Richard's Section 2

Richard-1

he means sentence meaning/character

p 431 A general treatment of *de se* ascriptions may be developed along the following lines. First, let us introduce some structure to meanings. Instead of thinking of a meaning as simply a function from contexts to propositions, think of it as a pair $\langle\langle s_1, \ldots, s_n\rangle, M^n\rangle$ ($n \geqslant 0$), where each s_i is a (demonstrative) term-meaning – a function from contexts to individuals – and M^n is an n-place predicate-meaning – a function from contexts to n-place properties. (I will, for the sake of expediency, identify n-place properties with functions from n-tuples of possible individuals to sets of

structured sentence meanings/characters

meaning of "I inform you of her danger"

=

<< meaning of "I", meaning of "you", meaning of "her [specific pointing event]" >, meaning of "inform _ of _'s danger" >

function from context c ↦ the agent of c

p 432 possible worlds; propositions with zero-place properties – *viz.*, sets of worlds.) The proposition such a meaning yields in a context c is, of course, the proposition p such that w is in p exactly if w is in $[M^n(c)](\langle s_1(c), s_2(c), \ldots, s_n(c)\rangle)$.

whereas Richard's sentence meanings/characters are structured, his propositions aren't

the intension of an n-adic predicate is usually a function from world w ↦ set of n-tuples that have the relevant property/relation at that world.
Richard is inverting order of worlds/n-tuples because it makes other things convenient.
Also is just identifying properties with these intensions, not taking them to be finer-grained

uses superscripts to signal "Adicity" (how many arguments are expected?)

Note, now, that we can 'partially interpret' such meanings, relative to a context. For example, if we start with a meaning $m = \langle\langle s_1, s_2\rangle, M^2\rangle$ and a context c, we can 'plug in' the values of s_1 and M^2 in c to get a 'reduced meaning' $m' = \langle\langle s_2\rangle, P^1\rangle$, P^1 the one-place property such that $w \in P^1(u)$ iff $w \in [M^2(c)](\langle s_1(c), u\rangle)$. The reduced meaning m', in turn, corresponds to the function from contexts to propositions which applied to a context c' yields the proposition that the value of s_2 in c' has P^1.

Sentence: "I see you"

its meaning m = << c ↦ c_{Agent}, c ↦ $c_{addressee}$ >, c ↦ seeing relation >

two "reduced meanings" of m, gotten by "partially interpreting" it, relative to a context c*

m' = << c ↦ $c_{Addressee}$ >, property of being seen by the Agent of c* >

m'' (Richard doesn't mention.) = << c ↦ c_{Agent} >, property of seeing the Addressee of c* >

Richard-2

p432

The basic intuition behind the general treatment of *de se* ascriptions we propose is this: A *de se* ascription

(3) a believes himself to be F

is true exactly if a's referent believes the proposition that he is F (*viz*., the proposition that he has the property which is expressed by ⌜is F⌝ relative to the context at which we interpret (3)) under a meaning m which has as one of its reduced meanings $\langle\langle\{I\}\rangle, F\rangle$, where $\{I\}$ is the meaning of 'I'. This, in turn, will be true precisely if a's referent accepts a meaning which is the meaning of a sentence of the form ⌜$\phi(I)$⌝, where $\phi(x)$ expresses, relative to his context, the property F. When someone believes a proposition under such a meaning, we will say that he self-attributes the property, allowing us to state our view in summary form as: (3) is true exactly if a's referent self-attributes the property expressed by ⌜is F⌝.[12]

{TERM} is Richard's confusing notation for "Mng of TERM," see below

could also be written as [[TERM]],

V(TERM)

|TERM|

TERM′

$\overline{\text{TERM}}$

Other uses of {TERM} notation

pp 435-6

we do using the notion of a reduced meaning, introduced above. Where $M = \langle\langle s_1, \ldots, s_n\rangle, M^n\rangle$ is a meaning, a reduced meaning corresponding to M, relative to a context c, is any function in $\mathscr{P}(W)^C$ which results (in the way indicated above) by interpreting M^n and one or more of the s_i, relative to c. An *i*-reduced meaning is any reduced meaning such that (a) not all the s_i's are interpreted; (b) the only s_i's not interpreted are $\{I\}$ ($\{I\}$, of course, is the function which yields c_A, when applied to a context c).

p.444

Let $m = \langle\langle s_1, \ldots, s_n\rangle, M^n\rangle$ be a meaning. The intuitive answer to the question – When does the agent of a context c attribute a property P, in virtue of believing under m? – is as follows. Consider, first of all, what one 'gets' if one (a) replaces M^n with $M^n(c)$ (*viz*., replaces the meaning M^n with the property which is its value in c); (b) replaces each s_i either with its value in c or with a variable; (c) doesn't replace distinct s_i's with the same variable. Call such entities the *proto-properties* associated with m in c.

(For example, proto-properties associated with

$$m_1 = \langle\langle\{t_1\}, \{y_1\}\rangle, \{F_1^2\}\rangle$$

– which could be identified with the meaning of "$F_1^2 t_1 y_1$" – in a context in which "t_1" denotes u, "y_1" denotes u' and "F_1^2" denotes P are

(i) $\langle\langle u, x\rangle, P\rangle$,

(ii) $\langle\langle x, u'\rangle, P\rangle$,

(iii) $\langle\langle x, x'\rangle, P\rangle$.

Proto-properties associated with

$$M_2 = \langle\langle\{t_1\}, \{\cancel{y_2}\}\rangle, \{F_1^2\}\rangle$$ typo: should be t_1

in such a context are all of the above and

(iv) $\langle\langle x, x\rangle, P\rangle$.)

note attached to p.447, next-to-last ¶ of paper

[20] Note, however, that it *is* very often important to us to get across that belief is held under a meaning involving $\{I\}$. One reason for this is that we seem to presuppose the truth of a psychological theory which predicts how people will behave when they so believe (and when they have certain desires, etc.). To effectively make use of such a theory in everyday affairs – in particular, to justify predictions of behavior *via* the theory – we need a way to say that a person believes in the relevant way. It is for reasons such as this that English has a *de se* belief operator like that discussed in Section II. That we have no very general need, as we do for beliefs held under meanings involving $\{I\}$, to say that someone holds a belief under the meaning of a sentence involving $\{that\}$ or $\{you\}$ explains, I think, the absence of belief operators in English which single out beliefs held under such meanings.

SYNTAX

Richard-3

p432

Consider, to begin with, the behaviour of 'believes' in *de dicto* and *de re* ascriptions and in *de se* ascriptions. In the first two sorts of ascriptions, the belief operator – use 'B^r' to represent it – appears to operate on an n-place predicate ($n \geqslant 0$) to yield an $n+1$-place predicate. For example, 'at the level of logical form', 'B^r' combines with 'x loves y' to yield 'zB^r(x loves y)'. The belief operator in *de se* ascriptions, on the other hand – let us use 'B^s' to represent it – apparently combines with an n-place predicate ($n > 0$) and a specification of an argument place to yield an n-place predicate. Thus, for example, applying 'B^s' to 'x loves y' and specifying the first argument place seems to yield something along the lines of 'zB^s (he himself loves y)'.

Richard treats this open sentence (with 3 free variables) as a 3-adic predicate

treats this open sentence (with 2 free variables) as a 2-adic predicate

p433

Of course, given that we do not want *de se* ascriptions to be implied by the corresponding *de re* ascriptions, we cannot assume that something like 'zB^s (he himself loves y)' is reducible to an expression involving 'B^r' and other syntactic operations. For example, we would not want to identify 'zB^s(he himself loves y)' with the result of applying the operation 'identifying the first two argument places' to 'zB^r(x loves y)'. For the latter object – 'zB^r(z loves y)' – will be true, relative to an assignment f, precisely if $f(z)$ believes *de re*, with respect to $f(z)$ and $f(y)$, that the former loves the latter.

A different 2-adic predicate

Thus, we will use two distinct belief operators, 'B^r' and 'B^s', in our formalization. 'B^r' will, as is usual, take a sentential complement. We will, however, have 'B^s' take as complement a 'property abstract' (something of the form $\ulcorner\hat{x}(\phi)\urcorner$, ϕ a sentence). The reasons for treating 'B^s' in this way have, for the most part, to do with elegance in presentation. We could, in principle, allow 'B^s' to take a sentential complement, so long as we introduced apparatus for indicating what argument positions in an embedded sentence are 'specified argument places' in the sense indicated above. Such a treatment, however, is messier than need be.

better to say "predicate abstract"

way of specifying some of the arguments rather than others

for example consider the 2-adic predicate expressed by open sentence:

x prefers x to y

this abstract selects the x arguments as special:

$\hat{x}$ (x prefers x to y)

this abstract selects the y argument as special:

$\hat{y}$ (x prefers x to y)

It should be stressed that the decision to treat the *de se* belief operator in this way does *not* constitute surrender of the view that the objects of belief (*viz*., the contents of belief, in the sense of Section I) are uniformly propositions, nor does it make it at all inappropriate to say that something of the form $\ulcorner\alpha B^s\hat{x}(\phi)\urcorner$ is (a representation of) an ascription of *belief*. Our semantics will take a formula of the form of $\ulcorner\alpha B^s\hat{x}(\phi)\urcorner$ to be true precisely if α's referent believes a *proposition* under a meaning m which has $\langle\langle I\rangle, \overline{\hat{x}(\phi)}\rangle$ as a reduced meaning, where $\overline{\hat{x}(\phi)}$ is the property the semantics associates with $\hat{x}(\phi)$. Furthermore, as we will show, a *de se* ascription will, in this treatment, imply its corresponding *de re* ascription (and thus imply that a certain proposition is believed), although the converse implication, of course, will not hold.

Reduced meaning [in a context c#]

$\langle\langle c\#\ c_{agent}\rangle$, property associated with $\ulcorner\hat{x}(\phi)\urcorner$ [in c#]$\rangle$

officially, Richard's predicates only pick out properties relative to contexts

though in fact he doesn't discuss any examples (except predicates using B^r or B^s) where the semantics gives us different properties/intensions in different contexts

p433

The vocabulary and formation rules for our treatment are as follows. As primitive vocabulary items we have: A denumerable set $V = \{x_1, x_2, \ldots\}$ of variables; denumerable sets $Y = \{y_1, y_2, \ldots\}$ and $T = \{t_1, t_2, \ldots\}$ of demonstrative terms (used to represent, respectively, uses of second person singular 'you' and third person singular demonstratives such as 'he', 'she', 'that', etc.); the singular term: I; for each n, a denumerable set F^n of n-place predicates; the truth functors: $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$; the belief predicates: B^r, B^s; the abstraction operator: $\hat{}$; the quantifiers: $\exists, \forall$; and, as

primitive non-logical vocabulary

V = set of variables

Y = set of uses of 'you'

T = set of uses of 'he' [specific pointing event], etc

term "I"

F^1 = set of 1-adic predicates

F^2 = set of 2-adic predicates

etc.

p434

punctuation, '(', ')'. We use D to name the set of demonstratives of the language, the set $Y \cup T \cup \{I\}$; $\mathscr{T}$, the set of terms, is $D \cup V$.

The definition of well-formed formula is:

1. If $\Pi \in F^n$ and $\alpha_1, \ldots, \alpha_n \in \mathscr{T}$, then $\ulcorner\Pi^n\alpha_1 \ldots \alpha_n\urcorner$ is a formula.
2. If ϕ and Ψ are formulas, then $\ulcorner\neg(\phi)\urcorner$, $\ulcorner(\phi) \wedge (\Psi)\urcorner$, $\ulcorner(\phi) \wedge (\Psi)\urcorner$, $\ulcorner(\phi) \rightarrow (\Psi)\urcorner$ and $\ulcorner(\phi) \leftrightarrow (\Psi)\urcorner$ are formulas.
3. If ϕ is a formula, $\alpha \in V$, then $\ulcorner\exists\alpha(\phi)\urcorner$, $\ulcorner\forall\alpha(\phi)\urcorner$ are formulas.
4. If ϕ is a formula, $\alpha \in \mathscr{T}$, then $\ulcorner\alpha B^r(\phi)\urcorner$ is a formula.
5. If $\alpha \in \mathscr{T}$ and Γ is a proper abstract, then $\ulcorner\alpha B^s\Gamma\urcorner$ is a formula, where a proper abstract is any expression of the form $\ulcorner\hat{\alpha}(\phi)\urcorner$, ϕ a formula and α a member of V which occurs freely in ϕ.
6. These are all the formulas.

we exclude abstracts like $\hat{x}$(y is in danger), where the open sentence has no free occurrences of variable x.

SEMANTICS 1

Richard-4

P.435

We define an interpretation for the language as a quartet $M = \langle U, W, C, V \rangle$ which obeys the following strictures:[14]

1. U, W, and C are non-empty and disjoint sets (which, intuitively, represent possible individuals, worlds, and contexts, respectively).
2. (a) Associated with each member c of C is four-tuple $\langle c_A, c_W, c_Y, c_T \rangle$,
   (i) $c_A \in U$ (c's agent),
   (ii) $c_W \in W$ (c's world),
   (iii) c_Y and c_T are denumerable sequences of members of U (the potential addressees and demonstrata of c).
   (b) $c = c'$ iff $c_A = c'_A$, $c_W = c'_W$, $c_Y = c'_Y$, and $c_T = c'_T$.
   (c) No world contains distinct contexts with the same agent.
3. V is a function which assigns
   (a) a member of $((\mathcal{P}(W))^{U^n})^C$ to each member of F^n, for each n;
   huh? (b) sets of meanings to each member of C, where a meaning is a pair $\langle\langle s_1, \ldots, s_n \rangle, M^n \rangle$ $(n \geq 0)$, each $s_i \in U^C$ and M^n a member of $((\mathcal{P}(W))^{U^n})^C$.

A word on the workings of V is perhaps in order here. V's assignments to predicate letters are, intuitively, predicate-meanings (taken to be functions from contexts to properties). V's assignments to contexts are to be understood as representing the class of meanings under which the agent of the context holds beliefs; in the terminology of Section I, $V(c)$ is the set of meanings which c_A accepts. Note that, for each context c, $V(c)$ determines a set of propositions, a proposition p being in the set so determined by $V(c)$ exactly if, for some m in $V(c)$, m, completely interpreted relative to c, yields p. These, of course, are the propositions which are objects of belief of the agent of c.

Model/structure/interpretation
= collection of all the parameters or tools necessary to specify meanings for primitive non-logical vocabulary

U = people (or other objects designated by names, demonstratives, variables)

W = worlds

C = contexts
= ⟨Agent in the context, world of the context, sequence [designated by you$_1$, ...], sequence [designated by she$_1$, ...]⟩

Agents are in (at most) one context/world
(no dealing w/ different times, possibility of addressing / demonstrating different people at different times)

(3A) V (A 2-adic predicate like 'sees') = member of $(\mathcal{P}(W)^{U^2})^C$ =

Some function from $C \to (U^2 \to \mathcal{P}(W))$
U^2 = pairs of people; $\mathcal{P}(W)$ = sets of worlds

Richard's way of representing 2-adic properties/intensions

(3b) V (A context c*) = the set of structured sentence meanings/characters accepted by the agent of c* (in the world of c*).

each is a ⟨⟨function from context to people, ...⟩, function from context and n-tuples of people to set of worlds⟩
(n-tuples of people to set of worlds = property intension)

SEMANTICS 2

$\boxed{\text{Richard-5}}$

p436

To define truth and denotation in an interpretation (reference to which is continually suppressed), we proceed as follows. The denotation of a term α, relative to a context c, assignment (member of U^V)f, and world w (write: $|\alpha|_{cfw}$) is defined: $f(\alpha)$, if $\alpha \in V$; c_A, if $\alpha = I$; C_{T_i}, if α is t_i; C_{Y_i}, if α is y_i. We begin the definition of ϕ, taken relative to c and f, is true at w (write: $cf[\phi]w$) as follows:

1. $cf[\Pi^n\alpha_1 \ldots \alpha_n]w$ iff $w \in [V(\Pi^n)(c)](\langle|\alpha_1|_{cfw}, \ldots, \alpha_n|_{cfw}\rangle)$
2. $cf[(\phi) \wedge (\Psi)]w$ iff $cf[\phi]w$ and $cf[\Psi]w$.

And so on, for the other truth functors.

3. $cf[\exists\alpha(\phi)]w$ iff $\exists u(u \in U$ and $cf^\alpha_u[\phi]w)$.

Analogously for $\forall\alpha(\phi)$.

4. $cf[\alpha B^r(\phi)]w$ iff $\exists c'(c'_A = |\alpha|_{cfw}$ & $c'_w = w$ & $\exists m(m \in V(c')$ & $m(c') = \{w' | cf[\phi]w'\}))$.

$m(c')$ here is the proposition yielded by m in c', defined as above.

The intuitive content of clause (4) is this. $\alpha B^r(\phi)$, taken relative to c and f, is true exactly if: There is a meaning m such that α's denotatum accepts it (formally: $m \in V(c')$, c' the context of α's denotatum), and m yields, relative to c', that proposition expressed by ϕ relative to c. Note that this clause has the result (given that a person believes a proposition p if he accepts a meaning which yields p relative to his context) that $\alpha B^r(\phi)$ is true iff what α denotes believes the proposition expressed by ϕ.

$\exists x \phi$ is true relative to c, f, w
iff
there's a person $u \in U$ such that
ϕ is true relative to c, f', w
where f' is like f except that it maps x to u,
I'd write f' as: $f[x := u]$
Richard writes as: f^x_u

$|\text{variable } x|_{cfw} = f(\text{'}x\text{'})$, that is, what person $\in U$ does assignment function f map variable 'x' to

$|I|_{cfw}$ = Agent of c

sentence ϕ is true relative to
context c
Assignment function f
world w

I might write as:
$[\![\phi]\!]_{cfw}$ = true

Richard writes as:
$cf[\phi]w$

Some write as:
$c, f, w \vDash \phi$

section 2's account of B^r
section 3 will make more complicated

SEMANTICS 3

Richard-6

Preliminaries to giving Semantic rule for B^s

p435

We must, in order to give a definition of truth, characterize the conditions under which the agent of a context self-attributes a property. This we do using the notion of a reduced meaning, introduced above. Where $M = \langle\langle s_1, \ldots, s_n\rangle, M^n\rangle$ is a meaning, a reduced meaning corresponding to M, relative to a context c, is any function in $\mathcal{P}(W)^C$ which results (in the way indicated above) by interpreting M^n and one or more of the s_i, relative to c. An *i*-reduced meaning is any reduced meaning such that (a) not all the s_i's are interpreted; (b) the only s_i's not interpreted are $\{I\}$ ($\{I\}$, of course, is the

Replace predicate's meaning/character with the property/intension it expresses in context c, filling some of its argument places with the objects $s_i(c)$

leave unreduced some (need it be all?) of the meaning/characters for "I"

p436

function which yields c_A, when applied to a context c). Where M is a meaning, we denote the set of *i*-reduced meanings of M, relative to c, by $M^{i,c}$. A member M_1 of $M^{i,c}$ is said to attribute a one-place property P just in case, for any context c' and world w

$$w \in M_1(c') \text{ iff } w \in P(c'_A).$$

When an $M_1 \in M^{i,c}$ and property P are so related, we write: $P \in [M^{i,c}]$. We can now say that the agent of a context c self-attributes the property P precisely if there is an M in V_c such that $P \in [M^{i,c}]$.

set of all such

one such reduced meaning M_1 $\boxed{\text{attributes a property } P}$ iff for every context c', the worlds where $M_1(c')$ are true = the worlds where the agent of c' has P.

Note that though Richard sometimes talks of the open sentence "x sees y" and the predicate abstract "$\hat{x}$(x sees y)" as both being 2-adic predicates, his semantics will treat the first like a sentence (assign it a truth value, relative to c, f, w) but will assign the second a (1-adic) property intension

p436

Let $\hat{\alpha}(\phi)$ be a proper abstract. We say that P is the implied property of $\hat{\alpha}(\phi)$, taken relative to c and f, if and only if P is the one-place property such that, for all u and w,

$$w \in P(u) \text{ iff } cf_u^{\alpha}[\phi]w.$$

An abstract $\hat{x}(\phi)$ $\boxed{\text{implies relative to } c, f}$ $\boxed{\text{that property } P}$ such that for all people u, worlds w, w is a world where u has P iff ϕ is true relative to c, f[x:= u], w

Writes P as $\overline{\hat{x}(\phi)}^{cf}$

p437

We use $\overline{\hat{\alpha}(\phi)}^{cf}$ to denote the implied property of $\hat{\alpha}(\phi)$, taken relative to c and f. We may complete our definition of truth by saying that a *de se* ascription $\alpha B^s\hat{\beta}(\phi)$, taken relative to c and f, is true at w precisely if: α's denotatum believes a proposition under a meaning which has, as one of its *i*-reduced meanings, one which attributes $\overline{\hat{\beta}(\phi)}^{cf}$ – that is, just in case α's denotatum self-attrubutes $\overline{\hat{\beta}(\phi)}^{cf}$. Formally, we have

5. $$cf[\alpha B^s\hat{\beta}(\phi)]w \text{ iff } \exists c'(c'_A = |\alpha|_{cfw} \,\&\, c'_w = w \,\&\, \exists m(m \in V(c') \,\&\, \overline{\hat{\beta}(\phi)}^{cf} \in [M^{i,c'}]).$$

These semantics adequately capture the view of the truth conditions of *de se* ascriptions discussed at the beginning of this section. In particular, they have the consequence that a *de se* ascription implies (what we will presently define as) its corresponding *de re* ascription, although the converse implication does not hold. Thus, something of the form $\alpha B^s\hat{x}(\phi)$ involves an ascription of *belief*: The ascription is true only if α's denotatum believes the proposition ϕ expresses, when the denotatum of α is assigned to x.

sentence $\alpha B^s \hat{x}(\phi)$ is true relative to c, f, w

iff $|\alpha|^{cfw}$ accepts a meaning/character that has a filled-in meaning that attributes property P

iff $|\alpha|^{cfw}$ self-attributes P

where P is the property $\hat{x}(\phi)$ implies relative to c, f $= \overline{\hat{x}(\phi)}^{cf}$

p437

We define the *de re* ascription corresponding to a *de se* ascription $\Psi = \alpha B^s\hat{x}(\phi)$ as follows. Let v be the least (i.e., with smallest subscript) variable not occurring in $\alpha B^s\hat{x}(\phi)$. The *de re* ascription corresponding to Ψ is then

$$\exists v(v = \alpha \wedge vB^r(\phi')),$$

where ϕ' is ϕ with all free occurrences of x replaced by v. (We of course understand the expression $\hat{\alpha}$ to bind free occurrences of α within its scope.) Thus, for example, corresponding to

$$IB^s\hat{x}_1(x_1B^s\hat{x}_1(Fx_1))$$

is

$$\exists x_2(x_2 = I \wedge x_2B^r(x_2B^s\hat{x}_1(Fx_1))).$$

It follows fairly directly from the above definitions that whenever a *de se* ascription, taken relative to c and f, is true at w, then so is its corresponding *de re* ascription. Of course, the converse does not hold. For example, if $V(c')$ consists solely of the meaning of 'Ft_s', 't_s' denotes c'_A relative to c',

$$\exists x_1(x_1 = I \wedge x_1B^r(Fx_1))$$

will be true, relative to c' and an assignment f, at c_w, but

$$IB^s\hat{x}(Fx)$$

will not.

going from de se ascription (here Ψ) to a "corresponding" de re ascription

eg. from I believe (myself to believe (himself to be F)) to I am an x_2 who believes of x_2 that he believes (himself to be F)

SECTION 3 – Intuitive Motivation

Richard – 7

p 442

Let us now return to the original case. It is clear what we will say about this case, if we accept the view of belief above labelled the triadic view. We will say that *A* believes the proposition – that *B* can be informed of her danger via the phone – under the meaning of the embedded sentence of

(1) I believe that I can inform you of her danger via the telephone.

but not under the meaning of the embedded sentence of

(2) I believe that I can inform her of her danger via the telephone.

This analysis shouldn't be terribly puzzling, even given that *A* understands both sentences and knows of each, and the proposition it expresses, that the former expresses the latter. For, as *A* doesn't know that his uses of 'she' and 'you' are co-referential, he can hardly be expected to know that the embedded sentences express the same proposition.

Compare, now, the position of *A* with that of a person *X*, who is in the same situation as *A*, but who knows that the woman he sees is the woman to whom he is speaking. *X* will hold a belief about *B* under both the meanings mentioned above. He will also differ from *A* in the following way: There will be a woman whom *X* believes to have the property *being such that she can be informed of her danger via the phone*. It seems that we cannot explain this difference between *A* and *X* in terms of proposition believed, since both of them believe the proposition that *B* can be informed of her danger via the phone. In order to explain the difference, we must appeal to *how A* and *X* hold their beliefs. It would seem that to believe the proposition expressed (relative to a context *c*) by a sentence in which demonstratives occur is to have a *de re* belief with respect to the objects denoted, in *c*, by the demonstratives in the sentence. If one has a *de re* belief with respect to an object, then one may be said to *attribute* certain properties to the object. However, it does not follow, from the fact that *x* and *y* each believe the proposition *p* expressed in *c* by a sentence *S*(*d*), *d* a

$\exists u\,(X \text{ believes } u \text{ is a } \underbrace{\hat{y}(y \text{ can be informed of } y\text{'s danger by phone})}))$

p443

demonstrative occuring in *S* and denoting *u* in *c*, that every property which *x* attributes to *u*, in virtue of his believing *p*, is one which *y* attributes to *u*, in virtue of this belief. For which properties one attributes to an object is determined by the meaning under which one's belief is held: *X*, for example, who believes the proposition, that he can inform *B* of her danger via the phone, under the meaning of 'I can inform her of her danger via the phone' will attribute to *B* the property *Being a thing that can be informed of its danger via the phone*: *A*, who doesn't believe the proposition under the meaning just mentioned, will not attribute this property to *B*.

If this much be accepted, we have the basis of an answer to the question: How can

(1) I believe that I can inform you of her danger via the telephone.

and

(2) I believe that I can inform her of her danger via the telephone.

diverge in truth value in a context in which their embedded sentences express the same proposition? For we may say: An ascription of belief ⌜a believes that *S*⌝, *S* a sentence in which demonstratives occur, not only implies that the proposition expressed by *S* is believed, but that certain properties are attributed to the referents of the demonstratives in *S*. What properties the ascription implies are attributed depends, in turn, upon the meaning of *S*. In the case in question, ascription (2) implies that a property (that assocaited with a use, in this context, of 'I can inform *x* of *x*'s danger by phone') is attributed which (1) does not imply is attributed. Hence, (1) may true be while (2) is not.

☆

Richard's A only attributes the 2-Adic Relation

$\hat{y}\hat{z}(y \text{ can be informed of } z\text{'s danger by phone})$

to the pair of the woman^B he's talking to and the woman^B he's seeing

(this is a pair of the same person^B taken twice, but A doesn't realize that).

SEMANTICS 4 – from Section 3

Richard-8

p.443

Let us consider how we might give a systematic development of this proposal. In order to simplify matters, we will do this for a language with only a *de re* belief operator; it will be obvious how the treatment would be generalized to a language including a *de se* operator such as that discussed in Section II.

We assume, then, that our language has the same primitive vocabulary as the language of Section II, minus the B^* operator and the abstraction operator; the formation rules are identical to those of Section II, save the omission of the clause of the *de se* operator. We preserve the definitions of interpretation, denotation, and the clauses of the truth definition for atomic,

p.444

(a)

(b)

truth functional, and quantified sentences. We now need to characterize, in terms of the formal structure, two things: When an individual, in believing a proposition under a meaning, attributes a property, and when a belief ascription, taken relative to a context, implies the attribution of a property.

p.444

Let $m = \langle\langle s_1, \ldots, s_n\rangle, M^n\rangle$ be a meaning. The intuitive answer to the question – When does the agent of a context c attribute a property P, in virtue of believing under m? – is as follows. Consider, first of all, what one 'gets' if one (a) replaces M^n with $M^n(c)$ (*viz*., replaces the meaning M^n with the property which is its value in c); (b) replaces each s_i either with its value in c or with a variable; (c) doesn't replace distinct s_i's with the same variable. Call such entities the *proto-properties* associated with m in c. (For example, proto-properties associated with

meaning/character of t_1 etc.

$$m_1 = \langle\langle\{t_1\}, \{y_1\}\rangle, \{F_1^2\}\rangle$$

– which could be identified with the meaning of "$F_1^2t_1y_1$" – in a context in which "t_1" denotes u, "y_1" denotes u' and "F_1^2" denotes P are

(i) $\langle\langle u, x\rangle, P\rangle$,

(ii) $\langle\langle x, u'\rangle, P\rangle$, metalanguage names/constants designating distinct variables

(iii) $\langle\langle x, x'\rangle, P\rangle$. even if u=u', won't get AS A proto-property of sentence that uses distinct demonstratives "she/her", "you"

Proto-properties associated with

$$M_2 = \langle\langle\{t_1\}, \{\cancel{t_2}\}\rangle, \{F_1^2\}\rangle$$ typo: should be t_1

in such a context are all of the above and

(iv) $\langle\langle x, x\rangle, P\rangle$.)

p.444

To each proto-property there corresponds, in a rather obvious way, a property. For example: to (ii) corresponds the one-place property P^1 such that $w \in P^1(u_1)$ iff $w \in P(\langle u_1, u'\rangle)$; to (iii) corresponds the two-place property P^2 such that $w \in P^2(\langle u_1, u_2\rangle)$ iff $w \in P(\langle u_1, u_2\rangle)$; to (iv) corresponds the one-place property P^3 such that $w \in P^3(u_1)$ iff $w \in P(\langle u_1, u_1\rangle)$.

Answer to (a)

We can now answer our initial question thus: An agent attributes a property P, in virtue of holding a belief under a meaning m iff P corresponds to one of the proto-properties associated with m relative to the agent's context. We will write

$$P \qquad P \in P(m, c)$$??

for: the agent of c attributes P, in vritue of holding a belief under m.

p.445

Re-asking (b)

A fully rigorous characterization of the above notion would disperse with the notion of a variable in the construction of proto-properties. It is easy enough to give such a characterization; we henceforth assume that the predicate $P(m, c)$ has been so defined in terms of our model structure. We now need a way to get from a sentence (taken relative to a context and an assignment) used to ascribe belief to the set of properties ~~it implies the~~ believer attributes. One way of doing this is as follows. Consider a sentence ϕ; let $\alpha_1, \ldots, \alpha_n$ be a complete enumeration of those demonstratives and variables (which occur freely) in ϕ. Let $v_1, \ldots, v_n$ be variables which do not occur in ϕ. We say that Ψ is a frame of ϕ just in case $\cancel{\phi}$ is the result of replacing one or more of the α_i's with v_i's, subject to the restriction that distinct α_i's are replaced with distinct v_i's.

typo: should be Ψ

Thus, for example, consider the sentences

(i) $F_2^2t_1y_1$,

(ii) $F_2^2t_1t_1$.

Frames of (i) are: $F_2^2t_1x_1$, $F_2^2x_1y_1$, $F_2^2x_1x_2$; frames of (ii) are the above and $F_2^2x_1x_1$. Note that this last is not a frame of (i).

SEMANTICS 5

Richard-9

p. 445

Answer to (6)

We say that a sentence ϕ implies the attribution of the property P^n, relative to c and f, just in case there is a frame ψ of ϕ, obtained by substituting the n distinct variables $v_1, \ldots, v_n$ for terms in ϕ and, for every w and $u_1, u_2, \ldots, u_n$:

$$cf^{u_1 u_2 \ldots u_n}_{v_1 v_2 \ldots v_n}[\Psi]_w \quad \text{iff} \quad w \in P^n(\langle u_1, u_2, \ldots, u_n \rangle).$$

We define the attribution class of a sentence ϕ, relative to c and f, as the set of those properties such that ϕ implies their attribution, relative to c and f; we denote this class with $A(\phi, c, f)$.

We now define truth for *de re* ascriptions of belief:

$$cf[\alpha B^r \phi]w \text{ iff } \exists c'(|\alpha|_{cfw} = c'_A \;\&\; c'_w = w \;\&\; \exists m(m \in V'_c \;\& \; m(c') = \{w' | cf[\phi]w'\} \;\&\; (f)(f \in A(\phi, c, f) \rightarrow f \in P(m, c')))),$$

where $m(c')$ is the proposition expressed by m relative to c'. Verbally, these truth conditions amount to this: $\alpha B^r \phi$, relative to c and f, is true exactly if there is a meaning m such that (i) $|\alpha|_{cfw}$ believes a proposition under m; (ii) m yields, relative to $|\alpha|_{cfw}$'s context, whatever ϕ expresses, relative to c and f, and (iii) whatever properties ϕ implies are attributed are such that belief under m requires their attribution.

☆

This permits that m (accepted by agent) may imply attribution of additional properties, that the report may be silent about.

Thus if A accepts "bRb", that attributes 1-adic property $\overline{\hat{x}(xRx)}$

2-adic relation $\overline{\hat{x}\hat{y}(xRy)} = \overline{R}$

A report "A believes (bRc)" can be true here (only implies that A attributes)

Generally, the reports are allowed to be less demanding than what agent has in their head

p. 446

It is easy to show that, given this semantics, representatives of sentences (1) and (2) can diverge in truth value relative to a context in which their embedded sentences express the same proposition.[17] On the other hand, the semantics validates, the claim, for which we argued above, that in any context in which the uses of 'she' and 'you' in

(3) I believe that she is in danger.

and

(4) I believe that you are in danger.

are co-referential, the truth of (4) is implied by the truth of (3).

17 We can also show that the semantics validates certain forms of 'quantifying in'. Precisely, given our semantics, we have:

> If β is a member of D which occurs in ϕ, then if $cf[\alpha B^r(\phi)]w$, then $cf[\exists v(\alpha B^r(\phi[\beta/v]))w$, provided that β is free for v in ϕ.

(If our semantics had allowed for the possibility that members of D failed to denote in some contexts, this rule would have to be weakened. For simplicity's sake, we have not allowed for this possibility.) That such a rule is sound justifies, in part, the claim that something of the form of $\ulcorner \alpha B^r(\phi) \urcorner$ is a *de re* ascription, provided that ϕ contains a member of D.

Note that not very 'way of quantifying in' is permitted by our semantics. In particular, from

(i) $t_1 = t_2 \wedge IB^r(F^2 t_1 t_2)$

the formula

(ii) $\exists x_1 \exists x_2 (x_1 = x_2 \wedge IB^r(F^2 x_1 x_2))$

follows, but

(iii) $\exists x_1 (x_1 = x_1 \wedge IB^r(F^2 x_1 x_1))$

does *not* follow. Given our reasons for adopting the treatment we have adopted, of course, one would not want (iii) to follow from (i).

WRAP-UP, explains why (3) and (4) sound different though

Richard-10

According to Richard they express the same proposition

p.446

It is, perhaps, worth discussing sentences (3) and (4) again. Many people, even after a rehearsal of the argument given above – that (4) is implied by (3) – are still uncomfortable with the claim that both (3) and (4) are true. A virtue of the semantics just presented, I think, is that it can be used to motivate an explanation of why the intuition, that (3) and (4) diverge in truth value, is so persistent.

Take a finite set of sentences and conjoin them; form what we called a frame of the result. (For example, if you start with {that$_2$ is sad, you$_3$ will make that$_4$ happy if that$_2$ helps you$_3$}, you will end up with something along the lines of 'x_2 is sad $\wedge$ x_3 will make x_4 happy if x_2 helps x_3'.) Call the property associated with such a sentence *a picture*; if all the members of the initial set are sentences, the meanings of which are accepted by an agent *u*, say that the resulting property is a picture *held by u*.[18]

The intuition motivating our semantical account is that an ascription is true provided is ascribes belief in a proposition which is believed and the ascription doesn't imply anything false about what pictures are held by the believer. Since sentence (4), as used by *A*, does not *when taken by itself* imply anything false about what pictures *A* holds, (4) so taken is true, since *A* believes *B* to be in danger.

Note, now, that a set of belief ascriptions may (conventionally) imply things about the pictures a believer holds that the conjunction of the members of the set does not (strictly) imply.[19] For example, the use of the ascription '*A* believes that you$_1$ are unhappy because she$_2$ spurned you$_1$' in a context in which the ascription '*A* believes that she$_2$ loves a Greek' has been used (and no one has disputed the truth of the latter ascription) will imply that *A* holds the picture associated with '*y* loves a Greek and *x* is unhappy because *y* spurned *x*'. Both ascriptions can be true, even if *A* doesn't hold

p.447

this picture; however, their joint use, in such a case, would be very misleading.

In general, we tend to avoid using an ascription $\ulcorner\alpha$ believes that $\phi\urcorner$, if an ascription $\ulcorner\alpha$ believes that $\psi\urcorner$ is assumed by all the parties to the conversation to be true (and we know this), and we think that the person to whom belief is being ascribed does not hold pictures associated with frames of $\ulcorner\phi$ and $\psi\urcorner$. Likewise, we will find an ascription $\ulcorner\alpha$ believes that $\phi\urcorner$ bizarre or objectionable if it is assumed by those conversing that the ascription $\ulcorner\alpha$ believes that $\psi\urcorner$ is true and we have good reason to think that the believer doesn't hold all the pictures associated with $\ulcorner\phi$ and $\psi\urcorner$.

All of this, I beleive, helps to explain why some find the assertion, that *A*'s use of

(4) I believe that you are in danger.

is true, counter-intuitive, even after a rehearsal of the argument that *A*'s use of (4) cannot be false if his use of (3) is not. For as we have just seen, without qualification and explanation, the claim that (4) is true relative to *A*'s context is very misleading. For obviously, in the case under consideration

(9) I believe that I am talking to you.

is true relative to *A*'s context. Thus, without further qualification, the claim that (4) is true implies that

(10) I believe that I am talking to someone who is in danger.

is true, relative to *A*'s context. But, obviously (10) is not thus true.

(3) I believe that she is in danger
—clearly true

(4) I believe that you are in danger
—Sounds bad, but Richard argues it's equivalent to

(6) The man watching you believes that you are in danger
—which sounds good.

On his view, all of these sentences express (in the context he describes) the same proposition. So they have to have the same truth-values.

[18] Strictly speaking, of course, we can associate properties with open sentences possibly containing demonstratives only relative to a context. My ignoring that here does not effect the point.

[19] I must stress that 'implies' is being used in two senses in this sentence. The first use of 'implies' is quite weak (certainly not the sort of implication which preserves truth). Roughly, the use I intend here is the sort present in (typical) uses of 'His saying that the movie was boring implies that he did not like it'.

Appendix 1

note ATTAched to p.429, sect.1

[8] I ought to say something here about what these meanings are, and how they differ; what needs to be made clear is what the meaning of terms like 'you' and 'she' is.

I presume the following (and do not suggest that it is an original view; it is a version of Kaplan's own view). There are what we might call 'modes of demonstrating' things and 'modes of addressing' things. These modes are such that the same mode can be used in different contexts or several times in one context. It is only when 'she' is accompanied by a mode of demonstrating ('you' is accompanied by a mode of addressing) that it refers to an object. Furthermore, although 'she' plus mode *m* of demonstrating ('you' plus mode *m'* of addressing) may pick out different objects in different contexts, "she" accompanied by one mode of demonstrating picks out the same object every time it is used in a context; analogously for 'you'.

The meaning (in Kaplan's sense of meaning as character) of 'she', then, is roughly this: 'she', accompanied by a mode of demonstrating, functions as a directly referential term; it denotes, relative to a context, what its accompanying mode of demonstrating demonstrates.

Thus, in giving formal representatives for sentences such as those mentioned in the text, what we really represent is the sentence type and aspects of the modes of demonstration or address. (For we wish to be able to assign the representatives of propositions to the formal representatives of sentences; the sentences being represented don't express propositions, on the view assumed here, unless accompanied by modes of demonstration or address.) We thus represent two occurrences of 'she' (of 'you') with the same term if and only if they are accompanied by the same mode of demonstration (or address).

note Attached to p.432, sect 2

Appendix 2

[12] To those familiar with views of *de se* belief advanced by Chisholm in [1] and Lewis in [4], this will sound somewhat familiar. Chisholm introduces a primitive notion *x directly attributes property P to y* which, according to Chisholm, is necessarily reflexive. Chisholm then says that to believe oneself to be F is to directly attribute F to oneself. Lewis suggests that we understand belief *de se* as the *self-ascription* of property.

There are several important differences between our approach and the approaches of Chisholm and Lewis. We do not hold that properties are the objects of *de se* belief, as do Lewis and Chisholm; we also hold that the objects of all beliefs are of uniform character, unlike Chisholm.

On Chisholm's view, it is somewhat mysterious as to why one can directly attribute properties only to oneself. Indeed, for Chisholm, there is no real correlate of direct attribution, relating distinct individuals and a property: Chisholm's indirect attribution (in terms of which Chisholm defines *de re* belief) is simply a complicated form of direct attribution.

On our view the reflexivity of self-attribution is not mysterious at all: It's reflexive because it involves meanings which contain $\{I\}$. Furthermore, we could define a perfectly analogous notion of indirect attribution, without invoking the notion of self-attribution, if we wished. Indeed, something like this is defined in Section III, below.

We have analogous differences with Lewis, who characterizes belief *de re* in [4] as a kind of belief *de se*. (For Lewis, as for us, the objects of belief are of uniform character; but, unlike us, he takes them to be all properties.)

It is worth noting that the formalization introduced in this section could be used, with some alterations, to regiment Lewis' view. (The major alterations would be to drop the 'B^r' operator introduced below, translating English sentences of the form of *a believes that S*, where S involves no reflexivies, as: $aB^s\hat{\alpha}\ (\alpha = \alpha \wedge \phi)$. One would also be required, in a formalization of Lewis' view, to prohibit quantification into 'B^s', and to come up with a scheme to represent *de re* ascriptions. This is discussed at the end of Section II.) This should not hide the fact that there are fundamental differences in motivation between Lewis and ourselves. Beyond those mentioned above, we note that this essay and its formalism is intended to function in the defense of the thesis of direct reference, a thesis which – insofar as it is bound up with what Lewis and Kaplan call 'haecceitism' – is anathema to Lewis.

end of sect 2, pp 438-9

There is a sense in which the semantics allows us to dispense with 'B^r' and make do with only 'B^s' as a belief predicate. For we can define 'B^r' using a schema along the lines of

$$\alpha B^r(\phi) = df \qquad \alpha B^s\hat{\beta}(\beta = \beta \wedge \phi)$$

With some minor tinkering, this would be an adequate definition. (The tinkering required is this: As it stands, it's not the case that

$$\alpha B^r(\phi)$$

and

$$\alpha B^s\hat{\beta}(\beta = \beta \wedge \phi)$$

always agree in truth value, since (speaking very loosely) the latter's truth requires that the believer believe under the meaning of $\ulcorner I = I \wedge \phi \urcorner$, while the former requires simply belief under the meaning of ϕ. Now, although these meanings are identical when *conceived as functions* from contexts to propositions, they are not identical when conceived, as in our semantical system, as ordered n-tuples of the meanings of constituent expressions. Thus, to implement the above definition, we'd need to impose a requirement on the function V in our models to the effect that $\{I = I \wedge \phi\} \in V(c)$, if $\{\phi\} \in V(c)$.)

However, such a definition has little, philosophically, to recommend it. The possibility of such a definition does not show that, in our regimentation, belief *de dicto* and *de re* are kinds of, or are reducible to, belief *de se*. (What it shows, I think, is that our system is committed to the thesis that anyone who believes a proposition p believes that he's himself and p, and the converse.) And it is certainly not the case that such a definition is what authors like Lewis [4] and Chisholm [1] have in mind when they suggest that belief *de re* is a kind of belief *de se*.

To take Lewis as an example: His view is that to believe *de re* of u that she's F is to self-ascribe the property *bearing R to one and only one thing, a thing that's F*, where R is a 'suitable' relation and one indeed bears R to u and u alone. On such a view, *de re* belief isn't to be represented via quantification into the belief context (as we have represented it), nor will someone with such a view be sympathetic with our treatment of belief ascriptions involving demonstratives other than "I" (which is, in part, designed to represent such ascriptions as ascriptions of belief in propositions 'singular' with respect to the referents of the demonstratives). What is critical to regimenting Lewis' view is not eliminating 'B^r' in favor of 'B^s' (although that's involved), but giving a procedure for representing ascriptions, which appear to involve quantifying in, as not involving it.