined with sufficient accuracy for the enterprise to make sense.) Then this value y will have for each possible value x for the true incubation time a certain probability conditional on the hypothesis that x is the true average incubation time; but these conditional probabilities will vary as a function of what x is. (It will vary as a function of the distance between y and x.) The situation here is like the one considered above, with the n alternative Hypotheses $H_1,...,H_n$ to account for E, except that we now have an infinite set of hypotheses, each corresponding to one of the possible values for x. So essentially the same conditionalization process as that described above can applied in this case too in order to determine a new probability distribution for x, in the light of the evidence y that we found.[100]

Applications like those outlined above have become so common in the sciences

---

[100]The computation of the new distribution is a little more complicated than my description of it here makes clear, since we have to work with continuous probability distributions, rather than with finite sets of n distinct conditional probabilities. It isn't possible to assign individual non-zero probabilities to individual real numbers x and y, since there are far too many of them. Non-zero probabilities are possibilities only for real number intervals surrounding individual numbers. That is, we can only talk about probability *densities* here, where the *density* of the distribution for the argument x is the is an (approximate) measure for the probability that the true value is within some small interval around x, e.g. the interval $[x - \epsilon, x + \epsilon]$ for some small fixed number ϵ . 0. It is the probability densities for the different possible values of x that get revised in the light of having found y, and it is this that leads to the new probability distribution for the average incubation time.

that it is hard to imagine what science today would be like without them. In those applications revision is possible because there are good ways (if sometimes very complex ones) of estimating the prior conditional probabilities which the method presupposes. (Usually these are conditional probabilities of the evidence on each of the two or more scientific hypotheses between which one wants to choose.) And it is because of these independent methods for determining those prior probabilities that the applications make sense and can be justified. Quite different is the appeal to the Bayesian method as a general model of how we update our doxastic states in the light of new information. There is a new trend within epistemology according Bayesian revision is the method by which we learn from experience, with each new experience potentially modifying our credences, whose degrees gradually better reflect the world because their increased power to explain the different facts that we encounter in the course of our lives.

Such a model of the dynamics of knowledge and belief is appealing because it seems fundamentally right, because in general outline it seems reasonably simple and because it seems widely and, one might have thought, universally, applicable. But as a general account of human learning it faces a fundamental objection. To put it somewhat formulaically, Bayesianism treats all new information as in some sense old.

This last sentence is meant to mean more than one thing. First, when a new bit of information reaches a cognitive agent, in the form of a proposition B that she recognizes as true, then the Bayesian method requires her to revise the credence degrees of all her credences with contents A from $P(A)$ to $P(A|B)$. And not only that, the credence degrees of all her conditional credences of contents A on contents C will have to be modified as well, from $P(A|C)$ to $P(A|(C \ \& \ B))$. For otherwise the agent might be unprepared for what she has to do when the next bit of information comes in. But how plausible is it that human agent should be equipped, throughout her life, with all conditional probabilities of the potential belief contents she can entertain? Having conditional probabilities for all these combinations is like having all the branches of your doxastic future mapped out before you from the start, and that seems implausible: Where are the degrees of all those conditional credences supposed to come from? Are we really equipped with this much wisdom from the cradle?[101]

---

[101]Perhaps a defense against this objection could go like this: Part of informational innocence is that one takes contents to be independent of each other. An agent may have many categorical credences – likelihood estimates about information contents, but far fewer conditional credences. When a new bit of information B becomes available to

But the problem isn't just one of omniscience about credence degrees. The Bayesian model also suggests that the agent must be equipped from the start with all contents that she may have to deal with at some later point in life. At the very least, the contents B that may become available to her as true evidence at some later time will have to be part of her cognitive structure beforehand, as conditions of the conditional credences that need to be turned into categorical credences when the time for conditionalizing on B has come. And from all that has been said so far, it would appear that they have to be part of her doxastic state from the moment go. For if not, then she could not perform conditionalization on B if and when she needs to.

And that assumption too is very unrealistic. Much of the information that comes our way at different times in our life is new to us in the sense that the propositional content is new to us *qua content*, and not just as something that we had been thinking about, but for which we had thus far insufficient evidence. It is true that some of the information we get is familiar content-wise; some of it may answer a question we have had, or contradict an opinion thus far held. But much of it is not. It is information that never occurred to us, and that quite possibly would never have occurred to us if we hadn't been confronted with it on the occasion when something drew our attention to it.[102]

the agent, she will then update those categorical and conditional credences for a content A for which she has conditional credences in which B is the condition, or a conjunct of the condition. But if she has only few such conditional credences, then the cognitive effort involved may be within reasonable bounds. In the limit, where the agent has no conditional credences in which B is or is part of the condition, no degree adjustment will be required. Note well, however, that now we need a further story about how conditional credences are acquired. I have no good idea of how such a story might be told.

[102]The way in which new information is sometimes new can be even more dramatic than this. Information may be new to an agent not only in the sense that it would never have occurred to her to formulate it (let alone ask herself whether it was true), but more radically, in that she would not have been *in any position* to formulate it, because she was missing some or all of the concepts in terms of which the information must be framed. Learning by receiving this kind of new information, and turning it into genuine information by acquiring the concepts that come as its constituents, and thus come within its wake, is perhaps the most important kind of learning of which human beings are capable. And how such information is acquired, and with it the novel concepts that are part of it, is perhaps the greatest mystery of our development into the cognitive agents we grow up to be, and one of the greatest challenges for cognitive science, and especially for cognitive linguistics as a branch of it is to understand better how concept acquisition works, from the earliest phases of our cognitive development onwards – I do not think that this ever ends; that's one of the reasons why life continues to be exciting.

A far as language is concerned, there are two aspects to the phenomena of concept for-

I conclude this subsection with an observation that is orthogonal, but that I want to make so in order to head off a misunderstanding to which at one point I fell victim myself. I earlier spoke of the Bayesian approach to learning and knowledge as revision of credences, and that could be taken as implying that Bayesian revision does with credences what Belief Revision, as discussed in the opening part of this section, does with degree-less beliefs. But that isn't right. Belief Revision comes into action only when the new information contradicts the beliefs so far held. When there is no contradiction, no revision is needed: the new information is simply added as a new belief to those one had already. And when Belief Revision does come into action, the processes described in the Belief Revision literature describe in great formal detail how this is to be done, so that as much of the most important parts of the cluster of old beliefs can be retained, while yet consistency with the new bit of information is restored.

Bayesian revision, we have seen, is quite different. It is to be performed when a new bit of information comes in, whether that information is consistent with what one had or not. In fact, it may not be immediately clear what we should understand in general by inconsistency of a new bit of information – let us assume that the agent treats it as beyond doubt by adopting it as

---

mation. One has to do with concept acquisition by individual agents: how do they learn concept words, in that they learn to align the concepts they associate with those words with how other members of the language community use those words, at least to the extent that speakers share their understanding of the truth conditions of sentences containing those concept words. The other factor has to do with the growth of the language. What goes on when a member of the community casts a new word for a new concept and that word and the concept it designates then become a common property of the language community as a whole, or of some significant part of it? One striking feature of such additions to the language as a whole is that there is a preference for using words that are part of the language already, but with a new sense, which distinguishes itself from the sense or senses already in place in that it denotes the new concept. Such novel 'metaphorical' uses of old words and the strong preference that speakers seem to have for them over the adoption of new words without any etymological connections to existing vocabulary, has been rightly recognized as a key to our understanding of concept formation of use, and therefore as equally important for our understanding of language and our understanding of cognition. Arriving at a better understanding of all this is, I believe, one of the most important challenges for the science of language and mind in the years ahead. In a set of reflections on the future as well as on present and past of logic and language this direction of research should have deserved greater prominence than in a mere footnote. But as this document developed, there never seemed a better place for going into this complex of issues in any detail – largely, I suppose, because I never managed to make any significant contributions to this topic myself (although that hasn't kept me from raising many other general issues that have been and are worrying, such as those in the present section).

a credence whose degree is 1 – and the set of categorical and conditional credences that according to the Bayesian picture constitute the agent's doxastic state when the new information is received. But there are some special cases where it does seem clear that we have inconsistency. One of these is where the concern is about a certain hypothesis H which entails some piece of evidence E, and where the agent is aware of this entailment. That means (I take it) that the agent assigns degree 1 to E conditional on H and, by the ams token, 0 to $\neg E$ conditional on H. Suppose now that at one point $\neg E$ becomes available to the agent and that as a proper Bayesian agent she revises her various credences according to Bayesian rule. That will then lead to the effect that she now assigns 0 to H (see (37)). That is where the Bayesian story I have been telling ends and that is where the Belief Revision story starts.

This conclusion is: Belief Revision and Bayesian revision aren't distinct methods for doing the same thing, but by different means and in application to different doxastic objects. They are complementary methods with one coming into action in situations when the other leaves us with an unpalatable result, without telling us what to do next. There is the further difference of course that Belief Revision belongs to the realm of the discrete, whereas Bayesian revision operates on objects that are continuous at least with regard to their degrees.

Since the two approaches are complementary rather than competitors, this difference in the structures they presuppose may seem puzzling. But perhaps this is one place for a hybrid architecture: An account of knowledge and belief and their acquisition in which the virtues of Bayesian revision and those of Belief Revision are combined should perhaps take the form of treating inconsistency as that special case where processing switches from credences to beliefs and the discrete modus operandi of Belief Revision comes into action.

Such a mixed architecture for a theory of knowledge and belief management will raise new questions. For instance, after Belief Revision has done its work and has come up with its revision, should the belief contents in this revision be turned back into credences? And if so, what should then be their degrees?

6.6.2 Bayesian Revision and Question Answering

The cases I sketched of the use of Bayesian Revision in science are all cases where the method is used to make progress with answering questions that have been consciously posed and for which answers are sought – e.g. the

question which of two possible hypotheses H_1 and H_2 is true, or what the true value is of a magnitude that is fixed by natural and/or social factors. I contrasted such cases, which are part of accepted and by now indispensable scientific practice, with Bayesianism as a general approach to the acquisition and modification of belief, which seems up against very serious objections. But we might also consider Bayesianism in a context that isn't limited to science but is nevertheless restricted to the problem of finding answers to questions posed. In this section I want to pursue some of the possibilities for this more restricted use of Bayesianism in accounts of knowledge acquisition and belief management.

To this end let me begin with a quick survey of questions and answers as they are standardly dealt with in formal semantics. As far as the semantics of questions is concerned, the standard assumption is that the semantic content of a question can be identified with the set of possible complete answers to it. Providing a complete answer to the question then comes to selecting the true answer from this set. But that can take all sorts of forms, and here pragmatics may come into it in all sorts of ways (Groenendijk & Stokhof 1984).

This general schema for analyzing the question-answer relation can be filled out in various different ways. One of the factors is how possible answers, and therewith the questions that they are the possible answers to, are to be identified, given or represented. That is important, for the way in which a question and the set of it is possible answers are given will determine to a considerable extent how the one or ones that have so represented them can and will proceed to find the answer. In the present context, were we want to explore how Bayesian revision can handle question answering, the natural assumptions that answers are given in the form in which the Bayesian method (as I have described it) assumes they are given, viz. as credences.

Before we go on first about different forms that questions can take. Here I will consider just three types, exemplified in (38.a,b,c).

(38) a. Did Mary fire the gun?

  b. Who fired the gun?

  c. When was the gun fired?

(38.a) is an example of a 'polar', or 'yes'-'no'-question. The answer sets of polar questions always have two members, the proposition expressed by the indicative sentence of which the interrogative version is used in the question

163

and the negation of that question. (38.b) is one of a variety of 'wh-questions'. What the answer set of such a question is depends on the context in which it is asked. For instance, when there are three people – a, b and c – who could have fired the gun, and if we assume that a gun can be fired by no more than one person at a time and that only one shot was fired, then the answer set will consist of the three propositions 'a fired the gun', 'b fired the gun' sand 'c fired the gun'. But of course in somewhat different circumstances the answer set could have been a different one. (38.c) is also a wh-question, but here it is more difficult to say what its answer set is. I do not know if there is a solution to this problem that has been widely accepted. But the following strikes me as a natural option: the true firing time was anywhere between two outer limits, e.g. 11.00 pm and and 1.00 am, with a stronger probability that it it was somewhere in the middle of this interval, around midnight, than at the edges; in short, the 'answer set' should take the form of a probability distribution over the interval between eleven and one. On this analysis question (38.c) poses a problem much like the one discussed earlier about determining the true average incubation time for the virus, with the implication I drew attention to there: In such cases we have continuity not only on the side of the semantic values (the probability values), but also on the side of the entities to which probability values are assigned: the possible incubation times, or, in the present example, the time of the firing.

Assume some agent A who wants to find the answer to some question Q. How should we assume, consistently with Bayesian revision account of question answering, what representation A should have of the answer set to her question? Here is what seems a plausible shot at a proposal: (i) The contents of the credences are the propositions that Q generates according to existing accounts that can be found in formal semantics of the meanings of questions. In particular, let us assume that the contents of the answers in the answer set are mutually exclusive and jointly exhaustive. (Thus, for any two answer contents ϕ and ψ in the set $\phi \vDash \neg\psi$, and therefore that A's doxastic state contains the conditional credence $P(\neg\psi|\phi) = 0$, and assuming that $\phi_1,...,\phi_n$ are all the credence contents in the answer set, then $P(\phi_1 + ... + \phi_n) = 1$ is also part of A's doxastic state.) (ii) The degrees of the credences in the set reflect the likelihood that A currently attributes to that answer. So if A currently has no idea what the true answer to her question is, then she should attach equal doxastic degrees to the credences in the set. So if there are N possible answers in the set each should get a degree of 1/N.

With this proposal we touch on another problem for identifying beliefs as credences. Consider the following example. The Department has been given

permission to make a senior hire and has decided to invite three people, a, b and c, to apply for the position. I, a member of the Department, have read some publications by each of a, b and c and think each of them is really good and I would be thrilled to have any one of them as my colleague. But otherwise I know nothing about them and I have no idea regarding any any of them whether they can be expected to react to the invitation and apply. The day after the decision has been made to send invitations to these three people a colleague asks me what I think will happen: Which subset of $\{a,b,c\}$, including the empty set, will apply? I reply, truthfully, that as far as I am concerned anything could happen: each of the eight propositions corresponding to the subsets of $\{a,b,c\}$, as far as I am concerned, equally likely. If this attitude of mine authentically reflects my attitude towards these eight possibilities, then it would seem that, translated into the language of credences, I attribute a probability of 1/8 to each of the eight propositions. That is, for each of these propositions $p_1,..., p_8$, I have a credence of the form $<1/8,p-i>$.

Two days after the decision has been made and the invitations have been sent out, another colleague asks me if I think anyone of the three candidates is going to apply. Again I reply, in perfect honesty, that I have no idea. Whether there will be anyone who will take up the invitation is, as far as I know, fully up in the air, or 'fifty-fifty', as I might put it. What I seem to have conveyed thereby is that for me the proposition that none will apply is just as probable as that at least one of them will apply. So, according to this reaction my credence with content that none of the three will apply has degree 1/2.

Have I thereby contradicted myself, or shown myself to be incoherent? In one sense yes: by committing myself to a value of 1/2 for the proposition that no one of the three will apply in the one situation and to a value of 1/8 in the other situation I have set myself up as a sitting duck for Dutch bookies. And yet, it feels like I was being sensible in each of the two situations in which I was being asked: I really had no idea, in either situation.

The problem, it would seem, has to do with translating total ignorance with what will happen in terms of absolute likelihood values to the different things that could happen. It may make sense to express the sense that one is utterly in the dark with regard to what the answer to a given question may be in terms of what probability one assigns to each of the different things that could happen; and it may make sense to formalize having no idea what will be the case in terms of distributing the total available probability mass equally between these different possibilities. But this seems to make sense

only in particular contexts, where one is confronted with some particular set of options.

If this is right, then it must be hopeless to capture an agent's doxastic state as a set of credences. There is no context-independent way of making sense of what an agent considers more probable than what or – if that is supposed to be the same thing – which options she believes in equally firmly. But this does not prevent us from applying the Bayesian perspective to particular situations in which agents are trying to arrive at answers to particular questions. It is in this sense that the use of the Bayesian method should be pursued in the remainder of this section.

Let us set aside, then, the assumption that an agent's doxastic state can be adequately modeled as a set of credences, whose degrees reflect how likely the agent considers their content, but limit the idea that graded beliefs can be thus represented only in the contexts of questions to which the agent wants an answer and where the answer set to the question consists of credences which reflect the degrees to which the agent considers them to be the answer that is correct. In other words, a question will be represented in the agent's mind, at any time between the moment when she first conceives the question and the one, if such there be, when she has fund the answer to her question, as a set of credences which satisfy the laws of probability together with the assumption that the members of the set are mutually exclusive and jointly exhaustive. The development of the degrees of the credences in the answer set over this period of time T will then trace the agent's progress with her question: at each time in T the credence degree of each possible answer in the set captures how likely the agent thinks it is that that credence is the answer to her question.[103]

Of the three types of questions distinguished in (38) the simplest are, from several points of view, the polar questions. This is so from our present per-

---

[103]I leave the Dutch Book problem for some other occasion. I can only hope that ducks like me are savvy enough to fly off before some Dutch bookmaker targets them, even if they feel there wasn't a real inconsistency in the two reactions to the two questions. (Suppose the first colleague offers you a bet where you bet on the proposition that someone will apply where you put in \$350,- (\$50,- for each of the non-empty sets of possible applicants), while he himself puts in \$250,- on the proposition that no one will apply, and you accept, since after all you take yourself to have a 7/8 chance of winning. If the second colleague then offers you the next day a bet in which you bet \$200,- on the proposition that no one applies and he bets \$250,- on its negation, then you better not accept that bet as well, unless you have come to the conclusion that you were a fool to accept the first bet and think that accepting the second one may help you to recover some of your losses.)

spective as well. They are therefore a good place to start for an exploration of a Bayesian story about what agents can do to find answers to their questions. Consider question (38.a). As in all crime stories – the real ones as well, I suppose, as the fictional ones I have read or watched – it takes work to find out whether it was Mary. As the investigations progress, things may go back and forth until at long last decisive information is found. For instance, it may at first look very much like it was Mary, and further evidence may then be found to support this. But eventually new evidence turns up which shows that it couldn't have been Mary and all the evidence that pointed in the direction of Mary was misinterpreted. At that point the credence degree of 'Mary fired the gun' has gone down to 0, or to some number low enough to justify the conclusion that it wasn't her. (There is also a question about who the agent is in cases like this. Is it the investigating team as a whole, whose members are fully agreed at every point in time about the likelihood that Mary fired the gun, or should the case be treated as involving as many agents as team members, with the possibility that at some or all times the likelihood estimates by the different agents can be wide apart?)

Some of the evidence that comes to light during the investigation will have been looked for purposefully. The team will have verified the whereabouts of Mary between eleven and one, since information about this will have certain bearings on the question one is trying to answer. For such bits of evidence it may be plausible to assume that what happens fits the story that Bayesian revisionism wants to tell. The team have conditional probabilities of the form $P(A|B)$, where A is the proposition that Mary fired the gun and B can be anyone of a range of propositions pertaining to where Mary was during the night when the gun was fired. The proposition or propositions about where Mary was when, which are brought to light through the efforts made by the team, may then match the conditions in some of these conditional probabilities and the probability that Mary fired the gun adjusted accordingly, following the Bayesian recipe. But many other events of potentially relevant information turning up may not fit the bayesian picture even to this extent. Facts are discovered that at first don't seem to fit what has become known about what happened and then much is done to figure out how these facts within the story one is trying to reconstruct. For such discoveries it makes no sense to assume that conditional probabilities with them as conditions were already in place, and that all that needs to be done is to conditionalize on the newly discovered fact. A more plausible account of what happens in response to such discoveries is that one is trying to construct a logically coherent story (no pun intended) of the events surrounding the firing of the gun in which all that has been discovered has its place and then to see whether this story

leads to an answer for the given question: Does the story become incoherent when we add the assumption that Mary fired the gun? Or does it become incoherent when we add the assumption that she didn't fire the gun? If and when either conclusion can be drawn, the conclusion might perhaps be reconstructed in Bayesian terms as involving a conditional probability in which A (or $\neg A$, as may be the case) is conditional upon the reconstructed story (understood as complex proposition). One can then conditionalize on the story and conclude A (or $\neg A$). But that seems little more than a Bayesian's afterthought, which leaves out all the real and hard work that goes into reconstructing the case.

To sum up: A Bayesian take on what goes on in certain situations where one is looking for the answer to a question may seem appealing in that it gives us a way of accounting for the gradual nature that such efforts often take, with the agent or agents getting to the truth only in small steps, and where their progress may be highly non-linear. But the discussion of our example suggests that Bayesian revision cannot be the whole story. One sort of thing that is missing from it are stretches of reasoning of the more or less traditional, discrete sort. For cases like the one discussed above some kind of hybrid account, involving both credence revision and discrete logical reasoning, may well be what we need.

What has just been said about the polar question (38.a) applies also to the question types exemplified in (38.b,c). But there are also further complications. First (38.b) and then (38.c).

One difference between (38.b) and (38.a) is that the answer set can now have any size from 2 upwards. The question as such doesn't reveal this; it is only the context in which it is asked that may tell us what the possible answers are. And as far as that is concerned, contexts can differ not only in that different contexts will fix different sets – for instance the context could specify a choice between just two possible firers, Mary and Justin, but the set could also be a larger one, Mary, Fred, Sarah and Louise, or everyone on Mary's soccer team. In any such case the context is needed for the contraction of the different contents represented in the answer set, a process that seems to belong to the purely discrete parts of semantics and logic. What the degrees should be of these credences is another matter. As implied in the discussion of (38.a), the initial degrees can be anything, so long as mutual exclusiveness and joint exhaustivity are warranted. (And for this the Bayesian approach can't be any help either, we have seen.) But let us assume once more, that the initial credence degrees are somehow given. How the facts established in

the course of trying to answer the question then modify these should then be the responsibility of Bayesian revision, but here the same problems will arise as for the polar case above. And there will be the further complexity that now the revisions will simultaneously affect the members of a set of cardinality N, where N may be larger than 2. That doesn't add a new complication in principle, but the extra practical complexities can be considerable, and go up steeply when N increases.

But trying to find an answer to (38.b) can add a further genuine complication. There need not be a given bound to the range of possible perpetrators. Perhaps the investigation starts out with Mary, Justin and Sara as the only contenders. But it may then be discovered that Justin's aunt too might have been the one who fired the gun. At that point the answer set has to be extended, by adding a credence whose content is that Justin's aunt fired the gun, adopting a certain degree d for this credence and recalibrating the degrees for the credences in the old set by distributing them proportionally over the probability mass $1 - d$. Such extensions of the answer set may be required repeatedly during the investigation and they may also involve adding more than one possibility at a time.

One special feature of (38.b) is the assumption that the gun was fired one time an that each time a gun is fired there is just one agent who does the firing (by pulling the trigger). In this respect the case of (38.b) differs from the one with the three people a, b and c who have been invited to apply for the faculty position. Here it is possible for more than one of them to apply; as we saw, the total number of possible answers to the question 'Who will apply?' is 8, the number of the set of all subsets of the set $\{a,b,c\}$. But there are now all logical connections between the propositions in this set and propositions of the form x is the one who applied or one of those who applied. For instance the proposition 'a is the one who applied or one of those who applied' is entailed by each of those propositions in our set of eight according to which a applied. So finding out gradually more information about who has applied (and also about who hasn't) it may be possible to narrow down the answering set. (But of course, one way in which this can be formally implemented is to set the degree of the credences that have creased to be candidates to 0.)

(38.c) differs from the types of questions exemplified by (38.a) and (38.b) in that the set of possible answers is continuous rather than discrete. In the paragraph following (38) suggested that one initial determination of the answering set for (38.c) could be the temporal interval $[t_1,t_2]$, where t_1 is

11.00 pm (on the day in question) and t_2 is 1.00 am (of the following day) and where any time within this interval is considered equally probable. In probability-theoretic terms this means that we have a constant probability density over the interval $[t_1, t_2]$ – graphically, a block curve starting at t_1 and going up vertically from 0 to the constant density c, then going horizontally to t_2 and then vertically down again to 0. The constant density c is such that $\int_{t_1}^{t_2} c \, dx = 1$. That is, $c = \frac{1}{|t_1, t_2|}$. But such a probability distribution isn't a particularly plausible one, even as a starting point. It is common for people who propose such intervals was a point of departure for an investigation to have in mind that there is a greater likelihood for the time one is trying to determine to be somewhere in the middle of the interval than near the edges, and often this is made explicit, by adding something like 'more likely to have been somewhere between 11.30 and 12.30.

In other words, the distributions that people have in mind when they set out to answer questions like (38.c) are typically not distributions with constant density. It may be difficult to know how to model what sort of distributions they have in mind. My impression is that when specific assumptions are made in connection with questions such as (38.c), the distributions assumed are based on a tradition where Gaussian distributions have become a kind of default, and that also in situations where the conditions that motivate such distributions need not be given. One can get by with such assumptions nevertheless because (a) they are hard to disprove on empirical grounds, (b) Gaussian distributions are food for calculations, partly because of their intrinsic properties and partly because there is so much that is known about their mathematical properties. In any case, the typical way in which such distributions get 'improved' in the light of new evidence is for the distribution to become more 'focused' by concentrating more of its total probability mass in a small area of high probability density – the bell curve changes in that the bell gets taller and narrower – would seem to reflect some of the revisions that occur in the course of answering *when*-questions. (Revisions from a bigger to a smaller interval, as from [11.00pm,1.00am] to [11.30pm,12.30am], may be thought of as special cases of bell focusing where the part of the total probability mass distributed over the intervals between the old and the new edges – here [11.00pm,11.30pm] and [12.30pm,1.00am] has become negligible upon revision of the Gaussian.) But other kinds of distribution updates should presumably be part of the revision repertoire as well. For instance, it ought to be possible to revise a symmetric distribution (e.g. a Gaussian) to a non-symmetric one, corresponding to a new view of the time that one is trying to determine as 'between 11.00 pm and 1.00 am, but more like in the later part, after midnight'.

The use of probability distributions to capture what is Bayesian about the revisions involved in coming to as good an estimate as possible of the answer to a question of this type seems not implausible in the light of how the Bayesian method is used in science. But the question may be raised if invoking the mathematics of probability distributions and their Bayesian updates isn't conceptually misleading overkill in attempts to come to a better understanding of cognition. This question can be raised with regard to the use of probability theory in explicating mental contents and processes, for precise numbers do not seem to play a role in the consciousness of those whose contents and processes are to be explained. (Certainly most human processors never seem to testify to anything like this when you ask them, expect for some loose uses of phrases like 'fifty-fifty', 'more than a half', which probably should be understood as qualitative in spite of the superficial numerical ring there is to them. More specific numerical statements tend to come from people who know something about probability theory, to the point of also applying it to their own daily lives.) But when it comes to applying mathematical techniques having to do with probability density functions, this question seems to become more urgent. I do not quite know what to think of this objection to the treatment I suggested for questions like (38.c) and will leave the matter, in the hope to be enlightened by others.

6.6.3 Conclusions to Section 6.6

This section has been about the intrusion into logic and semantics of what I have been referring to as the 'continuous' . The dominant aspect of this intrusion has been the role of probabilities (elements from the continuous domain [0,1]) as an alternative to a discrete set of truth values (the set {True,False}). The first subsection explored this intrusion from a static perspective: what can we say about logical validity when sentences have probability values instead of values from {True,False}? We can also see this part as a discussion of doxastic states, where the beliefs are 'credences': pairs of a content and an acceptance degree. The second subsection then focused on the dynamic aspects of belief from such a perspective: How do belief states, as consisting of sets of credences, change under the influence of new information? Here we compared discrete accounts of belief change, as studied in Belief Revision, with the revision policies of Bayesianism. Our explorations of this pointed to several problems with the Bayesian view that doxastic states can be modeled as coherent sets of categorical and conditional credences, but that it might nevertheless be useful to adopt a Bayesian analysis of certain cognitive processes, such as that of finding answers to questions. Some aspects of this

171

proposal were considered in Section 6.6.2. But here we came to the conclusion that even for this restricted set of cognitive processes the Bayesian picture is incomplete. The final, tentative, conclusion was that some kind hybrid theory, in which purely discrete processing modules interact with probabilistically driven ones, may give us a more realistic account of what is going on when people's doxastic states develop over time.

As an attempt to raise the general issue of how continuity-based concepts and methods have insinuated themselves into the world of logic and semantics this section has been two things that it should not have been: (a) limited in outlook, and (b) superficial in what it has had to say about the aspects that its limited outlook includes. These shortcomings are a direct consequence of my lack of expertise in this general area. Perhaps that should have been a reason for not writing this section at all. But I believe that the issue is so important that something had to be said here, and hoping for corrections from people who may see this.

6.7 Logical relations between different description levels: Cognition, NLP, Robotics

This final subsection has not yet been written. What follows is merely a declaration of intent, and with that an acknowledgement that what is missing should also be regarded as belonging to logic, when one adopts the broad conception of 'logic' that I have tried to use as guide in this section.

6.7.1 Cognition

Linguistic competence is central aspect of human cognition and linguistic performance among cognition's most distinctive manifestations. Most of these reflections have been concerned with the logical, linguistic and psycholinguistic study of this complex of properties of the human mind and the ways it reveals itself to the world. What has not so far been touched upon is the question what is responsible for this competence and its application in observable linguistic acts at the level or levels at which we can study and describe the structure and processes of the brain, and how descriptions at that or those levels are related to descriptions of linguistic competence and performance at logical, linguistic or psycholinguistic levels. This of course is nothing but the old question about the relations between mind and brain recast in the modern form as a question about the relations between descriptions of linguistic competence and performance at the different levels that

we still associate with, respectively, mind and brain.

Often this new formulation of the mind-brain relation talks about the relations between descriptions at two levels, the 'psychological level' and the 'physiological level'. But looked at critically today, that is an oversimplification that one cannot let pass without at least making a note of it. In the preceding sections we have seen indications of different levels at which it is possible to describe linguistic competence and performances, all of which qualify as 'linguistic' descriptions, and thus a fortiori as descriptions of certain aspects of the human mind. On the other hand there are also different levels for describing the brain. We know a lot about the structure and metabolism of individual brain cells, which explains at a physiological level how cells can fire and thereby communicate with other cells. But it is also possible to study brain functions as the executions of networks of brain cells which are connected by paths along which the cells send pulses to other cells. Information processing by the brain, in a sense of information that is interpretable at the level of cognition alluded to above, should then be accounted for in terms of firing patterns.

A crucial aspect of brain cells is that their trigger potentials are adaptable, through adaptation of the synapses on its body that connect it with the dendrites coming from other cells. It is in particular this capacity for the modification of firing behavior that has become a central feature of various proposals of abstract models for neural cell networks, which are intended to capture the representation and processing of information as interpretable at a level at which we describe mental states and mental activity in psychology or linguistics. Given the enormous effort that has been made over half a century to describe aspects of the structure and use and the acquisition of language in terms of such models, it is a sobering fact how little progress has been made on this project. The disappointment has been associated in particular with the work on PDP ('Parallel Distributive Processing') developed in the eighties by Rumelhart, McClellan and others (see in particular (Rumelhart & McClelland 1986)).[104] There may have been two quite different factors at play here. For one thing, it may have been a mistake to focus on the neural network modeling of linguistic tasks in isolation, from other linguistic tasks, and, more generally, from other non-linguistic tasks with the execution of which the executions of linguistic tasks are often interwoven. On the other

---

[104]**From what I know this was the state of play until ten years ago. Is this still true? Help! (Note well, I do not include at this point the remarkable success of Deep Learning in NLP. See point 2 below.)**

173

hand it may be that the neural networks and their ability to learn and then apply what they have learned on new task-specifying inputs simply do not model brain performances in the right way; perhaps the anatomy and dynamics of ensembles of brain cells support representations and processes at the cognitive level in ways that are radically different from what the models assume.

In light of how little we still seem to know about how to describe processes of apparently connected tasks in sufficient detail (and with predications that can be experimentally checked) at either of the two description levels (the level of neural network modeling and the psychological or linguistic level) questions about the relations between descriptions at these two levels (of a non-trivial set of processes and their in- and outputs) can only be asked at a purely abstract level. That is so in particular for questions about the logical relations between descriptions of the same phenomena or processes at the cognitive and the neural network level. But these are questions that belong to logic under its present broad conception, and once the descriptions will be available to which these questions are applicable, it will of the utmost importance to find answers to them, as central issues in the philosophy of Mind and in our understanding of ourselves.

6.7.2   NLP

Since its inception some six or five decades ago Natural Language Processing has been dominated by two fundamentally different conceptions of what it is one should be doing and how to go about it. The first, to which I will be referring here as 'Compositional Linguistics' (for lack of a better name and in the awareness that this is a non-canonical use of the term) saw its task to be that of writing algorithms that could compute linguistic structure, both in the abstract and implemented as computer programs. This paradigm was dominant originally. It may have had its heyday during the nineteen nineties and the early years of the present century; but then it was overtaken by the main direction with which I am contrasting 'Computational Linguistics' here: the approach which people identify increasingly as a branch of 'Deep Learning'. Today that is the dominant approach by a long stretch and one that is still growing. (Some people, who are judges of these matters in ways that I am not, have asked the question if this dominance is going to be for good or if the pendulum may swing back within a not so distant future. I think they may well be right. But that opinion, for the little that it is worth, might be dismissed as due to personal bias and wishful thinking.)

Computational Linguistics subscribed to a view of the division of labor between 'theoretical linguistics' and 'Computational Linguistics'.[105] 'Theoretical linguistics' was responsible for the systemic description of linguistic structure, in terms of the morphology and syntax of the expressions of the given language and the semantics associated with morpho-syntactically well-formed expressions; and also, to some extent, with processes of building or recognizing the postulated structures, as for instance in the definition of syntactic parsers, but leaving the implementation to their computational colleagues. Until one has tried to go this way for as long and intensively as it has been pushed, it is difficult to see what might not be right with this approach, given the vast amount of knowledge that has been collected about the morphology, syntax and semantics that has been brought together for a substantial (and still growing) number of human languages, and especially in the light of all that has been established about English. But experience has taught us that it is much harder to build properly working automated natural language processing systems in this way than had been expected.

After the fact it is easier to see why the Computational Linguistics approach hasn't been more successful than thus far it has been. First, one wasn't sufficiently aware at the time of how complex natural languages are – how many different grammatical construction can be found, for instance in English. A formal grammar for a language that can deal with all these challenges – a 'total coverage grammar' – is extremely difficult to develop and when a computational system is the implementation of such a grammar, it is likely to crash sooner or later when let loose on arbitrary text transcribed speech. A serious problem here is also that the typical implementations of Computational Linguistics are merciless. Consider for instance an implementation of a parser for a rule-based grammar developed for a language L that does not cover all the constructions of L. When such a system is asked to parse a sentence with an uncovered construction, it will simply fail – perhaps return with a message to that effect. Parsing spoken language is another problem. Many of the utterances people actually produce do not conform to the rules of the native speakers' grammar (and would not be acceptable when written). That is on the one hand because language producers easily get their wires crossed when the sentences they utter exceed a certain degree of complex-

---

[105]'Theoretical linguistics' is an often used term in discussions on the topics of this section. It is a particularly unfortunate one insofar as it suggests that Computational Linguistics is not or less 'theoretical', a position that people don't seem to beg willing to substantiate and that I would like to emphatically deny. I am using the term 'theoretical linguistics' here with this emphatic disclaimer and only in scare quotes.

ity, but also because we often knowingly offend against the official rules, for instance because this way it will be easier for our addressees to understand what we are saying. The need for parsers that can deal with input that is ungrammatical for either of these reasons was recognized early on. But parsers with just the right tolerance for grammar violations proved to be a lot harder to build than parsers that lack tolerance.

But by far the greatest problem, I believe, was that 'theoretical linguistics' has failed to provide Computational Linguistics with a suitable form of semantics. For one thing this may have had to do with the circumstance that semantic representations are assumed to play no part in the denotational approach to meaning, which was one the core ideas behind Montague Grammar and has dominated formal semantics for the past five decades. But Montague Grammar does make use of expressions from some formal language (the Lambda Calculus and Predicate Logic as a subsystem of it) and there was no reason for Computational Linguistics not to employ those expressions. as its semantic/logical forms. But even when that was done, it didn't do the Computational systems as much goods one might have hoped. This became clear especially when the NLP community started to organize RTE competitions ('RTE' = 'Retrieval of Textual Entailment'), where the competing systems have to make correct predictions about whether a given sentence is or is not entailed by a given textual passage. One might have thought that this would be the kind of task that would show the Computational Linguistics approach to its maximal advantage, especially since so much was known about the logical forms it was using. The logical forms of sentences – lambda terms of type t – are normally equivalent to formulas of predicate logic and entailment between formulas of predicate logic has long been something about which much has been known, not only on paper, but also in the form of effective computer-implemented theorem provers. This entailment machinery was something that Computational Linguistics could take more or less off the shelf. But in spite of that performance of such systems at RTE tasks was not particularly good.

Even in RTE tasks the Computational Linguistics approach was soon outperformed by systems based on Deep Learning.[106] This is only one of many

---

[106]At some point the only Computational Linguistics system that was a serious RTE competitor for Deep Learning systems was the 'Boxer' system developed by Johan Bos and colleagues **reference to Bos**. Boxer uses DRSs as logic forms. Its performance is presumably enhanced by the fact when the system has to evaluate the claim that sentence S follows from a text bit T consisting of several sentences, a single DRS K will be computed as representation of T and the representation of S will take the form of an extension K'

NLP tasks where at the present time systems based on Deep Learning do significantly better than Computational Linguistics systems, and often by a long stretch.

Deep Learning systems are in man respects like the systems proposed by PDP in the eighties. But the crucial difference is that between the nineteen eighties and the second decade of the 21st century the power of computing machinery increased by a stunning factor, both with regard to storage of information and the speed of actual computation. So many of the applications of the neural network approach that forty or thirty years ago weren't even dreamt of are now possible and among these there are many that have to do with NLP.

It is not easy to give a description of the essence of Deep Learning. Here is an attempt from one singularly under-informed.[107] The networks used by Deep Learning today typically consist of vast numbers of cells arranged in often many layers. Since the connections between the cells can be modified when the network is run on the right kind of input – inputs of the kind to which the network is intended to be applied once it has been properly trained – the network can be trained by being run on such inputs. The training can be seen as the determination of suitable values for an often very large number of parameters, which enable the network to analyze each new input as a combination of a large number of features and compute the output in accordance with the feature combination that it attributes to the given input.

The remarkable fact about such networks is that they can discover all sorts of regularities in bits of unanalyzed text, including the regularities that linguists (and the many generations of grammarians before them) have identified as grammatical rules and principles. But in addition the networks are apparently also capable of discovering many other regularities in the bits of language that they receive as inputs. It is – must be – this that enables them to succeed where Computational Linguistics systems fail.

To someone like me, who was brought up and worked with the convictions and prejudices of formal logic as it existed when I became a student and the convictions and prejudices of formal linguistics that were established while I was a student, the successes of Deep Learning-based NLP systems is nothing short of miraculous. But the successes, such as they are, are there and they

---

of K. Validity of the entailment is then the question whether K' can be derived from K.

[107]**replace this description as soon as I can**.

need an explanation. Much of the explanation must have to do with the fact that as language is used – the ways it is in the corpora that are used to train the networks – there are many more regularities than the linguist ever dreamt of (and that perhaps he should ever dream of as linguistically relevant principles – principles that are part of the internalized grammars of the speakers of the language. But whatever the status of these regularities, as long as they are to be found in the linguistic material that is used in training and then becomes the application domain of the trained network, and keep repeating themselves in the inputs to which the network is exposed as time goes on, they can be expected to help the network make the right predictions about its outputs.

When I spoke of the successes of Deep Learning-based NLP systems I added the qualification 'such as they are'. The qualification reflects the fact that impressive as the successes may be, there remain many more language-related tasks that current systems of this kind cannot do. Not surprisingly this is so increasingly when the tasks are less 'local', i.e. when they involve larger spans of written or spoken language. The patterns that the network would have to identify to produce the wanted outputs cannot be expected to occur as often in the training material as more 'local' patterns. That means that larger corpora are needed. And somewhere along the line the technology won't be able to cope with the sheer magnitude of the training material (as well as, perhaps, the size and speed of the network itself). This becomes a problem in particular if training is unlikely to lead to useful results unless the input corpora are annotated, where the annotations serve the network as feed-back on whether the predictions it makes are correct executions of the given task.

Perhaps there are some language-related abilities that human speakers have and that can never be emulated by Deep Learning methods as currently practiced, not even if there were no limits to network speed and size and to the quantity of training materials. Whether that is or isn't so, the actual limits may, as I suggested, place their absolute limits, because of the physics of computer hardware, and the total amount of available language data, of which there can never be more than what the speech community has left behind and is adding to that on a daily basis. But what also matters is that there is another limit that Deep Learning-based NLP crossed some time ago and that it has already far behind. This is the amount of linguistic input that children need to learn their mother tongue and also the amount of language they will have been exposed to by the time they have evolved into competent speakers. This means that such systems need and use resources that are

different form those available to first language learners. That is one reason why what is going on during network training cannot be a model for human language acquisition. There is nothing wrong or disturbing about this, as long as Deep Learning-based NLP systems aren't seen as such models, which they aren't. Children learn their languages by using them in direct relation to non-linguistic concerns. They ask for the things they want or tell what has just happened to them. The 'referring' dimension of language is there from the very start. How that helps the child to learn her language may still not be very well understood in detail. But it is hard to imagine that it doesn't help, and quite probably we humans could not learn our languages without this referential dimension. Until recently Deep Learning-based NLP systems had no access to this dimension, so they have had to compensate for that by pattern extraction from vastly larger amounts of pure language than we ever need. Nevertheless, that it is at all possible to compensate in this way, even if there are limits to that, remains, to one like me, something of a miracle.

When speaking in Section 6.7.1 about PDPas providing some abstract model of the brain as a structure of connected cells I speculated about the relationship between descriptions of certain linguistic phenomena as processes executed by such structures and descriptions of those same phenomena in linguistic or psychological terms. That relationship, I ventured, is a contemporary guise of the mind-body problem. And the investigation of that relationship, I declared, should be seen as falling within the scope of logic, in the wide sense I am understanding this here.

What I have been saying about what I have been calling 'Computational Linguistics' in the present section, and the Deep Learning-based systems discussed above may suggest a similar kind of relationship: Suppose for the sake of argument that for a substantial set of phenomena we have both a 'Computational Linguistics' system and a Deep Learning-based system that get these phenomena right. (e.g. give the same, correct answers to a given range of RTE tasks). What can we then say about the relationship between the two systems? Again I am asking this question in the abstract, for as things stand there is no such pair of systems known to me. In that respect the situation is not different here than the question I raised in 6.7.1. But I suspect that there may nevertheless be a difference between the question asked there and the one I am asking here. A logical comparison of the two systems, it seems to me, should make it possible to compare how they deal with particular tasks, i.e. by what steps they arrive from particular inputs to particular outputs. But the descriptions familiar to me of Deep Learning-based systems do not seem to give this kind of access. The system is described as a kind of black

box, which after adequate training will produce its outputs for given inputs, but without giving us much if any access to the operations it had to perform to go from input to output.

Whether that leaves with any *logical* question to ask I do not dare to say.

6.7.3 Robotics

In Section 7.2.2 I suggested that what makes first language learning so effective is that from the moment the child starts to produce utterances it uses language as a means for extralinguistic purposes - for getting what it wants, to protest against what it doesn't want and so on. I also speculated that if its use of language wasn't embedded in such non-linguistic contexts, it might not be able to learn its language at all.[108]

Even if it may be true that children need non-linguistic environments to earn their languages the way they do, the mental structures and operations that must be theirs if these non-linguistic embeddings can be of any use to them have considerable complexity. This is one of the morals that can be drawn from foundational research on robots with the capacity to use language in the service of some larger goal. My thoughts here are guided by what I have seen of a project that ran at the Technical University Munich some ten years ago[109], in which the aim was to develop a robot Clara that could not only herself perform certain basic mechanical acts (such as joining two wooden slats together with a drew an bolt), but who also is able to use natural language to ask a human assistant for help and to tell him what he should do.

The problems that arise in the programming of Clara rapidly grow in complexity when one wants her to deal with a larger and more flexibly managed range of actions. But even when that range is very narrowly defined, the problems are considerable. Suppose for instance that Clara only comes into action when she is told to do so, and that there is only one command to which she can respond, viz. 'Join the red and the green slat!' But if her response is to be as intuitively it should be, then she should begin by identifying the red and the green slat (from the objects that are spread out in front of her), then make a plan for how to execute the task, deciding whether she needs

---

[108]This may sound like a counterfactual with little bite: If a child wasn't motivated to use language by exterior motives, why should it try? Though we shouldn't forget that children also engage in a lot of language play.

[109](Pross 2010)

her assistant and what the assistant is to do – e.g. pick up the slats and hold them so that the right holes in them are aligned while she herself picks up a screw and a bolt, aligns the bolt with the two holes and then inserts the screw through the two holes and then turns it until screw and bolt hold the two slats tightly together. There are quite a few non-linguistic acts that Clara herself has to perform in the execution of this plan, and each one has to be guided by the kind of motor control via sensory feedback that is essential to almost any physical action that is performed by humans (and many other animals). Such actions and the control mechanisms needed for them have been a central concern of robotics from the start and I am assuming that we can rely on existing work in our design of Clara's mind. Oner of the difficult parts of the system that still have to be built is a component that enables the robot to use language to get the assistant to participate – to let him know that he is needed and also what he is expected to do. Designing this component is a complex exercise, but it is also an illuminating one. I am assuming here that the speech acts Clara will perform in order to get her assistant on board are genuine speech acts, as opposed to 'bits of canned speech: Clara first computes the message she is to get across to her assistant – as part of her over-all plan, in her 'language of thought' – and then turns this representation of the message into words, by making use of a language production system that can translate representations from the inner language into representations of the outer language.

As said, even if we are content for the time being with a Clara whose mind is severely limited to a small repertoire of plans, a correspondingly restricted inner language for the representation of messages and a fairly small fragment of the public language (e.g. English) that suffices for turning those representations into words, building such a production system is a true challenge. And this challenge multiplies when Clara is to be able to react to replies from her assistant, so that it is possible for the two of them to have meaningful verbal exchanges about the tasks they are to jointly perform.

These and other issues were addressed in Pross' dissertation, but the dissertation didn't lead to an actual working implementation at the time and to my knowledge this gap haven't been closed since. In any case trying to work out what is needed for a robot that can perform a number of comparatively simple tasks and use language towards the realization of them remains a difficult but potentially rewarding exercise, as it shows how deeply the linguistic and non-linguistic are interwoven in much of our daily use of language. I suspect that this is a direction for further work which may give us many important clues for how human acquisition and use of language is possible.

The present section about robotics is part of a larger section on logic. But what does this have to with logic? There are two answers. First, further design explorations for language-using robots can be seen as work on the logic of planning and action. In fact, this is perhaps the best avenue for making further progress in this area. The second answer is more tentative. Robot design also gives rise to mind-body-like questions of sorts: How is a description of what the robot does, at the level at which we pitched our discussion in the last few paragraphs, related to a hardware level description of the robot's brain? But there is a difference here. Questions about 'mind-level' and 'brain level' descriptions of the wirings of robots have been the subject of a vast literature in computer science, which includes both the semantics of programming languages and the theory of complier building. High level programming languages are connected by a cascade of languages to the level of machine code. The connections between these different levels may be complex, but they have been made fully explicit and are well understood. (If it weren't for that, high level programming could never have been made to work.) When we want to think about the human mind and the human brain in terms of analogies with the structure and workings of computers, there are important things to be learned. But I expect that most of them have been learned by now. Certainly many have been extensively discussed in the A.I./Philosophy of Mind literature. (**references. Dennett?**)

There is also another long term scenario. It is not inconceivable that we may find that when the mental capacities of a robot like Clara are extended beyond some very limited repertoire, it is better to construct the robot in such a way that she can learn from her own interactions with the world, much in the way that humans do – and not as a kind of metal-mind-Minerva that is born ready to face the world with a fully developed intellectual equipment. And it might then be found that the best way to put such a robot together is to give it a neural network for a brain, perhaps with some of its functions preprogrammed, but at the same time with the ability to adjust to the needs of a world that will be revealed to her only as she interacts with it. For such a robot it will still be possible – and perhaps indispensable to a successful design – to describe her behavior in the high level terms implied in our discussion of Clara. True, we would now have to distinguish between descriptions of the robot's mental processes at different stages of her development, since these will now depend on how much experience she has behind her. But perhaps such robots too 'grow up' in the sense that once the robot is 'grown up' her mental processes no longer change all that much; they mostly vary only in response to the different challenges that are coming

from a changing world. If so, then we may think for present purposes of the high level description as targeting the mind and doings of the grown up robot.

Building a robot of this sort will require some kind of hardware to implement the neural network – *some* kind of hardware will of course be necessary. One possibility might be that of actually growing human or human-like brain tissue. That possibility is fraught with moral dilemmas, quite apart from all the technological issues involved and I want to assume that this option is out of bounds. An alternative – and the only one I am prepared to entertain – is that we use standard computer hardware for this purpose. We are then faced with two versions of our original interface question. On the one hand there is the connection between the neural network and the hardware that implements it. Once again this is a question that is sit, through the implementation of neural networks on existing computer hardware. The second question is about the connection between the high level descriptions of the structure and functioning of the robot's mind and the description of the neural network that enables it to develop and work in the way the first description articulates. That too is a question I mentioned earlier, and of which I concluded that it might be unanswerable for fundamental reasons.

Perhaps this is not a good place to stop. But is probably it is no worse than would be many others. In any case, for a part of these reflections that hasn't yet been written Section 7 has ended up rather longer than seems reasonable. Perhaps the section should be dropped altogether. But the next task is to see how the main lines and most useful and least offensive bits from all of the above can be distilled from it (and paired down to something between 5 and 10 pages, or even less). This of course will in the first instance be my task. but I need and will grateful for your help.

References

Adams, E. (1998), *A Primer of Probability Logic*, CSLI Publications.

Adams, E. & Levine, H. (1975), 'On the uncertainties transmitted from premisses to conclusions in deductive inferences', *Synthese* **30**, 429–460.

Alchourron, C., Gärdenfors, P. & Makinson, D. (1985), 'On the logic of theory change', *Journal of Symbolic Logic* **50**, 510–530.

Anderson, J. R., Bothell, D., Byrne, M., Douglass, S. & Lebiere, C. (2004), 'An integrated theory of mind', *Psychological Review* **111**, 1036–1060.

Asher, N. (1986), 'Belief in discourse representation theory', *Journal of Philosophical Logic* **15**, 127–189.

Asher, N. (1987), 'A typology for attitude verbs and their anaphoric properties', *Linguistics and Philosophy* **10**, 125–198.

Asher, N. (1993), *Reference to Abstract Objects in Discourse*, Kluwer, Dordrecht, Dordrecht.

Asher, N. & Lascarides, A. (2003), *Logics of Conversation*, Cambridge University Press.

Asher, N. & Wada, H. (1988), 'A computational account of syntactic, semantic and discourse principles for anaphora resolution', *Journal of Semantics* **6**(3), 309–344.

Barwise, J. (1981), 'Scenes and other situations', *The Journal of Philosophy* **78**, 369–397.

Barwise, J. & Perry, J. (1983), *Situations and Attitudes*, MIT Press, Cambridge Mass, Cambridge, MA.

Beaver, D. (2001), *Presupposition and Assertion in Dynamic Semantics*, Folli, CSLI Publications, Stanford Unicersity.

Bittner, M. (2014), *Temporality*, Wiley.

Brasoveanu, A. (2011), 'Sentence-internal *different* as quantifier-intrnal anaphora', *Linguistics and Philosophy* **34**, 93–168.

Brasoveanu, G. & Dotlačil, J. (forthcoming), *Computational Cognitive Modeling and Linguistic Theory*, Springer.

Chierchia, G. & Rooth, M. (1984), Configurational notions in drt, *in* 'Proceedings of NELS 14', Amherst.

Cooper, R. & Kamp, H. (1991), Negation in situation semantics and discourse representation theory, *in* 'Situation Theory and Its Applications, Volume 2', CSLI Lecture Notes, pp. 311–335.

Demey, L., Kooi, B. & Sack, J. (2019), Logic and probability, *in* 'Stanford Encyclopedia of Philosophy', CSLI.

Earley, J. (1970), 'An efficient context-free parsing algorithm', *Communications of the ACM* **13(2)**, 94–102.

Eijck, J. v. & Stokhof, M. (2006), The gamut of dynamic logics, *in* D. Gabbay & J. Woods, eds, 'The Handbook of History of Logic, Vol 7: Logic and the Modalities in the Twentieth Century', Elsevier, pp. 499–600.

Elbourne, P. (2005), *Situations and Individuals*, MIT Press, Cambridge Mass.

Farkas, D. (2002), 'Specificity distinctions', *Journal of Semantics* **19**, 1–31.

Fine, K. (1975), 'Vagueness, truth and logic', *Synthese* **30**, 265–300.

Fine, K. (1985), *Reasoning With Arbitrary Objects*, Blackwell, Oxford.

Fraurud, K. (1990), 'Definiteness and the processing of noun phrases in natural discourse', *Journal of Semantics 7(4)* **7**, 395–433.

Gallin, D. (1975), *Intensional and Higher Order Logic; with Applications to Montague Semantics*, North Holland, Amsterdam.

Ganesalingam, M. (2013), *The Language of Mathematics: A Linguistic and Philo- sophical Investigation*, Springer. FoLLI Publications on Logic, Language and Information.

Gärdenfors, P. (1988), *Knowledge in Flux: Modelling the Dynamics of Epistemic States*, MIT Press, Cambridge Mass.

Geurts, B. (1999), *Presuppositions and Pronouns*, Elsevier Science.

Groenendijk, J. & Stokhof, M. (1984), Studies on the semantics of questions and the pragmatics of answers, PhD thesis, University of Amsterdam, Amsterdam.

Groenendijk, J. & Stokhof, M. (1990), Dynamic montague grammar, *in* L. Kalman & L. Polos, eds, 'Papers from the Second Symposium on Logic and Language', Akademiai Kiadoo, Budapest, pp. 3–48.

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

Groenendijk, J. & Stokhof, M. (1991*b*), Two theories of dynamic semantics, *in* J. v. Eijck, ed., 'Logics in AI—European Workshop JELIA '90', Springer Lecture Notes in Artificial Intelligence, Springer, Berlin, Berlin, pp. 55–64.

Gupta, A. (1982), 'Truth and paradox', *Journal of Philosophical Logic* **11**, 1–60.

Gupta, A. & Belnap, N. (1993), *The Revision Theory of Truth*, Bradford Books.

Hales, T. (2019), 'An argument for controlled natural languages in mathematics', *Thales* .

Halliday, M. A. K. & Hasan, R. (1976), *Cohesion in English*, Longman. London.

Hansson, S. (1999), *A Textbook of Belief Dynamics. Theory Change and Database Updating.*, Kluwer.

Harel, D. (1984), Dynamic logic, *in* D. Gabbay & F. Guenthner, eds, 'Handbook of Philosophical Logic', Reidel, Dordrecht, pp. 497–604. Volume II.

Haug, D. (2014), 'Partial dynamic semantics for anaphora: Compositionality without syntactic coindexation', *Journal of Semantics* **31**, 457–511.

Heim, I. (1982,1988), *The Semantics of Definite and Indefinite Noun Phrases*, Garland Press.

Heim, I. (1984), Artikel und definitheit, *in* A. v. Stechow & D. Wunderlich, eds, 'Semantics: An International Handbook of Contemporary Research', de Gruyter, pp. 487–535.

Herzberger, H. (1982), 'Naive semantics and the liar paradox', *Journal of Philosophical Logic* **79**, 479–497.

Higginbotham, J. (1983), 'The logic of perceptual reports: An extensional alternative to situation semantics', *Journal of Philosophy* **80**, 100–127.

Jasinskaja, K. & Rossdeutscher, A. (2012), Through narrative planning towards the preverbal message: A drt-based approach, *in* A. Benz, M. Stede & P. KÃ¼hnlein, eds, 'Representing and inferring discourse structure', John Benjamins, New York, pp. 45–75.

Kamp, H. (1975), Two theories about adjectives, *in* E. Keenan, ed., 'Formal Semantics for Natural Language', Cambridge University Press, Cambridge, pp. 123–155. Reprinted in: K. von Heusinger and A. Ter Meulen, 'Meaning and the Dynamics of Interpretation. Selected Papers of Hans Kamp'. Brill, 2013.

Kamp, H. (1981), A theory of truth and semantic representation, *in* J. Groenendijk et al., eds, 'Formal Methods in the Study of Language', Mathematisch Centrum, Amsterdam. Reprinted in: K. von Heusinger and A. Ter Meulen, 'Meaning and the Dynamics of Interpretation. Selected Papers of Hans Kamp'. Brill, 2013.

Kamp, H. (1983), A scenic tour through the land of naked infinitives. Manuscript.

Kamp, H. (1988), Comments on stalnaker, belief attribution and context, *in* R. Grimm & D. Merrill, eds, 'Contents of Thought', University of Arizona Press.

Kamp, H. (1990), Prolegomena to a structural account of belief and other attitudes, *in* C. A. Anderson & J. Owens, eds, 'Propositional Attitudes—The Role of Content in Logic, Language, and Mind', University of Chicago Press and CSLI, Stanford, chapter 2, pp. 27–90. Reprinted in: K. von Heusinger and A. Ter Meulen, 'Meaning and the Dynamics of Interpretation. Selected Papers of Hans Kamp'. Brill, 2013.

Kamp, H. (1999), Intentions, plans and their execution: Turning objects of thought into entities of the external world. ms. Stuttgart.

Kamp, H. (2001), The importance of presupposition, *in* C. Rohrer, A. Rossdeutscher & H. Kamp, eds, 'Linguistic Form and its Computation', CSLI, Standord. Reprinted in: K. von Heusinger and A. Ter Meulen, 'Meaning and the Dynamics of Interpretation. Selected Papers of Hans Kamp'. Brill, 2013.

Kamp, H. (2003), Einstellungszustände und einstellungszuschreibungen in der diskursrepräsentationstheorie, *in* U. Haas-Spohn, ed., 'Intentionalität zwischen Subjektivit"at und Weltbezug', mentis, Paderborn, pp. 209–289.

Kamp, H. (2015), 'Using proper names as intermediaries between labelled entity representations', *Erkenntnis* **80**, 263–312.

Kamp, H. (2017), 'Events, discourse representations and temporal reference', *Semantics and Pragmatics* **10(2)**, 1–45. A French translation of this article appeared in *L*angages, 1981.

Kamp, H. (2019*a*), Entity representations and articulated contexts. an application to the analysis of deictic demonstratives, *in* J. Alden Pepp & J. Almog, eds, 'Themes from Kaplan, Thirty Years Later,', OUP.

Kamp, H. (2019*b*), Semantics i, university of texas. Unpublished notes for a graduate course in Linguistics offered at the University of Texas, Austin.

Kamp, H. (2020), Sharing real and fictional reference. To appear in E. Maier and A. Stokke (eds.) 'Language in Fiction'.

Kamp, H. & Bende-Farkas, A. (2019), 'Epistemic specificity from a communication-theoretic perspective', *Journal of Semantics* **36 (1)**, 1 – 51.

Kamp, H. & Reyle, U. (1993), *From Discourse to Logic*, Kluwer.

Kamp, H. & Reyle, U. (2011), Discourse representation theory, *in* K. von Heusinger, C. Maienborn & P. Portner, eds, 'Handbook of Semantics', De Gruyter.

Kamp, H. & Roßdeutscher, A. (1994), 'Drs-construction and lexically driven inferences', *Theoretical Linguistics* **20**, 165 – 236.

Kamp, H., van Genabith, J. & Reyle, U. (2011), Discourse representation theory: An updated survey, *in* D. Gabbay, ed., 'Handbook of Philosophical Logic, Vol. XV', Elsevier, pp. 125 – 394.

Kaplan, D. (1989), Demonstratives, *in* H. W. J. Almog, J Perry, ed., 'Themes from Kaplan', Oxford University Press.

Kripke, S. (1975), 'Outline of a theory of truth', *Journal of Philosophy* **72**, 690–715.

Kühlwein, D., Cramer, M., Koepke, P. & Schröder, B. (2008), The naproche system, *in* 'Intelligent Computer Mathematics - 11th International Conference, AISC 2012, 19th Symposium, Calculemus 2012, 5th International Workshop, DML 2012, 11th International Conference, MKM 2012, Systems and Projects, Held as Part of CICM 2012, Bremen, Germany, July 8-13, 2012. Proceedings.', Springer.

Lascarides, A. & Asher, N. (1993), 'Temporal interpretation, discourse relations, and common sense entailment', *Linguistics and Philosophy* **16**, 437–49.

Lorenzen, P. & Lorenz, K. (1978), *Dialogische Logik*, WBG, Darmstadt.

Maier, B. G. . E. (2013), Layered discourse representation theory, *in* A. Capone, F. L. Piparo & M. Carapezza, eds, 'Perspectives on Linguistic Pragmatics: The Resource Logic Approach', Springer, pp. 311–327.

Mann, W. & Thompson, S. (1987), 'Rhetorical structure theory: A framework for the analysis of texts', *International Pragmatics Association Papers in Pragmatics, I, 79-105* **1**, 79–105.

Montague, R. (1963), 'Syntactic treatments of modality, with corollaries of reflexion principles and finite axiomatizability', *Acta Philosophica Fennica* **16**, 153–167.

Montague, R. (1970), 'Universal grammar', *Theoria* **36**, 373–398. Reprinted in: R. Thomason (ed), Formal philosophy; selected papers of Richard Montague, Yale University Press, 1974.

Murray, S. (2014), 'Varieties of update', *Semantics and Pragmatics* **7(2)**, 1–53.

Muskens, R. (1991), Anaphora and the logic of change, *in* J. v. Eijck, ed., 'Logics in AI / European Workshop JELIA '90 / Amsterdam, The Netherlands, September 1990 / Proceedings', Lecture Notes in Artificial Intelligence 478, Springer Verlag, pp. 412–427.

Muskens, R. (1994*a*), A compositional discourse representation theory, *in* P. Dekker & M. Stokhof, eds, 'Proceedings 9th Amsterdam Colloquium', ILLC, Amsterdam, pp. 467–486.

Muskens, R. (1994*b*), Tense and the logic of change, *in* U. E. et al., ed., 'Interface Aspects of Syntax, Semantics and the Lexicon', W. Benjamins.

Muskens, R. (1996), 'Combining montague semantics and discourse representation', *Linguistics and Philosophy* **19**, 143–186.

Neale, S. (1990), *Descriptions*, Bradford Books, MIT Press.

Nederpelt, R. P., Geuvers, J. H. & de Vrijer, R. C. (1994), *Selected Papers on Automath*, Elsevier. Vol. 133 of Studies in Logic and the Foundations of Mathematics.

Potts, C. (2005), *The logic of conventional implicature*, Oxford University Press.

Pross, T. (2010), Grounded Discourse Representation Theory, PhD thesis, University of Stuttgart.

Reyle, U. (1993), 'Dealing with ambiguities by underspecification: Construction, representation and deduction', *Journal of Semantics* **10**, 123–179.

Roberts, C. (1987), Modal Subordination, Anaphora, and Distributivity, PhD thesis, University of Massachusetts, Amherst.

Rumelhart, D. & McClelland, J. (1986), *Parallel Distributed Processing. Explorations in the Microstructure of Cognition*, MIT Press, Cambridge, Mass.

Schilder, F. (1998), An underspecified segmented discourse representation theory (usdrt), *in* 'Proceedings of the 17-th International Conference on Computational Linguistics (COLING '98) and of the Annual Meeting of the Association for Computational Linguistics (ACL '98)', Université de Montréal.

Searle, J. R. (1969), *Speech Acts: An Essay in the Philosophy of Language*, Cambridge University Press, Cambridge.

Simons, M., Tonhauser, J., Beaver, D. & Roberts, C. (n.d.), What projects and why, *in* 'Proceedings of SALT 20, Year = 2010, Address = Vancouver, British Columbia, Editor = N. Li and D. Lutz, Pages = 309–327 Morgan Kaufman'.

Slobin, D. (1987), Thinking for speaking, *in* N. Li & D. Lutz, eds, 'Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society', Berkeley (Ca).

Stalnaker, R. (1988), Belief attribution and context, *in* R. Grimm & D. Merrill, eds, 'Contents of Thought', University of Arizona Press.

Stone, M. (1998), Modality in Dialogue: Planning, Pragmatics and Computation, PhD thesis, University of Pennsylvania.

Van Den Berg, M. (1996), Some Aspects of the Internal Structure of Discourse. The Dynamics of Nominal Anaphora., PhD thesis, University of Amsterdam.

Van Der Sandt, R. (1992), 'Presupposition projection as anaphora resolution', *Journal of Semantics* **9**, 333–377. Special Issue: Presupposition, Part 2.

Van Der Sandt, R. & Geurts, B. (1991), Presupposition, anaphora, and lexical content, *in* O. Herzog & C.-R. Rollinger, eds, 'Text Understanding in LILOG', Springer.

Venhuizen, N. (2015), Projection in Discourse: A Data-driven Formal Semantic Analysis, PhD thesis, Groningen University, Groningen.

Venhuizen, N., Bos, J., Hendriks, P. & Brouwer, H. (2014), How and why conventional implicatures project, *in* 'Proceedings of SALT 24', CSLI, Palo Alto, pp. 63–83.

Venhuizen, N., Bos, J., Hendriks, P. & Brouwer, H. (2018), 'Discourse semantics with information structure', *Journal of Semantics* **35**, 127–169.

Vermeulen, C. (1994), Explorations of the Dynamic Environment, PhD thesis, OTS, Utrecht.

Zeevat, H. (1989), 'A compositional approach to discourse representation theory', *Linguistics and Philosophy* **12**, 95–131.

Zinn, C. (2003), 'Computational framework for understanding mathematical discourse', *Logic Journal of IGPL* **11**(July), 457–484.