Language and Cognition: Our Mental Representations of what we read and hear.

Hans Kamp

University of Stuttgart
University of Texas, Austin

March 19, 2018

The Main concerns of the seminar

- To develop ways of setting up a theory of natural language meaning which:

  a. retains the explicitness and formal rigor of Formal Semantics as it was initiated by Richard Montague.

  b. allows mental representations of content into accounts of how contents are expressed in language and communicated to an addressee or an audience.

  Representations of the mental state of speaker and hearer have proved to be of special importance in the theory of reference; and in particular to the question:

  How is reference effected by the different types of definite noun phrases?

The Main concerns of the seminar

N.B. In English the different types of definite NPs are:

(i) proper names

(ii) definite descriptions

(iii) the 'indexical' first and second person pronouns *I* and *you*

(iv) demonstratives (*that bird over there*, *this chair*, *that one* and so on)

(v) the third perseon pronouns *he*, *she*, *it*.

The Main concerns of the seminar

- To develop an account of the mental representation of content that:

  a. can explain how agents may be expected to react to new information (which can be either verbal or non-verbal), for instance by forming new plans of action or modify old ones.

  b. can serve as the foundation for an account of *attitude reports* – sentences and multi-sentence discourses whose purpose is to describe the mental states of agents.

  N.B. The attitude reports that are still most often discussed within philosophy are sentences like

  'X believes that ϕ', 'X desires that ϕ', in which an attitudinal verb (*believe*, *desire* etc) is followed by a sentential complement.

  But the repertoire of sentences we use to report attitudes is much more varied. Also, we often use more than one sentence..

Some remarks on the nature and history of Formal Semantics

- The standard (and still widespread) picture of languages that is presupposed by classical Formal Semantics:

  A language L is an autonomous, user-independent system, with a vocabulary (lexicon), a syntax and a semantics.

  Therefore a proper description of a language L must:

  (i) specify its vocabulary (e.g. by listing its members).

  (ii) define its syntax (typically in the form of a set of 'generative' principles, which generate increasingly complex expressions from the vocabulary).

  (iii) spell out the semantics for the grammatically well-formed expressions of L, and in particular for its sentences.

Some remarks on the nature and history of Formal Semantics

- The semantics is typically given in the form of a *model-theoretic semantics*, which specifies truth values of the well-formed sentences of L in different models.
  The specification makes use of the recursive structures that the syntax of L assigns to its well-formed expressions.

  Some of the details of how model-theoretic semantics can be implemented will be important in what follows.

  A central distinction is that between:

  (i) *Semantic Value* theories. These specify, for each model M from the model class for L the semantic values in M of all the well-formed constituents of the sentences S of L, including on the one hand the sentences themselves and on the other constituents that are not sentences (but could be names, predicates etc).

Some remarks on the nature and history of Formal Semantics

(ii) Logical Form theories. These proceed in two stages:

(a) specify a formal language – the *Logical Form Formalism*, or LFF – along the lines of language specification described above, including a Semantic Value semantics of the kind indicated under (i).

(b) assign to each sentence S of L as *logical form* a formula $LF(S)$ from this LFF. The semantics of S is thereby identified with the (independently defined) semantics of $LF(S)$.

Logical Form theories may differ substantial form each other with regard to how they specify the assignments of Logical Forms to sentences (and possibly other expressions) of L.

In fact, many Semantic Value theories can also be regarded as Logical Form theories: they can so long as they assign expressions of some formal language to the well-formed expressions of L

Some remarks on the nature and history of Formal Semantics

An illustration of the difference:

How to deal with the sentence

The European Union surrounds Switzerland.

[More is meant to be said here. Will fill this out at a later time.]

Basics of Discourse Representation Theory (DRT)

- DRT is a Logical Form Theory.

  It has been making use of various types of LFFs.

  These are its so-called *DRS languages*.

  DRSs (short for 'Discourse Representation Structures') are DRT's content representations.

- It is important to distinguish between (i) the design of DRS-languages, and (ii) the use of some given DRS-language in giving the semantics for (fragments of) natural languages.

Basics of Discourse Representation Theory (DRT)

- The original motivation for DRT was to find a satisfactory treatment of certain *inter-sentential semantic links* for which Formal Semantics at the time had no proper tools.

  Some examples:

  (1) a. John proved a well-known conjecture in twenty pages. Mary proved it in ten pages.
  b. John proved well-known conjecture in twenty pages. Mary had proved it in ten pages.

  (2) a. When Alan opened his eyes he saw his wife who was standing by his bedside. She smiled.
  b. When Alan opened his eyes he saw his wife who was standing by his bedside. She was smiling.

Basics of Discourse Representation Theory (DRT)

- Another well-known problem for classical formal Semantics are so-called ‘donkey-sentences’ and ‘donkey-discourses’:

(3) a. If Peter owns a donkey, he beats it
    b. Peter owns a donkey. He beats it.

Basics of Discourse Representation Theory (DRT)

John proved a well-known conjecture in twenty pages.

$e \quad t \quad j \quad y$
$t \prec n \quad e \subseteq t \quad$ John'(j) $\quad$ conjecture'(y) $\quad$ well-known'(y)
e: prove'(j,y)
'in-twenty-pages'(e)

Basics of Discourse Representation Theory (DRT)

Mary had proved it in ten pages.

$$<\{ \begin{array}{|c|}\hline t''? \\ \hline t'' \prec n \\ \hline \end{array}_{\,TPpt}, \begin{array}{|c|}\hline v? \\ \hline \text{non-human}(v) \\ \hline \end{array}_{\,an.pr.3d.sg.} \}, \begin{array}{|c|}\hline e' \quad t' \\ \hline TPpt := t'' \\ t' \prec TPpt \quad e' \subseteq t' \\ e'\text{: prove'}(m,v) \\ \text{`in-ten-lines'}(e') \\ \hline \end{array} >$$

Basics of Discourse Representation Theory (DRT)

John proved a well-known conjecture in twenty pages. Mary had proved it in ten pages.

$$\begin{array}{|c|}
\hline
e \quad t \quad j \quad y \quad e' \quad t' \quad m \quad t'' \quad v \\
\hline
t \prec n \quad e \subseteq t \quad \text{John'}(j) \quad \text{conjecture'}(y) \quad \text{well-known'}(y) \\
e : \text{prove'}(j,y) \\
\text{`in-twenty-pages'}(e) \\
\\
t'' = dur(e) \quad t' \prec t'' \quad e' \subseteq t' \quad \text{Mary'}(m) \quad v = y \\
e'\text{: prove'}(m,v) \\
\text{`in-ten-pages'}(e') \\
\hline
\end{array}$$

The Unity of Content and Context

- One feature of early DRT was the 'Unity of Content and Context':

  The very same structures that serve as semantic representations of the sentences or sentence sequences from which they have been derived also serve a discourse contexts for the semantic representation of the next sentence.

  This seemed to suggest that there was a genuine psychological plausibility to the DRT account of discourse anaphora:

  That the structure of DRSs captures some of the essential properties of the representations that human interpreters construct of the information they obtain through what they hear or read.

The Unity of Content and Context

- Support for this conclusion seemed to be that the following form a tight package:

  (i) the form of DRSs

  (ii) the rules of the *DRS Construction Algorithm*, which defines how those DRSs are constructed, and

  (iii) the way in which the Construction Algoritm exploits already contracted DRSs as discourse contexts

- The thought was:

  If this works as well as it does, then that is presumably because it is also how we do it.

The Unity of Content and Context

- But are we to say about this inference to psychological relevance in the light of what happened since?

  Other Construction Algorithms have been developed since the original formulation of DRT.

  These work quite differently from the original one.

  This is so in particular the Algorithm responsible for the temporal DRSs shown last week.

  This Algorithm first constructs a preliminary DRS for a given sentence and then, from this one the full DRS.

  (It is only during this second step that the discourse context is brought into play.)

The Unity of Content and Context

- As a matter of fact the importance of the structural form of DRSs and the plausibility that this form tells us something about human semantic representation are not affected by these changes.

  In particular, which entities are represented by discourse referents in the discourse context DRS and which are not is just as crucial for the new Construction Algorithms as it was for the original one;

  and which entities are represented as discourse referents and which not hasn't changed either.

  (The various Construction Algorithms do not differ from each other on this point.)

The importance of not being represented

- An important consideration in this debate has to do with entities that are *not* represented as discourse referents in a DRS, although the DRS entails their existence.

Here are two of the classical examples:

(4) (Partee)

a. One of the ten marbles is not in the bag. It is under the sofa.

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

The importance of not being represented

$$\begin{array}{|c|}\hline x \quad Y \quad z \\ \hline |x| = 1 \quad x \in Y \quad |Y| = 10 \quad \text{marble'*}(Y) \quad \text{bag'}(z) \\ \neg \begin{array}{|c|}\hline \\ \hline \text{in'}(x, z) \\ \hline \end{array} \\ \\ \begin{array}{|c|}\hline x' \\ \hline x' \in Y \\ x' \neq x \\ \hline \end{array} \; \left\langle \begin{array}{c} \forall \\ x' \end{array} \right\rangle \; \begin{array}{|c|}\hline \\ \hline \text{in'}(x', z) \\ \hline \end{array} \\ \hline \end{array}$$

The importance of not being represented

$$
\begin{array}{|c|}
\hline
X \quad Y \quad z \\
\hline
|X| = 9 \quad X \subseteq Y \quad |Y| = 10 \quad \text{marble'*}(Y) \quad \text{bag'}(z) \\
\\
\begin{array}{|c|}\hline x' \\ \hline x' \in X \\ \hline \end{array}
\left\langle \begin{array}{c} \forall \\ x' \end{array} \right\rangle
\begin{array}{|c|}\hline \\ \hline \text{in'}(x', z) \\ \hline \end{array} \\
\\
\neg \begin{array}{|c|}\hline x'' \\ \hline x'' \in Y \quad \neg \boxed{x'' \in X} \quad \text{in'}(x'', z) \\ \hline \end{array} \\
\hline
\end{array}
$$

The importance of not being represented

(5) a. Half of the shareholders didn't go to the annual meeting.
They found out about the decisions the next morning from the media.

b. Half of the shareholders went to the annual meeting.
They found out about the decision the next morning from the media.

The importance of not being represented

(6) a. Two of the ten marbles are not in the bag. They are under the sofa.

b. Eight of the ten marbles are in the bag. They are under the sofa.

The importance of not being represented

- The examples involving plural pronouns are especially remarkable.

  For the discourse referents that are needed as antecedents for plural pronouns need not be present in the discourse context as given, but may be constructed from other discourse referents.

(7) John took Mary to Acapulco. There they met Peter and Annabel.
The next day they set off on their sailing trip.

(8) (Partee)
When John comes to visit, we play duets. But sometimes he brings a cellist along and then we play trios.

The importance of not being represented

- DRS for the sentence 'John took Mary to Acapulco.'

$t \quad e \quad j \quad m \quad a$
$t \prec n \quad e \subseteq t \quad \text{john'}(j) \quad \text{mary'}(m) \quad \text{acapulco'}(a)$ $e : \text{take'}(j, m, a)$

  The plural pronoun *they* of the second sentence ('There they met Peter and Annabel') gets its antecedent by forming the complex DRT term '$j \oplus m$'.

- Amalgamation through the formation of complex DRT terms is also involved in the interpretation of the other occurrences of plural pronouns in the two examples from the last slide as well as in the next example.

The importance of not being represented

(9) Bill has found most of the books that Susan asked him to look for. He has put them on her desk.

- Note well: amalgamation operations are permitted in the construction of anaphoric antecedents for plural pronouns but a 'negative' operation like set subtraction is not.

  We cannot for instance form the subtraction of the set of two marbles that are in the bag from the set of all ten marbles to get an antecedent for *they* in the second sentence of (9.b).

  (This should not be allowed, since *they* in (9.b) cannot be used as referring to the two missing marbles.)

  (9.b) Eight of the ten marbles are in the bag.
  They are under the sofa.

The importance of not being represented

- But let us not forget:

  These are questions of detail that concern just one particular kind of noun phrase, viz. third person pronouns.

  Note that in (9.b) we can refer to the two missing marbles by using a definite description like *the (two) missing marbles*.

  The interpretation of this definite description also requires the DRS from the first sentence as discourse context.

  But in this case the discourse context is exploited in a different way, viz. as a source of information from which the unique satisfaction of certain predicates may be deducible.

Two ways of looking at DRT

- As a way of doing formal semantics of natural language

  Following a ‘Logical form’ approach rather than a ‘direct semantic value’ approach

- As a theory of *natural language interpretation*, which tells us something about the ways in which human interpreters build representation of what they hear or read.

- An argument in favor of this second view:

  The Unity of Content and Context.

Reference in the philosophy of logic and language

- There has been a tendency to reduce reference to satisfaction.

  This tendency is natural in the context of formalizing mathematics; mathematics is mostly concerned with *general propositions* (e.g. about arbitrary numbers of certain kinds, e.g. arbitrary primes).

- This changed only partially when Strawson insisted that an elementary speech act is composed of two quite distinct acts: (i) referring and (ii) predicating.

Reference in the philosophy of logic and language

- A dramatic change came in the late sixties and early seventies with what Kripke had to say about names, Kaplan about demonstratives and Donnellan about referential uses of definite descriptions.

  But this was about direct reference relations between expressions and objects.

  Little if anything was said about the mental representation of entities.

  (In part this was no doubt because mental representation was frowned upon generally in the Philosophy of language and Semantics).

Reference in the philosophy of logic and language

- DRT, as a theory that has got something to say about mental representation, can be seen as a first hint at the importance of the representation of entities.

  However, as a theory of the mental representation of entities it is arguably not radical enough.

  Partly this is because discourse referents have to play a number of distinct roles, of which the representation of particular entities is only one.

The structure of mental states

- We will now start with the development of a proposal for part of the structure of mental states, which is loosely inspired by DRT.

  We will eventually embed this proposal in an extension of DRT, called 'MSDRT' (for 'Mental State DRT').

  But that will be a second step.

- Consider the case of a person A walking along the sidewalk, seeing what she thinks is a gold coin lying in the middle of the road, forms the desire to have this object and forms the intention to go to the middle of the road and pick it up.

  What follows are two proposals for the relevant part of the mental state of A, consisting of her belief, desire and intention.

The structure of mental states

$$\left\{ \begin{array}{l}
\left\langle BEL, \begin{array}{|c|}\hline x \quad s_1 \quad s_2 \\ \hline n \subseteq s_1 \quad n \subseteq s_2 \\ s_1 : gold\ coin(x) \\ s_2 : x\ be\ lying\ in\ front\ of\ i \\ \hline \end{array} \right\rangle \\[2ex]
\left\langle DES, \begin{array}{|c|}\hline s_3 \\ \hline n \subseteq s_3 \\ s_3 : i\ have\ x \\ \hline \end{array} \right\rangle \\[2ex]
\left\langle INT, \begin{array}{|c|}\hline t \quad e \\ \hline n < t \quad e \subseteq t \\ e : i\ pick\ up\ x \\ \hline \end{array} \right\rangle
\end{array} \right\}$$

The structure of mental states

- Note the reuse of the dref x, which is 'declared' in the Universe of the BEL DRS, in the DES DRS and the INT DRS.

  This is a potential problem for the semantics of the this state representation:

  What could be the propositional content determined by these DRSs (the DES DRS and the INT DRS)?

  Referential dependence will play a major part in what follows.

- There is also another way in which we might describe the mental state of our agent, in which we make us of an *Anchored Entity Representation* for the thing the agent perceives (or thinks she is perceiving):

The structure of mental states

$$
(10)\ .\left\{
\begin{array}{l}
\left\langle [ANCH, x], \begin{array}{|c|}\hline x \quad s_1 \quad s_2 \\ \hline n \subseteq s_1 \quad n \subseteq s_2 \\ s_1 : i\ see\ x \\ s_2 : x\ be\ lying\ ifo\ i \\ \hline \end{array} \right\rangle \\[2em]
\left\langle BEL, \begin{array}{|c|}\hline s_3 \\ \hline s_3 : gold\ coin(x) \\ \hline \end{array} \right\rangle \\[2em]
\left\langle DES, \begin{array}{|c|}\hline s_4 \\ \hline n \subseteq s_4 \\ s_4 : i\ have\ x \\ \hline \end{array} \right\rangle \\[2em]
\left\langle INT, \begin{array}{|c|}\hline t\ e \\ \hline n < t\ e \subseteq t \\ e : i\ pick\ up\ x \\ \hline \end{array} \right\rangle
\end{array}
\right\}
$$

The structure of mental states

A first proposal for an Anchored Entity Representation (AER):

$$\left\langle\ [ANCH, x]\ ,\ \begin{array}{|c|}\hline x \quad s \quad s' \quad s'' \\ \hline n \subseteq s \quad s: i\ see\ x \\ n \subseteq s' \quad s': coin(x) \\ n \subseteq s'' \\ (s'': 1.5cm \leq diameter(x) \leq 2.5cm) \\ \hline \end{array}\ \right\rangle$$

The dref x that occurs together with 'ANCH' in an AER is called its *distinguished discourse referent*

- The *internalist reduction* of a mental state description that contains AERs is the conversion of its AERs into beliefs that the there is some entity that is represented by the distinguished dref of the AER which satisfies the anchoring DRS of AER in question.

  Example: Internalist reduction of the mental state description two slides back.

The structure of mental states

(11) .
$$\left\{ \begin{array}{l}
\left\langle BEL, \begin{array}{|c|}\hline x \quad s_1 \quad s_2 \\ \hline n \subseteq s_1 \quad n \subseteq s_2 \\ s_1 : i\ see\ x \\ s_2 : x\ be\ lying\ ifo\ i \\ \hline \end{array} \right\rangle \\[2em]
\left\langle BEL, \begin{array}{|c|}\hline s_3 \\ \hline s_3 : gold\ coin(x) \\ \hline \end{array} \right\rangle \\[2em]
\left\langle DES, \begin{array}{|c|}\hline s_4 \\ \hline n \subseteq s_4 \\ s_4 : i\ have\ x \\ \hline \end{array} \right\rangle \\[2em]
\left\langle INT, \begin{array}{|c|}\hline t\ e \\ \hline n < t\ e \subseteq t \\ e : i\ pick\ up\ x \\ \hline \end{array} \right\rangle
\end{array} \right\}$$

Some background about the model theory of DRT

<u>*Def* 1</u> (Extensional models for our basic DRS language L_0)

An *extensional model M for L_0* is a structure

$< T, U, EV, LOC, \approx, Name, Pred >$, where

i. T is a time structure <T, $\prec$>, consisting of a non-empty set of instants T and a precedence order $\prec$ on T; $\prec$ is assumed to be a total ordering.

IT will be the set of intervals of T.

ii. U is a function which assigns to each t $\in$ T the set U_t of 'entities that exist at t in M.

Some background about the model theory of DRT

iii. EV is an event structure, i.e. a triple $<$EV, $<$, $O>$. with EV $=$ E $\cup$ S: a set of eventualities which consists in one part of events (the set E) and in another of states (the set S);

$<$ and O are relations of complete precedence and overlap on the set EV, satisfying some obvious postulates:

(i) $<$ is irreflexive and transitive, (ii) O is reflexive and symmetric; (iii) for any two eventualities ev_1 and ev_2 either $\mathrm{ev}_1 < \mathrm{ev}_2$ or $\mathrm{ev}_1 \; O \; \mathrm{ev}_2$ or $\mathrm{ev}_2 < \mathrm{ev}_1$; (iv) for any four eventualities ev_1, ev_2, ev_3, ev_4, if $\mathrm{ev}_1 < \mathrm{ev}_2 \; O \; \mathrm{ev}_3 < \mathrm{ev}_4$, then $\mathrm{ev}_1 < \mathrm{ev}_4$.

(iv) LOC is a function which maps the members of EV onto intervals of T (i.e. onto members of IT ; intuitively LOC(ev) is the temporal interval occupied by the eventuality ev).

(v) $\approx$ is the relation of same duration between temporal intervals.
($\approx$ is an equivalence relation on IT)

Some background about the model theory of DRT

(vi) $Name$ assigns to each name from L_0 an element u from $U = \cup_{t \in T} U_t$.

(vii) $Pred$ assigns to the set Pr of predicates of L_0 the following kinds of values:

when N is a nominal predicate of L_0, then $Pred(N) \subseteq U$;

when V is an n-place verbal event predicate, then $Pred(V)$ is a set of tuples $<$e, u_1, ..., $u_n$$>$, where e $\in$ E and u_1, ..., $u_n \in U$;

when V is an n-place verbal state predicate, then $Pred(V)$ is a set of tuples $<$s, u_1, ..., $u_n$$>$, where s $\in$ S and u_1, ..., $u_n \in U$.

Some background about the model theory of DRT

- A DRS K from L_0 that represents an utterance made at the time t_u from the time structure of a model M *is true in* M *as representation of an utterance made at* t_u iff there is a verifying embedding f of K in M such that $f(n) = \mathrm{t}_u$.

  A *verifying embedding of* K *in* M is a function f that maps the Universe U_K of K into the Universe U_M in such a way that the Conditions of K are all satisfied in M under the assignment that f provides for the discourse referents in U_K.

  Instead of the mouthful 'K is true in M as representation of an utterance made at t_u' we will often say, more succinctly: 'K is true in M at t_u'

Some background about the model theory of DRT

Def 2 An *intensional model for* L_0 is a structure M

$< W, T, U, EV, LOC, \approx, Name, Pred >$, where

(i) W is a non-empty set (of 'possible worlds');

(ii) Each of $T, U, EV, LOC, \approx, Name, Pred$ is a function whose domain is W.

(iii) For each $w \in W$, $< T_w, U_w, EV_w, LOC_w, \approx_w, Name_w, Pred_w >$ is an extensional model in the sense of $Def 1$.

We refer to this extensional model as M_w.

An intensional model M for L_0 is *temporally uniform* iff for all $w, w' \in W_M$, $T_w = T_{w'}$.

Some background about the model theory of DRT

- Let M be an intensional model, $w \in W_M$, $t_u \in T_{M_w}$, K a DRS from L_0.

  Then K is *true in M in w at t_u* iff there is a verifying embedding f of K in M_w such that $f(n) = t_u$.

  <u>*Def* 3</u> Let M be an intensional model. A *proposition relative to* M is a subset of W_M.

  <u>*Def* 4</u> Let M be a (temporally uniform) intensional model for L_0, K a DRS from L_0, t_u an instant from the time structure of M.

  The *proposition expressed by K in M with respect to* the time t_u is the set of all worlds $w \in W_M$ such that K is true in M in w at t_u.

  We denote this proposition as $[[K]]_{M,t_u}$.

Some background about the model theory of DRT

- A DRS is much like an open formula; the drefs in its Universe function in a certain sense as ‘free variables’.

  This shows in particular in the way truth is defined for DRSs, via functions that embed their Universe in the Universe of a given model.

  Suppose that M is an intensional model, K a DRS. Then we can consider besides the proposition expressed by K in M also the *open proposition determined by K in M at t_u*.

  This is the set of all pairs $<w, f>$ such that $w \in W_M$ and f is verifying embedding of K in M_w such that $f(\mathrm{n}) = t_u$

  Note well that the functions f which occur as second members in pairs $<w, f>$ belonging to the open proposition determined by K in M all have the same domain.

  For each such f, $\mathrm{Dom}(f) = U_K$.

Some background about the model theory of DRT

- Note that the open proposition determined by K in M is in general more informative than the proposition that K expresses in M.

  For the open proposition captures all the different ways in which K can be verified in M_w, and not just whether there is any verification of K in M_w.

- Open propositions – sets of pairs $<w, f>$ – are also known as *information states.*

  The notion of an *information state* is one of the central notions of Dynamic Semantics as it was developed by the 'Amsterdam School' in the eighties and nineties.

  (A central role in this development was played by Groenendijk, Stokhof and Veltman, but others made important contributions too.)

Some background about the model theory of DRT

- As noted, the open proposition determined by a DRS K in an intensional model M consists of pairs $<w, f>$ in which all functions f have the same domain.

  In Dynamic Semantics this property is not always assumed, but the open propositions/information states that we will be using will always have it.

  An open proposition/information state I that has this property can be said to have a *Base.*

  More formally, let M be an intensional model and I an open proposition/information state relative to M.

  The *Base of* I is the common Domain of all the functions f such that $<w, f>$ belongs to I.

Some background about the model theory of DRT

- There is one case for which this definition does not work.

  This is the case where I is the empty open proposition or information state, also called the *inconsistent open proposition/information state* – the open proposition/information state $\emptyset$, which isn't satisfiable in any world of M.

  Since this state doesn't contain any pair $<w, f>$, there is no f from which the Base can be recovered.

  For technical reasons it will be convenient to distinguish between empty information states with different Bases.

  For this reason we identify *the empty open proposition/information state with Base* X with the pair $<\emptyset, X>$.

  Non-empty information states, for which the base can also be recovered from the pairs they contain will be identified as sets of such pairs, in accordance with the definition above.

Some background about the model theory of DRT

- There is an alternative way of characterizing information states.

  Instead of defining them as sets of pairs $<w, f>$, we can also define them as functions from worlds to the sets of verifying embeddings for those worlds:

  Suppose that M is an intensional model and I a set of pairs $<w, f>$, with $w \in W_M$ and f a function into U_{M_w}.

  We call such functions from worlds to function sets *functional information states.*

  Each open proposition I can be transformed into the functional information state J defined by:

  for each $w \in W_M$, $J(w) = \{f :<w, f>\in IS\}$.

  We also can get back from J to I via $I = \{<w, f>: f \in I(w)\}$.

Some background about the model theory of DRT

- The operation going from open propositions I to functional information states J is called *lifting* and denoted as I^*.

- The operation that goes from functional information states J to open propositions I is called *flattening* and denoted as $\overline{J}$.

Some background about the model theory of DRT

- We have got ourselves into a bit of a terminological tangle at this point.

  On the one hand we have two terms – 'open proposition' and 'information state' – for the same notion.

  And on the other we have the somewhat unwieldy term 'functional information state' to denote a related but different construct.

  Let us simplify this by:

  (i) using for the constructs that in the above we have been denoting alternatively with 'open proposition' and 'information state' only as 'open propositions

  (ii) refer to functional information states also with the simpler expression 'information state'.

Some background about the model theory of DRT

- Our primary need in what follows will not be for open propositions and functional information states as just defined, but for more general notions:

  that of a *partial open proposition* and that of a *partial functional information state*.

  The need for these arises for more that one reason.

  One of these has to do with resolving anaphoric presuppositions of preliminary sentence DRSs K in the discourse context provided by some other DRS K'. Recall for instance the example:

  (1.b) John proved a well-known conjecture in twenty pages.
  Mary had proved it in ten pages.

- We repeat the DRS for the first sentence and show the result of resolving the anaphoric presuppositions of the second sentence.

Some background about the model theory of DRT

$e \quad t \quad j \quad y$
$t \prec n \quad e \subseteq t \quad$ John'(j) $\quad$ conjecture'(y) $\quad$ well-known'(y)
e : prove'(j,y)
'in-twenty-pages'(e)

$e' \quad t' \quad m \quad t'' \quad v$
$\mathbf{t'' = dur(e)} \quad t' \prec t'' \quad e' \subseteq t' \quad$ Mary'(m) $\quad \mathbf{v = y}$
e': prove'(m,v)
'in-ten-pages'(e')

The bold-faced conditions of the second DRS K_2 contain drefs (e and y) that do not occur in its Universe.

This makes K_2 an *improper* DRS. It cannot be evaluated for truth in the usual way because embedding functions that are defined just on its Universe won't provide values for its free dref occurrences.

Some background about the model theory of DRT

- It is only in the context provided by some other information state that such a DRS can be properly evaluated.

  More specifically, K_2 can be evaluated against the background of any open proposition with the property that its Base contains the drefs e and y.

  Suppose that M is an intensional model and that OP is an open proposition relative to M with a base X such that $\{e,y\} \subseteq X$.

  Then K_2 can be evaluated for every $w \in W_M$ such that for some f $< w, f > \in OP$:

  For each such w K_2 is true in w relative to OP iff there is a $< w, f > \in OP$ and a mapping g of the Universe of K_2 into U_{M_w} such that $f \cup g$ is a function which verifies all the conditions of K_2 in M_w.

Some background about the model theory of DRT

- Another setting in which the need for partial propositions arises is that of mental state descriptions with Anchored Entity Representations.

  Anchored Entity Representations are proper only when their internal anchors are witnesses of some *external anchor*, to the object that the ER represents.

  In other words, a state description $\mathbb{K}$ with one or more such ERs is proper, with respect to an intention model M and a world w of M, only if there is a function A which maps (or 'anchors', as we will also say) the drefs of its anchored ERs to objects from the Universe U_{M_w}.

  In this case $\mathbb{K}$ will be evaluable only in those extensional models $M_{w'}$ belonging to M in which the external anchors of $\mathbb{K}$ exist – that is: at those $M_{w'}$ such that $\mathrm{Ran}(A) \subseteq U_{M_{w'}}$

Some background about the model theory of DRT

- There is more than one way in which partial open propositions and partial functional information states can be formally defined.

  The following definitions have proved useful for our needs.

  We start with the definition of *partial (functional) information state.*

  $\underline{Def\ 6}$ Let M be an intensional model, X a set of discourse referents.

  A *partial functional information state with base X relative to M* is a function J with the following properties:

  (i) Dom(J) is some open proposition I relative to M such that Base(I) $\subseteq X$.

  (ii) For each $<w, f> \in I$, $J(<w, f>)$ is a set of functions $g \supseteq f$ with Dom(g) $= X$ which map the discourse referents in X onto objects from M.

Some background about the model theory of DRT

- When J is a partial functional information state, we refer to the Domain of J also as J's *presupposition*.

- Functional information states can be seen as partial functional information states a special kind.

  To make this explicit we must first define the notion of the *trivial open proposition relative to an intensional model M*, $\top_M$:

  The *trivial open proposition relative to* M, $\top_M$, is the set $\{<w, \emptyset>: w \in W_M\}$.

  N.B. Intuitively, $\top_M$ is the open proposition relative to M that doesn't carry any information whatever:

  (i) $\top_M$ doesn't exclude any worlds from W_M and (ii) $\top_M$ does't make any commitment to 'given' discourse referents (since its Base is empty).

Some background about the model theory of DRT

- Using $\top_M$ we can 'recode' J as the partial functional information state J' whose domain is $\top_M$ and which assigns to each pair $<w, \emptyset>$ in $\top_M$ the same value that J assigns to w.

- The notion of a *partial open proposition* is a generalization of that of an open proposition in much the same sense in which the notion of a partial functional information state is a generalization of that of a functional information state.

  Partial open propositions are like open propositions except that they come with (usually non-trivial) presuppositions.

  As in the case of partial functional information states, these presuppositions are open propositions.

Some background about the model theory of DRT

- Formally:

  A *partial open proposition relative to* M is a pair $<Pr, P>$, where Pr is an open proposition relative to M and P is an open proposition relative to M which extends Pr

  (i.e. for each $<w, g> \in P$ there is an $f \subseteq g$ such that $<w, f> \in Pr$).

- Open propositions can be seen as a special kind of partial open propositions, just as functional infermation states can be seen as a special kind of partial functional information state:

  > The open propositions P relative to M correspond one-to-one with the partial open propositions whose presupposition is the trivial open proposition $\top_M$
  >
  > (i.e. those partial olden propositions that are of the form $<\top_M, P>$).

Some background about the model theory of DRT

- The operations of flattening and lifting can also be applied to partial open propositions and partial information states:

  If J is a partial functional information state, then the flattening of J is the pair $<$Dom(J),$\overline{J}$$>$.

  If $<Pr, P>$ is a partial proposition, then its lifting is the partial functional information state J defined by:

  (i) $\text{Dom}(J) = Pr$;

  (ii) for $<w, f> \in Pr$, $J(<w, f>) = \{g\text{: } <w, g> \in P \ \ \& \ \ f \subseteq g\}$.

Some background about the model theory of DRT

- Each partial open proposition and each partial information state involves two Bases.

  One of these is the Base of the partial open proposition or information state itself and one is the Base of its presupposition.

  For a partial open proposition $<PR, P>$ the Base of the partial open presupposition itself is the set Dom(g), where g is any embedding function that occurs as second member of a pair $<w, g>$ of P for some w.

  Furthermore the Base of the presupposition Pr of $<PR, P>$ is the set Dom(f) such that for some w $<w, g>$ belongs to Pr or X iff the presupposition is $<\emptyset, X>$.

  N.B. The first base is not defined when P is empty. We could repair this by making provisions for this special case, but the effort doesn't seem worth it.

Some background about the model theory of DRT

- For partial information states the Base of the presupposition is defined as for partial open propositions.

  The Base of a partial information state J is the set $\text{Dom}(g)$ where g belongs to any set $J(<w, f>)$ where $<w, f>$belongs to the Domain of J.

- We have defined partial propositions as pairs $<Pr, P>$.

  Here Pr is an open proposition and P an open proposition that can be seen as a 'reinforcement' of P: P entails Pr.

  So when we drop Pr from $<Pr, P>$ the open proposition P that remains can be thought of as the result of 'accommodating' Pr and then conjoining it with P .

Some background about the model theory of DRT

- Likewise for the presuppositions of partial information states:

  We have identified the presupposition of a partial information state J with its domain.

  Here presupposition accommodation leads to an open proposition P_J, defined by

  For all $w \in W_M$ and all functions g such that $\mathrm{Dom}(g) = \mathrm{Base}(J)$ $<w, g> \in P_J$ iff there is an f such that $g \in J(<w, f>)$

- Note that our use here of the term ‘presupposition accommodation’ conforms to the established understanding of this term.

Some background about the model theory of DRT

- This is easy to see for partial information states

  When we go from the partial information state J to the open proposition P_J, we remove the restriction that the Domain of J imposes on its truth-conditional evaluability.

  At worlds w such that for no f $<w, f> \in \mathrm{Dom}(J)$ P_J is defined, but false.

  This is precisely what presupposition accommodation amounts to, as the term is ordinarily understood:

  The propositional content of the presupposition is added to the content of what presupposes it by stipulating that the latter content is junsatisfiable in all worlds in which the presupposition fails

Some background about the model theory of DRT

- The same applies, *mutatis mutandis*, to partial open propositions.

  Going from the partial open proposition $<Pr, P>$ to the open proposition P by eliminating the presupposition Pr has the following effect:

  the worlds w incompatible with Pr (and thus outside the applicability range of P in $<Pr, P>$) now become worlds incompatible with the open proposition P.

- Note that there is an important difference between the operations of lifting and flattening on the one hand and presupposition accommodation on the other¿

  Lifting and flattening are reversible operations.

Some background about the model theory of DRT

- But presupposition accommodation is not.

  By amalgamating a content with its presupposition the distinction between what is presupposed and what holds given that the presupposition is satisfied is obliterated.

- The transition from a partial information state to an open proposition via presupposition accommodation is a crucial ingredient to the semantics for IAADRSs that will presented below.

  Consider for instance the content specifications K_{BEL} and K_{DES} of the BEL and DES components of our first example of an ADRS (see slide 33).

  K_{DES} is referentially dependent on K_{BEL} via the dref x.

Some background about the model theory of DRT

- In the semantics that will be given below, this dependence will be captured by assigning K_{DES} a partial information state whose Domain (and thus presupposition) is the open proposition that is assigned to K_{BEL}.

  Accommodation of this presupposition – forming, intuitively speaking, the conjunction of the contents of the belief and the desire – will play a central part in any account of practical reasoning based on the description of mental states as ADRSs.

  The intuitive justification for this is that practical reasoning by someone whose mental state can be described by an ADRS will be limited in the following way:

  Inferences involving desires that are referentially dependent on certain beliefs will be made only on the assumption that those beliefs are true (and the presuppositions of the desires are thus satisfied).

Some background about the model theory of DRT

- We will need on further auxiliary notion of a technical nature..

  Components of mental state descriptions can be referentially dependent on other components of the description.

  One component $<MOD, K>$ of a mental stated description is *referentially dependent on* another component $<MOD', K'>$ of the description iff the content DRS K of the first component contains a dref which is free in K but also occurs in K' and is bound there.

  (In this sense of referential dependence the DES component and the INT component of our first example of a mental state description are both referentially dependent on the BEL component .)

Some background about the model theory of DRT

- Suppose that $\mathbb{K}$ is a mental state description, that $<MOD, K>$ and $<MOD', K'>$ are components of $\mathbb{K}$ and that $<MOD, K>$ is referentially dependent upon $<MOD', K'>$.

  Then we write: $<MOD', K'> \prec_{\mathbb{K}} <MOD, K>$

- Often a component will be referentially dependent on more than one other component.

  In such cases we need to be able to form the *'union* of the semantic values of all the component DRSs that a given component of the description referentially depends on.

Some background about the model theory of DRT

<u>*Def 7*</u> Let M be an intensional model and let $\mathbb{I}$ be a set of open propositions relative to M.

The *merge* of the open propositions $I \in \mathbb{I}$ is the open proposition $\cup\, \mathbb{I}$ relative to M defined by:

$\cup\, \mathbb{I} = \{<w, h>$: there exists a function F with Domain$(F) = \mathbb{I}$ which maps each I in $\mathbb{I}$ to the second component f of some member $<w, f>$ of I such that $h = \cup\, \{F(I) : I \in \mathbb{I}\}\}$

N.B. The functions F spoken of in this definition. collect 'verifying' embeddings $F(I)$ in w from each of the open propositions I in into a single joint 'merge' embedding h in w for the entire family $\mathbb{I}$ of I's.

Articulated Discourse Representation Structures

- We now proceed to a model-theoretic semantics for mental state descriptions that are like our descriptions of the 'gold coin' scenario.

  First a definition of the syntax of such descriptions.

  Def 8 An *Internally Anchored Articulated DRS* ($IAADRS$) is a set of pairs of the form $< MOD, K >$, where

  (i) MOD is an 'mode indicator';

  (ii) K is a DRS

  N.B. In what follows we will be mostly concerned with the attitudinal mode indicators BEL, DES, INT and the Anchored Entity Representation indicators [ANCH,x].

  When $\langle\ [ANCH, x]\ , K\ \rangle$ is an internally Anchored Entity Representation, then x is called its *distinguished discourse referent.*

Articulated Discourse Representation Structures

- The distinguished drefs of different Entity Representations that belong to the same Internally Anchored Articulated DRS must all be different from each other.

  This is necessary in order that when such a dref occurs as an argument within the content DRS of some other component of the IAADRS it points towards a unique AER in the IAADRS that determines the value of this argument.

  If $\mathbb{K}$ is an IAADRS, we denote as '$\mathrm{AN}_{\mathbb{K}}$' the set of internally anchored drefs of $\mathbb{K}$, i.e. those drefs x that occur as the distinguished drefs of AERs $\langle\ [ANCH, x]\ , K\ \rangle$ in $\mathbb{K}$.

Articulated Discourse Representation Structures

- Let $\mathbb{K}$ be an IAADRS, AN the set of its internally anchored drefs and M an intensional model.
- An *external anchor for* $\mathbb{K}$, *relative to* M *and* a world w of M is a function **a** which maps some subset of AN into the Universe U_{M_w} of the model M_w.
- Where $\mathbb{K}$, M and **a** are as above, $<\mathbb{K}$, **a**$>$ is called an *externally anchored ADRS, relative to* M.
- $<\mathbb{K}$, **a**$>$ is called *properly anchored in* M iff **a** is defined for all drefs in AN.

Articulated Discourse Representation Structures

- Let $\mathbb{K}$ be an IAADRS, AN the set of its internally anchored drefs and M an intensional model.
- An *external anchor for* $\mathbb{K}$, *relative to* M *and* a world w of M is a function $\mathbf{a}$ which maps some subset of AN into the Universe U_{M_w} of the model M_w.
- Where $\mathbb{K}$, M and $\mathbf{a}$ are as above, $< \mathbb{K}, \mathbf{a}>$ is called an *externally anchored ADRS, relative to* M.
- $< \mathbb{K}, \mathbf{a}>$ is called *properly anchored in* M iff $\mathbf{a}$ is defined for all drefs in AN.

Articulated Discourse Representation Structures

- Anchored Entity Representations are the result of interactions between the agent and her environment.

  Therefore an external anchor **a** should always map the discourse referents in its domain to entities that exist in the world in which the agent lives and interacts with her environment (in particular: in the ways that give rise to anchored entity representations).

  We will therefore assume from now on that one of the worlds in the world set of an intensional model plays the part of the actual world w_0.

  External anchors will always be mappings to entities that exist in tho world (and thus belong to the Universe of the model M_{w_0}).

Articulated Discourse Representation Structures

- Internally Anchored ADRSs should be *coherent*.

  Coherence has to do with *referential dependence*.

  When one component $< MOD, K >$ is referentially dependent on another component $< MOD', K' >$, then $< MOD', K' >$ should not be dependent on $< MOD, K >$.

  In general, for an Internally Anchored ADRS $\mathbb{K}$ to be coherent, the referential dependency relation between its components must be

  (i) *well-founded* and

  (ii) it must be in accordance with the *attitudinal hierarchy*.

The Attitudinal Hierarchy

- If you believe in the existence of something, then you can form the desire to stand in a certain relation to that thing.

  The relation can be of any number of sorts, depending on whether your attitude towards things of the kind of the thing you believe exists (e.g have the thing, find the thing, meet the thing, avoid the thing, destroy the thing.)

- Your belief that something of a certain kind exists may also give rise to further beliefs about that thing.

  For instance, if you have come to believe, through the media or because someone told you, that there just was an earthquake of 7.9 on the Richter scale somewhere in South America you may associate with that belief a second one to the effect that some people died in it.

The Attitudinal Hierarchy

- But referential dependence of a belief on a desire is not coherent.

  When you desire there to be a certain thing, then you cannot form a belief about the desired thing just on the strength of that desire.

- You can form a conditional belief: a belief about the thing you desire ***provided your desire comes true***.

  But that will then a referentially independent belief.

  Its content will be that if the content of the desire is or will be true, then the thing in question – which in this case will exist – will have further properties (e.g. that it will give you pleasure).

  (A kind of donkey pronoun content!)

The Attitudinal Hierarchy

- Intentions can also be referentially dependent on beliefs.

  For instance, if you believe that a thing of a certain kind exists, then you may form an intention or plan to go and find that thing.

  (Recall the legend about Ponce de Leon, who believed there was a/the Fountain of youth in Florida and formed the intention to look for it (and then implemented that intention by going on an expedition to search for it).

  Apparently the story is a fabrication, but it can serve as an illustration all the same.

The Attitudinal Hierarchy

- Can a belief be referentially dependent on an intention?

  That is perhaps a more difficult question than the referential dependency of belief on desire, which as far as I can see just does not seem coherent.

  Can I for instance given my intention to make something of a certain kind, form a belief of what that thing will be like?

  No. Not, I think, in the sense in which I understand referential dependence between different attitudes.

  I may associate with my intention a strong belief that I will carry it out, so that a thing of the intended kind will comes into being.

  But then it is this belief which supports further beliefs about th thing that I intend to make, and not the intention itself.

The Attitudinal Hierarchy

- Can an intention be referentially dependent on a desire?

  Here I think the answer is yes.

  I can form the desire for there to be something of a certain kind and/or to which I stand in a certain relation, e.g that of using it for a certain purpose.

  And I can then resolve to make the thing I need, or to look for it in what may seem to me a plausible place (given what kind of thing it is that I desire)

The Attitudinal Hierarchy

- Often an intention is the result of a combination of a desire and one or more beliefs.

  Recall our first example, in which the intention to pick up the coin that the agent perceives (or thinks she perceives) was supposed to result from the belief that there is something lying in the middle of the road, the belief that it is a gold coin and the desire to have that object.

  In this case it is arguably the belief on which both the desire and the intention are referentially dependent.

  (Our representation of the mental state of our agent who walks along the road and (thinks she) sees an object in the middle of the road that she takes to be a gold coin assumed such a dependency of intention on belief.)

  But cases in which an intention referentially depends directly on a desire do seem to exist as well.

The Attitudinal Hierarchy

- It is common for the contents of propositional attitudes to be referentially dependent on Anchored Entity Representations .

  That in fact is the most important way in which AERs and propositional attitude components of mental states interact.

  (This view will be central to much that we will discuss in the seminar.)

  It is also possible for the anchor specification of one AER to be referentially dependent on another AER.

  Here is one example. Jones has an AER for his colleague Smith, whom he often interacts with at a professional level.

The Attitudinal Hierarchy

- On a given occasion Smith tells Jones that he has to cancel a meeting because he has to take his daughter to the hospital.

  This informs Jones of the fact that Smith has a daughter (the one that on that occasion Smith had to tale to the hospital).

  I will argue later at some length on that verbal communications like what Smith said to Jones on the given occasion can be a legitimate ground for the recipient to form an AER of an individual mentioned (such as in this case Smith's daughter).

  In such cases the anchor specification of the recipient's AER will refer to the source by using the AER that the recipient has for the source.

  In particular, Jones' AER for Smith's daughter will be referentially dependent on Jones' AER for Smith.

The Attitudinal Hierarchy

- These are just some elements of a theory of referential dependence between components of propositional attitudes.

  As far as I know there is still much about referential dependence that has to be explored and thought through.

  For one thing the questions about referential dependence become more complex the more Mode Indicators are included in the investigation.

  For instance, how do the propositional Mode Indicators DOUBT, HOPE, FEAR, WONDER interact with those of our minimal set {BEL, DES, INT, [ANCH,α]}?

The Attitudinal Hierarchy

- Another dimension of this problem is how the mentioned Mode Indicators interact with Mode Indicators for non-propositional attitudes (which we briefly discussed last week in class).

  I am thinking of attitudes like fear, love or abhorrence, which web can have towards things, without any propositional content – e.g. a belief content – that is an essential part of them).

- Yet another direction in which the investigation can be extended arises when attitudinal strength is allowed for.

  For instance, we might distinguish between between beliefs of different strength – or *credences*, as they are often referred to – and admit a family of doxastic Mode Indicators $\{ \text{BeL}_r \}$, where r ranges over some set of strength indicators.

  An extreme but often mentioned case of this is that in which r is assumed to range over the real number interval [0,1].

The Semantics of IAADRSs

- The next definition gives the semantics of IAADRSs.

  Before we give the definition itself, first an observation that could have been made earlier but that has become crucial at this point.

  The embedding functions that are involved in the semantic values of the different components of mental state descriptions, given as ADRSs should provide values to all the discourse referents that occur within any given DRS.

  For only then can the DRS be semantically evaluated relative to the assignments (of entities in the model to discourse referents) that the function provides.

The Semantics of IAADRSs

- The DRSs that occur as constituents of IAADRSs can contain two types of discourse referents whose values are fixed 'in advance'.

  These are:

  (i) the self-reflective discourse referents i and n, and

  (ii) the distinguished discourse referents of externally anchored AERs.

  The values of i and n are fixed 'indexically':

  The value of i is the agent whose mental state is being described.

  The value of n is the time at which she is in this mental state.

  The values of the distinguished discourse referents of externally anchored AERs are – obviously – their external anchors.

The Semantics of IAADRSs

- Suppose that an IAADRS has a proper external anchor $\mathbf{a}$

  (that is, $\mathbf{a}$ provides external anchors for all the AERs that are components of the IAADRS).

  Then an evaluation should treat each distinguished discourse referent of an AER in the IAADRS as having the value that is assigned to it by $\mathbf{a}$.

  The simplest way to deal with these additional sources of values for discourse referents is this:

  Assume that all functions referred to in the semantic value definition for IAADRSs are defined for the set $\{i,n\} \cup$ AN (in addition to the Universes of the DRSs those functions apply to).

  Moreover, for the discourse referents in $\{i,n\} \cup$ AN all functions return the values just described.

The Semantics of IAADRSs

- We now turn to the definition of the semantic values ofof IAADRSs.

  Def 8 Let $\mathbb{K}$ be a coherent IAADRS, AN the set of its internally anchored drefs, M an intensional model, $\mathbf{a}$ a proper external anchoring of $\mathbb{K}$ relative to M and t an instant from the time structure of M.

  The *semantic values* of the components $<Mod, K>$ of $\mathbb{K}$ in M at t given $\mathbf{a}$, $[[[K]]]_{M,\mathbb{K},\mathbf{a},t}$, are defined as follows:

  (i) Suppose $<MOD, K>$ has no predecessors in $B_{\mathbb{K}}$.

  Then $[[[K]]]_{M,\mathbb{K},\mathbf{a},t} = ([[K]]_{M,\mathbf{a},t})^*$.

  (i.e. the functional information state corresponding to the open proposition expressed by K in M given $\mathbf{a}$ at t).

The Semantics of IAADRSs

(ii) Otherwise, let P be the open proposition which is the merge of all the open propositions determined by the components $<MOD', K'>$ of $\mathbb{K}$ such that $< Mod', K' > \leq_{\mathbb{K}} < MOD, K >$.

(That is, $P = \cup \, \{\overline{[[[K']]]_{M,\mathbb{K}',\mathbf{a},t}}\}$ where the K''s in this set range over the set consisting of all the content representations of those components $<MOD', K'>$).

Then $[[[K]]]_{M,\mathbb{K},\mathbf{a},t}$ is the functional information state whose domain is P and which assigns to each $<w, f> \in P$ the set $\{g : g \supseteq f \;\; \& \;\; \mathrm{Dom}(g) = \mathrm{Dom}(f) \cup U_K \;\; \& \;\; g$ verifies K in M at w relative to $\mathbf{a}\}$

The Semantics of IAADRSs

- This definition is not easily penetrable. I hope the following example will help.

  Consider the following case.

  A child – let us call him Hamid – is in an orphanage but does not believe he is an orphan. In fact he believes that he has a mother and a father who are both alive. He wants to find each of them and to bring the three of them, his mother, his father and himself, together.

- The next slide displays one way in which Hamid's mental state can be described in the form of an ADRS.

  Note that the first DES DRS of this ADRS is referentially dependent on the first BEL DRS and the second DES DRS on the second BEL DRS. Moreover, the third DES DRS referentially depends on both the first and the second BEL DRSs.

The Semantics of IAADRSs

$$
(12)\ .\left\{
\begin{array}{c}
\left\langle BEL, \begin{array}{|c|}\hline m \quad s_1 \quad s_2 \\ \hline n \subseteq s_1 \quad n \subseteq s_2 \\ s_1 : mother(m,i) \quad s_2 : alive(m) \\ \hline \end{array} \right\rangle \\[2ex]
\left\langle BEL, \begin{array}{|c|}\hline d \quad s_3 \quad s_4 \\ \hline n \subseteq s_3 \quad n \subseteq s_4 \\ s_3 : father(d,i) \quad s_4 : alive(d) \\ \hline \end{array} \right\rangle \\[2ex]
\left\langle DES, \begin{array}{|c|}\hline e_1 \\ \hline n \prec e_1 \\ e_1 : find(i,m) \\ \hline \end{array} \right\rangle \quad \left\langle DES, \begin{array}{|c|}\hline e_2 \\ \hline n \prec e_2 \\ e_2 : find(i,d) \\ \hline \end{array} \right\rangle \\[2ex]
\left\langle DES, \begin{array}{|c|}\hline e_3 \\ \hline n \prec e_3 \\ e_3 : bring-together(m,d,i) \\ \hline \end{array} \right\rangle
\end{array}
\right\}
$$

The Semantics of IAADRSs

- Let M be an intensional model that consists of a single extensional model. That is, the world set of M consists of a single world w_0.

  We assume that Hamid, the referent of the indexical i, belongs to the Universe of M_{w_0} and that the time t from the time structure of M is the time at which Hamid is in the mental state described by the ADRS from the last slide.

  Note that since the ADRS $\mathbb{K}$ of the previous slide contains no AERs, no external anchoring **a** plays a part in the semantic evaluation of the content DRSs of its components.

  So we can simplify the terms $[[[K]]]_{M,\mathbb{K},\mathbf{a},t}$ denoting the semantic values denoted by the content DRSs of the components of $\mathbb{K}$ to '$[[[K]]]_{M,\mathbb{K},t}$'.

The Semantics of IAADRSs

- For ease of discussion let the content DRSs of the five components of $\mathbb{K}$ be K_1, ..., K_5, going from the top down.

  Furthermore, let F_1, ..., F_5 be the semantic values that K_1, ..., K_5 determine in M. (That is, $F_i = [[[K_i]]]_{M,w',\mathbb{K}}$ for i = 1,..., 5.).

- We are now going to say in more detail what the semantic values in M are of the five terms F_1, ..., F_5. We start with F_1.

- F_1 ($= [[[K]]]_{M,w',\mathbb{K},t}$) is that functional information state the Domain of which consists just of the world w_0 (i.e. $\mathrm{Dom}(F_1) = \{w_0\}$).

The Semantics of IAADRSs

- When applied to its only argument F_1 returns the set of all verifying embeddings f of K_1 in M_{w_0} such that $f(i)$ = Hamid, $f(n) = t$.

  (As noted above, these embeddings have to be defined for the self-reflective drefs i and n as well as for the drefs that occur explicitly in the Universe of K_1. So $\mathrm{Dom}(f) = \{i, n, m, s_1, s_2\}$.)

  Thus the verifying embeddings f that belong to the set returned by F_1 for its one argument w_0 are those functions f such that (in addition to the constraints that $f(i)$ = Hamid and $f(n) = t$):

  (i) $t \subseteq_{T_M} f(s_1)$, (ii) $t \subseteq_{T_M} f(s_2)$, (iii) $< f(s_1), f(m), f(i) > \in I_M(\text{mother})$ and (iv) $< f(s_2), (m) > \in I_M(\text{alive})\}$.

- Note that $\overline{F_1}$ is the set $\{< w_0, f >: f \in F_1(w_0)\}$.

The Semantics of IAADRSs

- The story about the semantic value determined by F_2 is completely analogous to that for F_1.

  So we just record:

  $F_2(w_0) = \{f\colon \text{Dom}(f) = \{i, n, d, s_3, s_4\}\ \ \&\ \ f(i) = \text{Hamid}\ \ \$\ \ f(n) = t$ $\&\ \ t \subseteq_{T_M} f(s_3)\ \ \&\ \ t \subseteq_{T_M} f(s_4)\ \ \&\ \ < f(s_3), f(d), f(i) > \in I_M(\text{father})$ and (v) $< f(s_4), (d) > \in I_M(\text{alive})\}$;

- $\overline{F_2}$ is the set $\{< w_0, f >\colon f \in F_2(w_0)\}$.

The Semantics of IAADRSs

- Next F_3.

This is the semantic value of the first DES DRS K_3 of $\mathbb{K}$.

It is referentially dependent just on the first BEL DRS.

So its value is a function whose Domain is $\overline{F_1}$.

Recall that $\overline{F_1}$ is a set of pairs $< w_0, f >$ such that f verifies K_1 in M_{w_0} at t.

Let $< w_0, f >$ be any one such pair. Then the value that F_3 returns for this argument is the set of all embedding functions g which extend f and whose Domain consists of that of f together with the drefs in U_{K_3}:

$\mathrm{Dom}(g) = \{i, n, m, s_1, s_2, e_1\}$

The Semantics of IAADRSs

- Moreover, each such g must verify K_3.

  That is: $g(n) \prec_{T_M} g(e_1)$ and $< g(e_1), g(i), g(m) > \in I_M(\text{find})$.

- The second DES DRS K_4 depends referentially on the second BEL DRS K_2 in the same way that the first DES DRS K_3 depends on the first BEL DRS $K - 1$.

  So again we just record the result:

  For any pair $< w_0, f >$ from $\overline{F_2}$, $F_4(< w_0, f >)$ is the set of those embedding functions g that extend f and are such that $\text{Dom}(g) = \{i, n, d, s_3, s_4, e_2\}$, $g(n) \prec_{T_M} g(e_2)$ and $< g(e_2), g(i), g(d) > \in I_M(\text{find})$.

The Semantics of IAADRSs

- The last content DRS, K_5, depends on both K_1 and K_2.

This means that F_5 is a function the Domain of which is the merge $\cup\{\overline{F_1}, \overline{F_2}\}$

So let us compute this merge first.

Recall that the pairs belonging to the merge of a set of open propositions are of the form $< w, h >$, where the function h is some consistent union of embedding functions from each of the open propositions in the set., but all involving the same w

In the present case that boils down to this:

The pairs in the merge of the two-membered set $\{\overline{F_1}, \overline{F_2}\}$ are pairs of the form $< w_0, h >$ where h is a function that is the union of a function f_1 such that $< w_0, f_1 >$ is in $\overline{F_1}$ and a function f_2 such that $< w_0, f_2 >$ is in $\overline{F_2}$.

The Semantics of IAADRSs

- Furthermore, in the present case the overlap of the Domains of any such functions f_1 and f_2 consists just of the self-reflective drefs i and n.

  Since the values for these discourse referents are fixed in advance, f_1 and f_2 will agree on these arguments.

  This means that any function f_1 from a pair $< w_0, f_1 >$ from $\overline{F_1}$ and any function f_2 from a pair $< w_0, f_2 >$ from $\overline{F_2}$ will be compatible with each other, in the sense that $f_1 \cup f_2$ is a function.

  So each such combination will give you a function h such that $< w_0, h >$ belongs $\cup\{\overline{F_1}, \overline{F_2}\}$

  So in this case the set of h such that $< w_0, h >$ is in the merge $\cup\{\overline{F_1}, \overline{F_2}\}$ is (speaking somewhat sloppily) the cross product of the f_1 such that $< w_0, f_1 >$ is in $\overline{F_1}$ and the f_2 such that $< w_0, f_2 >$ is in $\overline{F_2}$.

The Semantics of IAADRSs

- F_5 is a function whose Domain is $\cup\{\overline{F_1}, \overline{F_2}\}$, i.e. the set of pairs $< w_0, h >$ just described.

For each such pair $< w_0, h >$ F_5 returns as value the set of all functions g such that (i) g extends h, (ii) the Domain of g is the set $\{i, n, m, s_1, s_2, d, s_3, s_4, e_3\}$ and (iii) g verifies K_5 in M_{w_0}.

That is, these functions g must have the additional property that $g(n) \prec_{T_M} g(e_3)$ and $< g(e_3), g(i), g(m), g(d) > \in I_{M_{w_0}}$(bring-together).

(N.B. The verb 'bring together' has been analyzed here as a 4-place predicate, which holds between an event e, an agent a, and two other individuals b and c iff e is an event of a bringing b and c together with her- or himself. That is not a particularly good analysis. But it doesn't to what the example is meant to illustrate.)

The Semantics of IAADRSs

- Is the description we have of Hamid's mental state the right one, or the most plausible one?

  An alternative description might be one in which Hamid's mental state contains anchored entity representations for his mother and father.

  Our ties to our parents, it might be argued, are transparent to such an extent that they enable us to have anchored representations for them, even if we have never consciously seen or met them.

  The internal anchors for such entity representations would be different from the perceptual anchors considered thus far. They should take the form of a special causal link to oneself.

- Let us use the same conditions that express the relationships of mother- and fatherhood in the beliefs of (12) for expressing these links.

  These assumptions lead to the following mental state description.

The Semantics of IAADRSs

(13)

$$
\left\{
\begin{array}{ll}
\left\langle [ANCH, m], \begin{array}{|c|}\hline m \quad s_1 \\ \hline n \subseteq s_1 \\ s_1: mother(m,i) \\ \hline \end{array} \right\rangle &
\left\langle [ANCH, d], \begin{array}{|c|}\hline d \quad s_3 \\ \hline n \subseteq s_3 \\ s_3: father(d,i) \\ \hline \end{array} \right\rangle \\
\left\langle BEL, \begin{array}{|c|}\hline s_2 \\ \hline n \subseteq s_2 \quad s_2: alive(m) \\ \hline \end{array} \right\rangle &
\left\langle BEL, \begin{array}{|c|}\hline s_4 \\ \hline n \subseteq s_4 \quad s_4: alive(d) \\ \hline \end{array} \right\rangle \\
\left\langle DES, \begin{array}{|c|}\hline e_1 \\ \hline n \prec e_1 \\ e_1: find(i,m) \\ \hline \end{array} \right\rangle &
\left\langle DES, \begin{array}{|c|}\hline e_2 \\ \hline n \prec e_2 \\ e_2: find(i,d) \\ \hline \end{array} \right\rangle \\
\multicolumn{2}{c}{\left\langle DES, \begin{array}{|c|}\hline e_3 \\ \hline n \prec e_3 \\ e_3: bring-together(m,d,i) \\ \hline \end{array} \right\rangle}
\end{array}
\right\}
$$

The Semantics of IAADRSs

- Let us assume that AERs for one's parents are always externally anchored.

  (If they weren't, one wouldn't exist.)

  If so, then in this last mental state description all drefs are either self-reflective or externally anchored.

  So all content representations of its beliefs represent singular propositions, which are about the values of the drefs they contain. And the referentially dependent content representations also denote singular semantic values, in a sense to be explained below.

  To see more clearly what these notions come to let us start with the content representation K_{BEL_1} of the first of the two belief components of (13).

The Semantics of IAADRSs

- The distinction between singular and non-singular propositions doesn't show up in 'quasi-extensional models like the model $\{< w_0, M_{w_0} >\}$ we used for our illustrative purposes above.

  So let us consider instead a non-degenerative intensional model M whose world set W_M consists of more than one world.

  The proposition expressed by K_{BEL_1} is singular with respect to the dref m (or,as we also say, with respect to the argument position filled by this dref), in the following sense.

  Let w and w' be two worlds from W_M. Then in both w and w' the content of the desire is about Hamid's actual mother.

  The reason is that all embedding functions f that can enter into the verification of K_{BEL_1} in either M_w or $M_{w'}$ will assign m the same individual from M, viz the external anchor for m's AER.

The Semantics of IAADRSs

- More precisely, suppose that (13) describes Hamid's state at t (were t is a time from the time structure of M).

  So for all relevant embedding functions f, $f(n) = t$.

  Then we have:

  K_{BEL_1} is true in M_w at t iff there exists an embedding function f such that $f(n) \subseteq_M f(s_3)$ and $< f(s_3), f(m) > \in I(M_w)(alive)$.

  and likewise:

  K_{BEL_1} is true in $M_{w'}$ at t iff there exists an embedding function f' such that $f'(n) \subseteq_M f'(s_4)$ and $< f(s_3), f(m) > \in I(M_{w'})(alive)$.

The Semantics of IAADRSs

- But for all embedding functions f, f', irrespective of whether they are embedding functions in M_w, $M_{w'}$ or $M_{w''}$ for any other w'' of W_M, we have that $f(i) = f'(i)$ and $f(m) = f'(m)$.

  So the set of worlds from W_M in which K_{DES_1} is true at t consists of those worlds w in which Hamid's 'real mother' is alive.

- This then is what singularity comes to in the present framework:

  A DRS K *expresses a proposition that is singular with respect to* a dref x occurring in a model M iff all embedding functions that can enter into the evaluations of K at different extensional models M_w of M map x to the same individual of M.

The Semantics of IAADRSs

- How does the proposition expressed by the DRS K_{BEL_1} of (13) differ from that expressed by the corresponding DRS K'_{BEL_1} of the first belief in our first proposal (12) for the description of Hamid's mental state ?

That depends on what we say about the motherhood relation.

If we assume, in the spirit of Kripke, that parenthood relations are necessary relations, then the proposition expressed in M by K'_{BEL_1} will be the same set of possible worlds as the one expressed by K_{BEL_1}.

Does that make the proposition expressed by K'_{BEL_1} a singular proposition?

I do not know of any discussion in the literature that helps us decide this question.

(Though there may very be such discussions somewhere. If anyone knows, then please tell me.)

The Semantics of IAADRSs

- According to this proposal the proposition expressed by K_{BEL_1} in M is singular with respect to m, but the proposition expressed by K'_{BEL_1} in M is not.

  This is because all the embedding functions that might be involved in the evaluation of K_{BEL_1} map m to the same individual (the external anchor of the first AER in (13)).

  But the embedding functions that might be involved in the evaluation of K'_{BEL_1} can in principle map m to different individuals.

  It is just that because the motherhood relation, which is expressed by the condition 's_1: $mother(m, i)$ in K_{BEL_1}, is assumed to be a necessary relation that any function f which verifies K'_{BEL_1} in a model M_w will have to map m to the real mother of Hamid.

The Semantics of IAADRSs

- On this view two DRSs can express the same proposition (in the sense that they are true in the same set of possible world)

  and yet the one can express a singular proposition while the other does not.

  On this conception of singularity, then, it isn't just the identity of a proposition, as a set of possible worlds, that determines whether or not it is singular, but also the way in which it is expressed.

  Sometimes the same proposition, qua set of possible worlds, may be expressed as a singular proposition but also as a non-singular one.

The Semantics of IAADRSs

- In the above the notion of singular expression has been given for DRSs that occur as components of IAADRSs.

  Transfer of this notion to sentences and discourses of English is not automatic.

  But there is a straightforward connection if we make the standard assumption of DRT about DRS construction: DPs of an English sentence introduce drefs in the DRS constructed for it.

  A further standard feature of DRS construction is that the DPs whose drefs get fixed values (as self-reflective or as externally anchored drefs) are those that are the ones that are introduced for DPs that treated as directly referential in accounts like those of Kaplan and others.

  So the English sentences/discourses that express singular propositions are precisely those that contain directly referential DPs in the sense of these theories.

The Semantics of IAADRSs

- The distinction between singular and non-singular semantic values can also be applied to components of mental state descriptions that are referentially dependent on other components.

  Consider for instance the content DRS K_{DES_3} of the third desire component of (13).

  Each of its drefs except for e_3 has a fixed value, in the sense that all embedding functions referred to in the truth evaluation definition will assign it the same value.

  The DRS K_{DES_3} can therefore also be said to express a singular semantic value in M at t, or more precisely a value that is singular with respect to n, i, m and d.

  This is perfectly well-defined even though the semantic value of K_{DES_3} is a functional information state and not a proposition.

MSDRT

- We have now formally specified the syntax and the model-theoretic semantics of mental state descriptions, given a set of Mode Indicators and a basic DRS language L_0. These are the IAADRSs we defined above.

  More specifically, let us assume for now that our repertoire of Mode Indicators consists of BEL, DES, INT and [Anch,α], where α can be any discourse referent of the type that stands for an individual or a group of individuals.

  The definition of IAADRSs is one step towards MSDRT. But it is only the first step.

  What we want is a formalism in which we can represent what it is for one agent to attribute a mental state to some other agent (or to herself).

MSDRT

- To this end we introduce the predicate 'Att' into our DRS language.

  An example of a DRS containing an instance of Att is shown on the next slide.

- Att is a 4-place predicate.

  Its first argument slot is filled by a state dref.

  Intuitively the state represented by this dref is that the atrributee, represented by the second argument of Att, is a in mental state of the kind described by Att's third argument.

  The third argument slot of Att is always filled by an IAADRS.

MSDRT

(14).

$$
\begin{array}{|c|}
\hline
s \quad a \quad x' \\
\hline
n \subseteq s \\
s : Att\left(a, \left\{ \begin{array}{l}
\left\langle [ANCH, x]\,, \begin{array}{|c|} \hline x \quad s_1 \\ \hline n \subseteq s_1 \\ s_1 : see'(i, x) \\ \hline \end{array} \right\rangle \\[2ex]
\left\langle BEL, \begin{array}{|c|} \hline s_3 \\ \hline n \subseteq s_3 \\ s_3 : gold - coin'(x) \\ \hline \end{array} \right\rangle \\[1ex]
\left\langle DES, \begin{array}{|c|} \hline s_4 \\ \hline n \subseteq s_4 \\ s_4 : have(i, x) \\ \hline \end{array} \right\rangle \\[1ex]
\left\langle INT, \begin{array}{|c|} \hline t \quad e \\ \hline n < t \;\; e \subseteq t \\ e : pick - up'(i, x) \\ \hline \end{array} \right\rangle
\end{array} \right\}, \{< x, x' >\}\right) \\
\hline
\end{array}
$$

MSDRT

- The fourth argument of Att is filled by a *linking relation*, a relation between AERs occurring in the third argument slot and entities that are 'identified from the outside' –

  identified not by the attributee's internal psychology, but by the attributor.

  In fact, there doesn't have to be identification in the usual sense.

  All that external linking requires is that the attributor assumes that there is something to which the AER in question is externally anchored.

  She herself may know no more about this entity.

  But there are of course also many cases in which the attributor has her own beliefs about this entity.

  And quite often she will have identifying knowledge about the entity herself.

MSDRT

- Formally, the terms occupying the fourth argument slot of Att are lists of ordered pairs $<x, x'>$, where x is the distinguished dref of an AER that is a component of the term occupying the third slot and x' is a dref that 'must be bound in the DRS K'.

  That is, x' should occur in the Universe of K.

  (x' may also be bound by belonging to the Universe of a DRS K' of which K is a sub-DRS of K'.)

MSDRT

- Note that the DRS (14) could be the content representation of some component of the mental state description to some other agent b.

  For instance, (14) could be the content representation of a belief of b's that is a component of this description.

  In that case the belief of b that is described by this component would be to the effect that some agent a is having a (veridical) perception of something that she takes herself to be seeing in the middle of the road and of which she thinks that it is gold coin, wants to have it and intends to pick it up.

  In other words, b only believes that there is something that a is seeing, but may have no further information about that thing.

MSDRT

- Perhaps it wouldn't be very common to entertain such a doubly existential belief (existential both with respect to e and to what a is seeing).

  A somewhat more plausible alternative would be one in which b has an anchored representation for a.

  In that case the relevant part of b's attitudinal state would be as in (15).

MSDRT

(15).

$$\left\{ \begin{array}{l} \left\langle [ANCH, a], \begin{array}{|c|} \hline a \\ \hline K_a \\ \hline \end{array} \right\rangle \\ \\ \left\langle BEL, \begin{array}{|c|} \hline s \quad x' \\ \hline n \subseteq s \\ s : Att(a, \left\{ \begin{array}{l} \langle [ANCH, x]\ ,\ K_{ANCH}\ \rangle \\ \\ \langle BEL,\ K_{BEL}\ \rangle \\ \langle DES,\ K_{DES}\ \rangle \\ \\ \langle INT,\ K_{INT}\ \rangle \\ \end{array} \right\} , \{< x, x' >\}) \\ \\ \hline \end{array} \right\rangle \end{array} \right\}$$

MSDRT

- Of course the combination of AER and belief component in (15) will specify a well-defined belief content only if b's AER for a has an external anchor.

  As before, such an external anchor for (15) could only be specified 'from the outside', i.e. by someone who attributes to b an attitudinal state containing the components shown in (15).

  Yet another possibility would be that in which b not only has an AER for a but also for the entity that he believes a is seeing; the diagram is shown on the next slide.

  Of course the external anchor for this second AER could also only specified from the outside.

MSDRT

(16).

$$\left\{\begin{array}{l}
\langle [ANCH, a], K_a \rangle \\[2ex]
\langle [ANCH, x'], K_{x'} \rangle \\[2ex]
\left\langle BEL, \begin{array}{|c|}\hline
s \;\; x' \\ \hline
n \subseteq s \\
s : Att(a, \left\{\begin{array}{l}
\langle [ANCH, x]\ ,\ K_{ANCH} \rangle \\[2ex]
\langle BEL,\ K_{BEL} \rangle \\
\langle DES,\ K_{DES} \rangle \\[2ex]
\langle INT,\ K_{INT} \rangle
\end{array}\right\} , \{< x, x' >\}) \\ \hline
\end{array}\right\rangle
\end{array}\right\}$$

MSDRT

- Before we go on two useful abbreviations.

1. Suppose

(i) that a DRS contains a condition of the form 's: $Att(a, \mathbb{K}, LINK)$',

(ii) that $\mathbb{K}$ is an IAADRS consisting solely of a single attitude component $< MOD, K >$ and a set of AERs whose distinguished drefs occur in K.

Then we will sometimes use the following abbreviation for this Att condition:

s: $Mod(a, K, LINK)$

MSDRT

- For instance, suppose that the only component of $\mathbb{K}$ that is not an AER is for the form $< BEL, \begin{array}{|c|}\hline \\ \hline \text{great(i)} \\ \hline \end{array} >$ and where LINK is the empty set.

Then the Att condition can be abbreviated as:

$$s\text{: } Bel(a, \begin{array}{|c|}\hline \\ \hline \text{great}(i) \\ \hline \end{array}, \emptyset)$$

or, even more simply, omitting $\emptyset$, as:

$$s\text{: } Bel(a, \begin{array}{|c|}\hline \\ \hline \text{great(i)} \\ \hline \end{array})$$

MSDRT; iterated attributions

- The predicate ‘Att’ is a powerful device.

In particular, it allows the representation of attitude nestings, as expressed in the following two sentences.

(17)

a. Bill thinks that John believes that Mary is sad.

b. Mary believes that John thinks that she is sad.

- Within our framework attitude attributing sentences tend to be multiply ambiguous:

They can be mapped onto distinct representations that our formalism makes available.

We will address this problem in detail later.

For now we will only look at some logical forms for these sentences that our formalism makes available.

MSDRT; iterated attributions

- For now we will only look, for each of these two sentences, at only one of the logical forms that our formalism makes available for them.

- In the DRS shown below as logical form for (17.a) the 'outer' name *Bill* is treated as in original DRT:

  via a discourse referent b and a condition '$Bill(b)$' which is intended to capture that b represents the person that (17.a) is used to talk about.

  The only difference is that we now, with an eye on the account of names that is coming, use the condition '$Named(b, Bill)$' in lieu of '$Bill(b)$'.

  This treatment of names is notoriously unsatisfactory. A more rfined account of roper names will be discussed later on.

MSDRT; iterated attributions

- The other two names in (17.a), *John* and *Mary*, are treated along lines that are closes to the ones we will adopt when looking at proper names more closely.

  The DRS below assumes that Bill has AERs for both John and Mary, that John has an AER for Mary, that this assumption is part of the mental state that Bill's belief attributes to John.

  (According to the DRS below Bill also assumes that John's AER is for the same Mary as Bill's own.

  This last assumption would probably have been taken for granted even if I hadn't stated it.

  But as part of our use of MSDRT as Logical Form Formalism for sentences like these it too is something that has to be stated explicitly.)

MSDRT

(18)

$$\begin{array}{|c|}\hline b\ \ j\ \ m\ \ s \\ \hline Named(b,\text{`}Bill') \quad Named(j,\text{`}John') \quad Named(m,\text{`}Mary') \\ n \subseteq s \\ s\!: Att\left(b, \left\{ \begin{array}{l} \langle [ANCH, j_b]\ ,\ \emptyset\ \rangle\ \langle [ANCH, m_b]\ ,\ \emptyset\ \rangle \\ \left\langle BEL, \begin{array}{|c|}\hline \\ \hline Named(j_b,\text{`}John') \\ \hline \end{array} \right\rangle \\ \left\langle BEL, \begin{array}{|c|}\hline \\ \hline Named(m_b,\text{`}Mary') \\ \hline \end{array} \right\rangle \\ \langle BEL,\ K_{BEL_b}\ \rangle \end{array} \right\}, \{<j_b, j>, <m_b, m>\}\right) \\ \hline \end{array}$$

MSDRT

(19) (K_{BEL_b})

$$
\begin{array}{|c|}
\hline
s' \\
\hline
n \subseteq s' \\
\\
s' \colon Att\left(j_b, \left\{ \begin{array}{l} \langle [ANCH, m_j],\ \emptyset\ \rangle \\ \\ \left\langle BEL, \begin{array}{|c|} \hline \\ \hline sad(m_j) \\ \hline \end{array} \right\rangle \end{array} \right\}, \{<m_j, m_b>\}\right) \\
\\
\hline
\end{array}
$$

MSDRT; iterated attributions

- (17.b) is most naturally understood as a self-attribution by Mary:

that she herself is the subject of John's belief.

- That is, she herself – represented as i – is the external anchor for an Anchored representation for her that is part of what she attributes to John.

A DRS capturing this is given on the next slide.

MSDRT

(20)

$$
\begin{array}{|c|}
\hline
m \quad j \quad s \\
\hline
Named(m, \text{`}Mary') \quad Named(j, \text{`}John') \\
n \subseteq s \\
s\text{: } Att(m, \left\{ \begin{array}{l}
\langle [ANCH, j_m] \,,\; \emptyset \,\rangle \\
\\
\left\langle BEL, \begin{array}{|c|} \hline \\ \hline Named(j_m, \text{`}John') \\ \hline \end{array} \right\rangle \\
\\
\langle BEL,\; K_{BEL_m} \,\rangle
\end{array} \right\}, \{<j_m, j>\}) \\
\hline
\end{array}
$$

(K_{BEL_b} on next slide)

MSDRT

(21) (K_{BEL_b})

$$
\begin{array}{|c|}
\hline
s' \\
\hline
n \subseteq s' \\
\\
s'\text{: } Att\left(j_m, \left\{ \begin{array}{l} \langle [ANCH, m_j]\,,\ \emptyset\ \rangle \\ \\ \left\langle BEL, \begin{array}{|c|} \hline \\ \hline sad(m_j) \\ \hline \end{array} \right\rangle \end{array} \right\}, \{<m_j, i>\}\right) \\
\\
\hline
\end{array}
$$

MSDRT: Model-theoretic Semantics

- We now turn to the model-theoretic semantics for MSDRT.

The principal task here is to provide verification conditions for DRS Conditions whose main predicate is 'Att').

We have given a model-theoretic semantics for the IAADRSs that can occur as fillers of the third slot of Att.

But what should be the relation between that semantics and the verification conditions for Att-Conditions?

One of our basic assumptions is that attitude attributions have truth conditions and that these truth conditions should be made explicit in model-theoretic terms.

But what information must models contain, against which the truth or falsity of such an attribution – the verifiability or non-verifiability of Att-Conditions – can be assessed?

MSDRT: Model-theoretic Semantics

- I do not think there is unique answer to this question and what follows is only one of several possible answers.

For starters something that I take to be uncontroversial and that any reasonable answer will have to accept.

Our models will now have to provide information about the mental states of agents – for how else could Att-Conditions be evaluated for correctness?

That is, we will now enrich our intensional models M with a further component 'AS' (for 'Attitudinal state').

AS is a partial function that assigns to triples $<a, w, t>$ of agents,worlds and times the attitudinal state that the agent is in in that world at that time.

MSDRT: Model-theoretic Semantics

- But what are the values of AS? This is the crucial question.

This is the point where more than one answer can be given.

One of these we present now.

(Alternatives will be discussed after the satisfaction conditions for Att-Conditions has been spelled out for this first answer.)

AS(a, w, t) is a set of components $<MOD, Z>$ which as a whole has the structure of a coherent IAADRS.

However, we want the second members Z of the components $<MOD, Z>$ of AS(a, w, t) to be independent of any conventions about representational form.

MSDRT: Model-theoretic Semantics

- To be more precise, these components should not only be independent from any particular description in a natural language like English, but also from the conventions of the formal language we use for the description of mental states.

  In the present context this means that the components should be independent from the particular MSDRT language we are using.

- The following proposal is a compromise between two desiderata:

  (i) To make the components Z as independent from the syntax of the MSDRT language we have adopted for the description of mental states;

  (ii) To preserve enough of the structure of semantic values of our mental state descriptions to allow for a substantive definition of the satisfaction conditions for Att-Conditions.

MSDRT; Model-theoretic Semantics

- Here is my proposal:

The values of AS have the same general form as the values that our mental state descriptions determine in models at times.

We will refer to such values as 'ISBAS's'. 'ISBAS' is short for 'Information State Based Attitudinal State'.

<u>Definition</u>: Let M be an intensional model. An *ISBAS* $\mathbb{I}$ *relative to M* is a set of pairs $<MOD, Z>$, where

(i) MOD is a Mode Indicator for a propositional attitude, and

(ii) Z is a partial information state.

(iii) the Referential Dependency relation between the members of $\mathbb{I}$ is well-founded and obeys the constraints of the attitudinal hierarchy.

MSDRT: Model-theoretic Semantics

- As it stands, this definition is incomplete because we haven't yet made explicit what the Referential Dependency relation is for an arbitrary ISBAS $\mathbb{I}$.

We define the Referential Dependency Relation for an ISBAS $\mathbb{I}$, $\prec_{RDR}(\mathbb{I})$, in terms of the Bases of the content components Z of pairs $<MOD, Z>$ in $\mathbb{I}$ and the Bases of the presuppositions of those components:

Definition: Suppose that $<MOD, Z>$ and $<MOD', Z'>$ are members of an ISBAS $\mathbb{I}$.

The *Referential Dependency Relation for* $\mathbb{I}$, $\prec_{RDR}(\mathbb{I})$, holds between $<MOD, Z>$ and $<MOD', Z'>$ iff there is a discourse referent x which belongs to the Base of Z' and the Base of the Presupposition of Z.

MSDRT: Model-theoretic Semantics

- To see the motivation behind this definition recall our first example (on slide 33) of an ADRS.

The DRSs K_{DES} and K_{INT} which specify the contents of the Desire and the Intention components of that ADRS are referential dependent on the specification K_{BEL} of the Belief component because DRSs K_{DES} and K_{INT} contain free occurrences of the dref x, which is 'declared' in K_{BEL} (it occurs in the Universe of K_{BEL}).

At the more abstract level of an ISBAS this combination – of a dref occurring free in one component and bound in another – is captured via the Bases of those components:

A dref x has a 'free occurrence' in a partial information state Z, where $<MOD, Z>$ is a component of a given ISBAS $\mathbb{I}$ iff x is part of the presupposition of Z, as a member of the Base of this presupposition.

And x is 'bound' in Z iff it belongs to the Base of Z itself.

MSDRT: Model-theoretic Semantics

- Note well in this connection that the self-reflective drefs n and i are not among the discourse referents that can occur be part of an ISBAS.

  All drefs that occur in ISBASs are 'regular discourse referents', which are subject to the general principles of interpretation that also govern the semantics of the DRSs of our underlying DRS language in extensional models.

  We will address the possible role of n and i in the values of the function AS later, after our account of the satisfaction of Att-Conditions has been completed.

MSDRT: Model-theoretic Semantics

- Another point worth noting is that ISBASs are just collections of propositional components.

  They do not contain abstract counterparts of Anchored Entity Representations

  Nevertheless they will be used in the evaluation of Att-Conditions in which the third slot of Att is filled by any IAADRS (including those that contain AERs as well propositional attitude components).

  What makes this possible is that the contents specified by an ISBAS may be singular contents (relative to the given intensional model M).

MSDRT: Model-theoretic Semantics

- However, as will be clear from the definitions that follow, the verification of an IAADRS containing AERs by an ISBAS is sensitive only to the external anchors of those AERS;

  Their internal anchors are ignored.

  The content of the internal anchors will be captured, however, by the internalist reduction of the IAADRS (see slide 36).

  It will also be captured by the IAADRS itself if the belief components into which the internal reduction converts internal anchors are added to the IAADRS (.i.e. without eliminating its AERs).

MSDRT: Model-theoretic Semantics

- The attitudinal states which the function AS assigns to agents and times can be expected to be very rich.

  They will typical contain large numbers of both propositional attitude components and AERs.

  Att-Conditions, in contrast, typically only describe some small part of the mental states of the attributee.

  So, in order that an Att-Condition be verified by a state AS(a, w, t) it will only be required that the components of the Att-Condition correspond to some small subset of AS(a, w, t).

MSDRT: Model-theoretic Semantics

- Formally that comes to this:

There must be a mapping F from the third argument of the Att-Condition (viz. its IAADRS) into AS(a, w, t) which

(i) F preserves Mode Indicators: if $F(<MOD, K>) = <MOD', Z>$, then $MOD' = MOD$

(ii) F is such that K is 'entailed' by Z.

In addition we require that F preserves the referential dependence between the components of the IAADRS from the Att-Condition.

MSDRT: Model-theoretic Semantics

- The relevant notion of entailment needs some spelling out for those cases where $<MOD, K>$ and $F(<MOD, K>)$ $(= <MOD, Z>)$ have predecessors in their respective referential dependency hierarchies.

  In essence the relation holds if the open proposition determined by $[[[K]]]_{M,\mathbb{K},\mathbf{a},t}$ is entailed by the open proposition determined by Z.

  Note that this entailment requirement introduces a further dimension of partiality into the verification conditions of Att-Conditions:

  Suppose for instance that $<BEL, K>$ is part of the third argument of an occurrence of Att in the Logical Form of some attitude attribution (to some agent a in some world w at some time t).

  The correctness of the attribution only requires that the belief component of a's mental state at t *entail* the semantic value of K, not that the two are the same.

MSDRT: Model-theoretic Semantics

- In other words, the content specification in the attribution needs to give only a partial account of this content.

  Note well, however, that the entailment relation between actual attitude and attitude attribution holds only for ‘upward monotonic’ attitudinal modes, such as belief, desire and intention.

  It does not hold for downward monotonic modes like doubt or reject, or ‘monotonicity-neutral’ attitudinal modes like *wonder*.

  (We will return to this point later on.)

MSDRT: Model-theoretic Semantics

- There is one point about the verification of Att-Conditions that we have so far overlooked.

  This concerns the discourse referents that are part of the content-specifying DRSs that occur as part of the IAADRSs in Att's third argument slot.

  The choice of the drefs in these DRSs is, as always, arbitrary. (All that matters is that the drefs are distinct.)

  There also as a certain arbitrariness in the choice that AS values make of the sets of discourse referents that form the Bases of their components.

  In fact, speaking of 'discourse referents' in relation to AS values is something that deserves comment in any case.

MSDRT: Model-theoretic Semantics

- Consider once more the example of agent a who walks along the road and thinks she is seeing a gold coin in the middle of the tarmac.

  The description of her state that we assumed when first discussing this case and that led us eventually to our formulation of MSDRT makes use of discourse referents.

  But these discourse referents have a different status from those that get introduced when MSDRSs are constructed as Logical Forms of attitude attributions made in English (or any other natural language).

  The former have an identity that is internal to the psychology of the agent a.

  This identity has nothing to do with whether any attitude attributions are made to a or with the Logical Forms that can be constructed for those attributions.

MSDRT: Model-theoretic Semantics

- This means that verification of the Logical Forms of attitude attributions must allow for ‘renaming of discourse referents’:

  The discourse referents occurring in the Logical Form must be correlated (one-to-one) with the discourse referents in the relevant value of AS.

  But these discourse referents have a different status from those that get introduced when MSDRSs are constructed as Logical Forms of attitude attributions made in English (or any other natural language).

  In other words, verification of any DRS containing Att-Conditions must include such renamings.

MSDRT: Model-theoretic Semantics

- Since we cannot expect the drefs in the content specifications K and those occurring $\mathrm{AS}(a, w, t)$ to match, we must allow renaming as part of the verification of Att-Conditions.

  So we assume that the verification of the IAADRS $\mathbb{K}$ that is filling the third slot of an Att-Condition involves a one-to-one mapping $\mathrm{REN}(\mathbb{K},\mathrm{AS}(a, w, t))$ from the discourse referents occurring in $\mathbb{K}$ to the discourse referents occurring in the relevant AS value $\mathrm{AS}(a, w, t)$.

  $\mathrm{REN}(\mathbb{K},\mathrm{AS}(a, w, t))$ can be used to transform $\mathbb{K}$ into an alphabetic variant $\mathbb{K}'$ in which each dref x occurring in $\mathbb{K}$ is replaced everywhere in $\mathbb{K}$ by $\mathrm{REN}(\mathbb{K},\mathrm{AS}(a, w, t))(x)$.

MSDRT: Model-theoretic Semantics

- It is then $\mathbb{K}'$ that gets embedded into AS(a, w, t) via the function F.

  By proceeding in this way we make sure that the referential dependency relations in $\mathbb{K}'$ can be preserved by F:

  The pair <REN($\mathbb{K}$,AS(a, w, t)),F> must be such that F preserves the referential dependency structure of $\mathbb{K}'$.

- The renaming procedure just described works unproblematically so long as $\mathbb{K}$ contains no $\mathbb{K}$-*free* occurrences of drefs –

  that is: no dref occurrences that are not bound somewhere inside $\mathbb{K}$.

  The intuitive meaning of this is:

  A thought can be directly about a thing only is that thing is represented by an Anchored Entity Representation.

MSDRT: Model-theoretic Semantics

- This is an assumption that so far we didn't impose explicitly.

But it is in the spirit of our general conception of direct reference and its formalization.

- When a DRS contains two or more Att Conditions, renaming the drefs occurring in IAADRSs only works properly if we impose a further restriction:

The sets of discourse referents occurring in the IAADRSs occupying the third argument slots of those Att Conditions must be disjoint.

- This constraint is always satisfied in DRSs that have been constructed from natural language input.

We now adopt this constraint as a general well-formedness condition on the DRSs of MSDRT.

MSDRT: Model-theoretic Semantics

- This gives us almost all the pieces for the verification conditions of Att-Conditions.

  There are a couple of issues that still need to be addressed.

  The first of these has to do with the handling of time:

  The function AS assigns attitudinal states to triples of agents, worlds and times.

  But what are the times in this case?

  In principle these could be instants of the time structure of the model or intervals.

  However, taking the times that figure as third arguments of AS to be intervals leads to awkward questions about consistency in those cases where those intervals overlap.

MSDRT: Model-theoretic Semantics

- We therefore assume that the times are instants.

This is an intuitively plausible decision:

At each instant of time t when the agent a is conscious in w, a is in a mental state $\mathrm{AS}(a, w, t)$.

The state will typically have some temporal inertia:

If it is the state that a is in at t, then a is likely to also be in that state at neighboring times t'.

But that will be reflected by AS if it assigns ISBASs to agents, worlds and temporal instants.

MSDRT: Model-theoretic Semantics

- At this point we should recall that the durations of the states that occur as first arguments of Att-Conditions are temporal intervals.

  Suppose for instance that an embedding function f that arises in the verification of a DRS containing the Att-Condition '$s : ATT(a, \mathbb{K}, EXTANCH)$' assigns the state $f(s)$ of the extensional model M_w to the state dref s of this Condition.

  Let the duration of $f(s)$ be the interval t_s of the time structure of M.

- Question: What is it for f to verify '$s : ATT(a, \mathbb{K}, EXTANCH)$' in M_w?

MSDRT: Model-theoretic Semantics

- One part of the answer should be intuitively clear:

  f verifies the Att-Condition 's :$ATT(a, \mathbb{K}, EXTANCH)$' in M_w

  (if and) only if there exists a combination $<$REN($\mathbb{K}$,AS(a, w, t)),$F>$ such that for each instant t belonging to the interval t_s,

  the components of the IAADRS occupying the third slot of the Att-Condition stand in the required entailment relations to the corresponding components of the ISBAS AS(a, w, t).

MSDRT: Model-theoretic Semantics

- This way of defining the satisfaction conditions of Att-Conditions is based on a presupposition that must be made explicit.

  The function F must give a correspondence that is *uniform over the different ISBASs $AS(a, w, t)$, for t within t_s (the interval $dur(f(s))$).*

  That is, for any component $<MOD, K>$ of $mathbbK$ F($<MOD, K>$) must belong to each of those ISBASs.

  This makes sense only if the component of mental states have a certain individual identity over time.

MSDRT: Model-theoretic Semantics

- Intuitively this is plausible enough:

As a rule our attitudes – our particular beliefs, desires etc – persist over time, at least for some time.

They may remain the same while our mental states over-all are changing, because new beliefs or desires become part of them and other attitudes are abandoned or modified.

(There is much that can and ought to be said about the dynamics of mental states, and the way in which IAADRSs and ISBASs make the structure of mental a states explicit can help with this.)

But this is an aspect of them that our definition of AS as a function of triples $<a, w, t>$ does not do justice to.

MSDRT: Model-theoretic Semantics

- Rather than adding a formalization of the persistence of individual attitudes to the complexity of the satisfaction definition for Att-Conditions, we will therefore be content with a conceptual stop-gap:

A correspondence function F that is part of the verification of an Att-Condition '$s : ATT(a, \mathbb{K}, EXTANCH)$' over a temporal interval t_s must assign to each propositional component $<MOD, K>$ of $\mathbb{K}$ a value that belongs to each of the ISBASs AS(a, w, t) with t in t_s.

It is implicitly understood that these values are the same attitudinal components in the sense alluded to above.

They are not components that are privy to the individual ISBASs AS(a, w, t), but that happen to be indistinguishable from the components of ISBASs AS(a, w, t') for t' *neq* t that are privy to their ISBASs and therefore must be distinct entities.

MSDRT: Model-theoretic Semantics

- There is also another aspect of the temporal dimension of the verification of Att-Conditions.

As a rule the content specifications K of components of $\mathbb{K}$ contain occurrences of the self-reflective dref n.

When the Att-Condition '$s : ATT(a, \mathbb{K}, EXTANCH)$' is evaluated in M_w by an embedding function f, what time of M should be assigned to these occurrences of n?

Intuitively the answer to this question is more or less clear:

f assigns some state $f(s)$ of M to the state dref s of the Att-Condition. The time $dur(f(s))$ indicates when the agent $f(a)$ is in the mental state described by the Att-Condition.

MSDRT: Model-theoretic Semantics

- The occurrences of n in the content specifications that are part of this description represent the 'psychological now' of the agent at that time.

So the time t that must be assigned to the occurrences of n must stand in some close relation to the time $dur(f(s))$.

The assumption we make is that the two times coincide:

The state $f(s)$ that is described by the Condition '$s : ATT(a, \mathbb{K}, EXTANCH)$' is to be thought of as determined, through its duration what the agent $f(a)$ takes to be her psychological now according to the evaluation of the Condition under f.

MSDRT: Model-theoretic Semantics

- Formally this means that the value under f of each content specification K of a component of $\mathbb{K}$ ought to be the partial information state $[[[K]]]_{M,\mathbb{K},\mathbf{a},dur(f(s))}$,

  as defined earlier when we dealt with the semantics of IAADRSs.

  But this suggestion leads us to yet another question:

  What could the subscript **a** stand for in our present use of this definition?

  Recall that when we defined $[[[K]]]_{M,\mathbb{K},\mathbf{a},dur(f(s))}$ earlier, **a** was a set of external anchors for the AERs of $\mathbb{K}$.

MSDRT: Model-theoretic Semantics

- But in an Att-Condition the external anchors for AERs in the third slot of Att are provided by the links EXTANCH that fills its fourth slot.

  EXTANCH is a set of pairs $<x, x'>$, where x is the distinguished discourse referent of an AER in the third slot and x' represents the external anchor of that AER.

  The embedding functions f under which an Att-Condition is evaluated will provide values for each such dref x'.

  This value is the external anchor that f provides for x (and thus for the AER of which x is the distinguished discourse referent).

  So the **a** of $[[[K]]]_{M,\mathbb{K},\mathbf{a},dur(f(s))}$ is in this case the function which maps each x occurring as first member of a pair in EXTANCH to $f(x')$, where x' is the second member of that pair.

MSDRT: Model-theoretic Semantics

- This completes our first pass at a verification definition for Att-Conditions and the DRSs containing them.

  Putting all the pieces of this definition together in a single statement leads to something rather monstrous.

  Instead we list once more the different ingredients of the definition.

- Simple and complex attitude attributions are represented by Att-Conditions.

- Verification of an Att-Condition s: $ATT(a, \mathbb{K}, EXTANCH)$ by an embedding function f in an extensional model M_w that is a component of an intensional model M makes use of what the model M_w has to say about the mental states of $f(a)$ at times included within the duration of $f(s)$.

MSDRT: Model-theoretic Semantics

- For all relevant combinations of a, w, t the specification $\mathrm{AS}(a, w, t)$ that M_w provides of a's mental state in w at t is an ISBAS (an 'information state-based attitudinal state').

  (An ISBAS is a complex of partial information states with a structure similar to that of the semantic values of IAADRSs.)

- For an embedding function f to verify $s :ATT(a, \mathbb{K}, EXTANCH)$ in M_w it is (sufficient and) necessary that at each time t within the duration of $f(s)$ the semantic values $[[[K]]]_{M,\mathbb{K},\mathbf{a},dur(f(s))}$ of the context specifications of the propositional components of $\mathbb{K}$ stand in the right semantic relations to the contents of corresponding components of $\mathrm{AS}(f(a), w, t)$.

MSDRT: Model-theoretic Semantics

- For the three propositional Mode Indicators BEL, DES, and INT these relations are entailment relations:

  The open proposition determined by $[[[K]]]_{M,\mathbb{K},\mathbf{a},dur(f(s))}$ must be entailed by the open proposition determined by the content of the corresponding component of $\mathrm{AS}(f(a), w, t)$.

- The correspondence between components of $\mathbb{K}$ and components of the ISBASs $\mathrm{AS}(f(a), w, t)$ is established in two stages.

  The first stage involves a renaming of the discourse referents in $\mathbb{K}$, so that they match the corresponding drefs in these ISBASes.

  Secondly, the propositional components of the result $\mathbb{K}'$ of this renaming are then mapped onto components of these ISBASs.

MSDRT: Model-theoretic Semantics

- The semantic relation between the IAADRS $\mathbb{K}$ and the relevant ISBASes AS$(f(a), w, t)$ is mediated by a renaming of the discourse referents in $\mathbb{K}$, so that they match the corresponding drefs in these ISBASes.

- After renaming, the entailment relations required by verification of the Att-Condition must hold between the semantic values of the content specifications K' of the components $<MOD, K'>$ of the result $\mathbb{K}'$ of renaming $\mathbb{K}$ and the corresponding components $F(<MOD, K'>)$ of the relevant ISBAS:

The open proposition determined by K' in M_w at $f(s)$ as part of $\mathbb{K}'$ must be entailed by the open proposition determined by the partial information state of $F(<MOD, K'>)$.

[End of Summary of the verification conditions for Att-Conditions]

MSDRT: Model-theoretic Semantics

- An aspect of the structure of attitudinal states that the definition above has ignored has to do with the self and its representation by means of the self-reflective discourse referent i.

In fact, we didn't so far address the question what occurrences of i in an IAADRS $\mathbb{K} that occupies the third slot of an Att-Condition$s:

$$ATT(a, \mathbb{K}, EXTANCH)$$

should be mapped onto by an embedding functions f.

Intuitively the answer is clear: these occurrences of i should be mapped onto the same individual as the agent dref a.

So we stipulate: determining the semantic value of the third argument $\mathbb{K}$ of an Att-Condition $s : ATT(a, \mathbb{K}, EXTANCH)$ by an embedding function f involves extending f by putting $f(i) = f(a)$.

MSDRT: Model-theoretic Semantics

- The stipulation $f(i) = f(a)$ has an effect, however, that may be considered undesirable:

It obliterates the difference between self-attribution and *de re* attributions to oneself that are not self-attributions.

(Those where one has an Anchored Entity Representation of oneself without realizing that one is the external anchor of that AER:

that it is your pants that are on fire, that it is the bag of sugar in your shopping cart that is producing the trail, that it is your voice that you are listening to on the recording of the singing contest, that it is your paw whose the imprints in the snow are puzzling and increasingly frightening you.)

MSDRT: Model-theoretic Semantics

- Given the way we have defined the values of the function AS there doesn't seem to be a way of doing better.

If we want to capture the difference between self-attributions and *de re* attributions to oneself that are not self-attributions at the level of truth conditions, then this distinction must be encoded in some form in the AS values.

As far as I can see, this would require a radical change to our set-up:

Presumably the values of AS would have to be assumed to be entities that are substantially different from the ISBASs we have been using.

But perhaps we should not be too worried that not all distinctions that MSDRT descriptions can make are captured by our model-theoretic semantics. Some tension between syntax and semantics seems inevitable in semantics and may be thought as part of its essence.

MSDRT: Model-theoretic Semantics

- 'now' attributions raise a similar issue as self-attributions.

I can have the '*de nunc*' thought that right now Mary is boarding a plane for Munich.

I can also have the thought that March 14, at 15.00 sharp, is Mary's boarding time, but without realizing that that time is right now.

Suppose this second thought is *de re* about March 14, at 15.00 sharp.

(I have an Anchored Entity Representation for this time.)

Then our semantics will not be able to distinguish between this *de re* thought and the *de nunc* thought.

Too bad perhaps, but not worse than the problems we just noted about distinguishing between *de re* and *de se*

MSDRT: Model-theoretic Semantics

- This concludes the presentation of the paper ‘Elements of the Attitudes’, of which so far only a slightly abridged German translation appeared under the title ‘Einstellungszustnde ind Einstellungsberichte in der Diskursreprsentationstheorie’ in ‘Intentionalitt zwischen Subjektivitt und Weltbezug’ (Mentis, 2003, ed. Ulrike Haas-Spohn).

  [I am still working on the latexed version of the English original (which will also be adapted somewhat to better fit the general interests of the seminar and will also be a little more explicit uncertain aspects of the formal implementation, in line with these slides.]

  (There are no substantive changes from the original, only some more details, which I hope will make this material more accessible to the reader.)

MSDRT: Model-theoretic Semantics

- Two important topics addressed in the original paper (but worked out in formal detail only in later, so far unpublished work) are:

(1) ‘Self-reflection’: Many of our thoughts are ’higher order’ in the sense that they are about other attitudes we have.

For instance I may wonder about a belief I have whether it is really true, or whether you share it with me.

Or I may have a certain desire and a higher order desire not to have that desire, or the belief that the first desire is the result of the irresponsible prescription of a certain drug

Such higher order thoughts are obviously a very important part of our mental lives.

MSDRT: Model-theoretic Semantics

A distinctive feature of second order thoughts is that their topics – the first order thoughts they are about – are transparent to them, in much the same way as the self and the psychological present.

A satisfactory rendering of this transparency is an important challenge for the representation of second order attitudes.

(2) Attitude attributions involving more than one attributee, as in:

(i) Mary thinks that the thing she is seeing in the distance is a man, but Ella thinks it is a tree trunk.

(ii) Mary thinks there is a gold coin in the middle of the road. Ella thinks there is nothing there at all.

MSDRT: Model-theoretic Semantics

These examples are just the tip of a very large iceberg.

The semantic analysis of the vast majority of these examples requires a proper analysis of what it is for two or more agents to 'share a referent'.

This is a notion that appears to have been hardly investigated in the literature.

Within an MSDRT setting referent sharing is analyzed in terms of coordination of Entity Representations

(For some remarks see Kamp, H. and J. van Genabith and U. Reyle: 'Discourse Representation Theory: An Updated Survey', in Handbook of Philosophical Logic, Vol. XV (D. Gabbay, ed.), Elsevier, 2011.)

Sharing – of propositional contents, reference or more generally parts of mental states – is a topic of the first importance for any account of verbal communication.

MSDRT: Model-theoretic Semantics

- What has also been missing from our presentation is any kind of Construction Algorithm for a fragment of English which contains ways of making attitude attributions, which converts sentences and discourses of this fragment into DRSs from some MSDRS language.

  We will look at some aspects of DRS construction for MSDRT later on (depending on how much time we will have).

- The second main topic of the seminar will be the use of MSDRT in a communication-theoretic approach to reference and to the semantics of referring phrases (i.e. of definite and specifically used indefinite noun phrases).