Greg Restall, *Arché, University of St Andrews*

Practice run for the 2025 Wendy Huang Lectures, NTU Taiwan

In my first lecture, I introduced INFERENTIALIST SEMANTICS, giving an account of how you can understand basic logical vocabulary—the propositional logical constants of conjunction, disjunction, negation, etc.—as *definable* using invertible rules of inference. Taking these inference rules as *definitions* means that you have a ready explanation of how it is that a proof can be completely gap-free. To adopt $\to Df$ as definitional of the material notion of implication is to take it that to show $A \to B$, it suffices to show B while taking A for granted, and conversely, if we have granted $A \to B$ and granted A, then B must follow. To ask for anything more is to fail to take '$\to$' to be defined by $\to Df$. The defining rules, together with the fundamental structural rules governing proof *as such* are the only transitions in our gap-free proofs. Since these defining rules conservatively extend those background structural rules, they are safe to adopt.

$$\frac{\mathcal{C}, A \succ B}{\mathcal{C} \succ A \to B}\ {\to}Df$$

In the second lecture, I showed how a this account generalises to quantifiers, and to modal operators, with one difference. The defining rules used for these concepts appeal to distinctive features of our linguistic and conceptual practices. For the first order *quantifiers*, our language must involve a category of *singular terms*, satisfying constraints on *substitution* of one term for another.

$$\frac{\mathcal{C} \succ A}{\mathcal{C} \succ \forall x A}\ \forall Df \quad (x \text{ is not free in } \mathcal{C}.)$$

To give defining rules for the modal operators, we expand our background structural account of proof to encompass not only acts of assertion, denial, inference and purely *material* supposition (consisting of temporarily granting something 'for the sake of the argument', by adding it to the common ground, leaving the rest unchanged), but also properly *modal* supposition. To modally suppose A is to consider what *would* be the case *had* A obtained. This is, of course, compatible with denying A, and it does not involve *withdrawing* your denial of A. To modally suppose A is to add A to the conversational context in a *new zone*. We can continue to use our reasoning capacities with this considered content in its zone, isolated from our commitments concerning how things are. The isolation is not total, however, because cross-zone commerce is underwritten by modal operators. To say take it that A is *necessary* is to take it that A not only holds, but *would* hold had things gone otherwise. Tagging our different zones with labels, there is no clash between asserting A at a and denying it at another zone b, but there *is* a clash between asserting $\Box A$ at a and denying A at b. To deny A at b is, at least implicitly, to deny that A is necessary.

$$\frac{\mathcal{C} \succ A \cdot i}{\mathcal{C} \succ \Box A \cdot j}\ \forall Df \quad (i \text{ is not present in } \mathcal{C}.)$$

This modal logic contains a varying domain inner *existence-entailing* quantifiers and constant domain outer *possibilist* quantifiers, given the discipline governing variables discussed in the previous lecture.

Combining these defining rules for the first-order quantifiers and modal operators gives an inferentialist semantics for first-order quantified modal logic, where the inference rules governing the logical connectives describe what one *does* to use the concepts in making assertions, denials, inferences, and suppositions, i.e., to use them in your talk and in your thought. ¶ To adopt an inferentialist semantics is not to *reject* the traditional truth-conditional model theory for that vocabulary. I showed how the familiar models for first-order modal logics arise as the limit of a process of settling issues in the structured *positions* that feature in deduction, and that the model-theoretic 'semantics' for modal logic can be adopted, on inferentialist lines. This development can be seen as a way to vindicate Arthur Prior's view that we understand possible worlds because we first understand the notions of possibility and necessity. This inferentialist semantics is one way to show *how* to acquire the concepts of possibility and necessity in such a way that possible worlds talk is an appropriate and useful *model* of its logical structure.

One motivation for this line of inquiry this challenge from Timothy Williamson. Speaking of Robert Brandom, he says his "inferentialism has remained at an even more programmatic stage than Dummett's, lacking an equivalent of Dummett's connection with technical developments in proof theory by Dag Prawitz and others. As a result, inferentialism has been far less fruitful than referentialism for linguistics. In that crude sense, referentialism beats inferentialism by pragmatic standards" [11, p. 34]. While I do not so much care about what perspective 'beats' another, Williamson does prompt the question of how inferentialism bears on model-theoretic semantics and its 'referentialist' commitments.

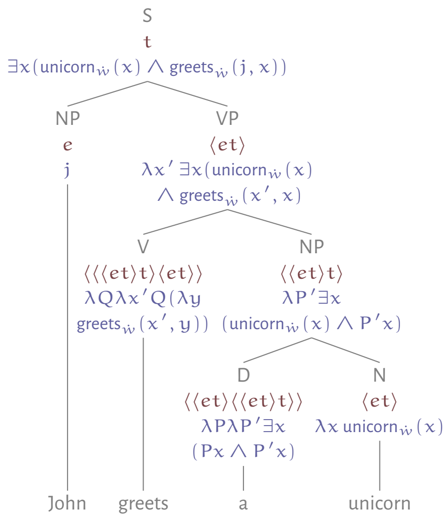
3.1 THERE IS MORE TO LANGUAGE THAN LOGIC

I take this inferentialist semantics to have many virtues. However, I have been playing on the logicians' *home turf*, focussing on logical connectives, quantifiers, and modal operators. There is much more to language and meaning than tiny fragment of our vocabulary. While it is not my job to give a comprehensive meaning theory for every concept that we can use in our thought and talk, I would do well to give some explanation of how the inferentialist 'semantics' I offered for the logical concepts might relate to the kinds of semantic theories offered by linguists and philosophers of language.

Semantics is an incredibly rich and diverse field, and I cannot do justice to it in a short lecture [6]. I will focus my attention to the broad field of *truth-conditional* semantics, pioneered by Richard Montague [10], David Lewis [5] and Barbara Partee [7, 9], in the 1970s. Modal model theory is central to this enterprise. Semantic values are supplied for natural language lexical items, in the vocabulary of an *intensional type theory*. Lexical items of different parts of speech are interpreted as different *types*, which compose in a compositional manner, to provide an interpretation for a complex construction out of the values of its constituents. The theory is *intensional* because the basic semantic values are not merely composed out of this-world referents of singular terms, and extensions of predicates. While we can understand what it is for John to *greet* a unicorn in terms of the relationship between John and any individual unicorns, what it is for John to *seek* a unicorn is a more subtle matter. In an intensional type theory, an intensional transitive verb expects the *value* of a noun phrase as one input (the *seeker*), but takes the *intension* of a noun phrase (what it is that is sought) as the other. After all, one can count as seeking a unicorn without there being any unicorns there to be found.

Barbara Partee (1940–)

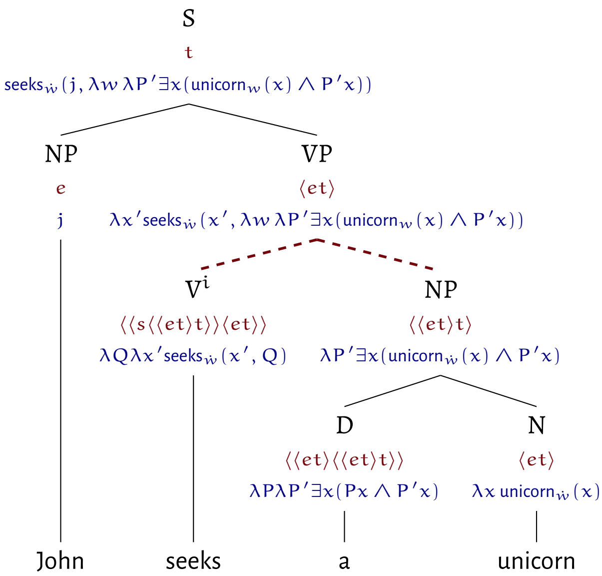
Here is a Ty_2 derivation for 'John greets a unicorn'. Notice that for it to be true that John greets a unicorn, there must *be* a unicorn that John greets.



In what follows, I will consider a simple, general framework for intensional type theory, the two-sorted type theory Ty_2 of Daniel Gallin [3]. It is a natural extension of a two-sorted first-order logic, with one sort e for *entities* and another sort s for *states*. These basic sorts are also the basic *types* of the vocabulary. *Sentences* in the language have type t (on the intended semantics, each sentence has a *truth value*). ¶ For any type α and β there is a *functional* type, $\langle\alpha\beta\rangle$ which can be thought of as functions from items of type α to items of type β. A one-place predicate, then, has type $\langle et \rangle$ since when combined with an item of type e (a singular term) as input and returns a sentence as an output. At the level of *values*, a term of type $\langle et \rangle$ can be interpreted as a function from objects to truth values. A term of type $\langle s\alpha \rangle$ is a function from *worlds* to values of type α.

Here is a Ty_2 derivation of a semantic value for the expression 'John seeks a unicorn'. Here, the NP 'a unicorn' has a complex semantic value given by applying the determiner 'A' to the unary predicate 'unicorn', where the latter is interpreted at $\dot{w}$, the world of evaluation. Since 'seeks' is an *intensional* verb, the input NP value is *intensionalised* by binding the world of evaluation. As a result, there is no requirement that John's quest be backed up by any unicorn existing at this world—or at any other.

Elizabeth Coppock and Lucas Champollion's textbook is a particularly comprehensive and accessible introduction to intensional type theory in general, and Ty_2 in particular [2].



I do not need to pursue the detail of this compositional meaning theory. It is enough to notice that it is a thoroughly denotational semantics, and as committed to a basic framework of possible worlds, entities, and truth values, and a hierarchy of grammatical and semantic types corresponding to functions between those different categories. ¶ It has been remarkably successful in structuring insightful analyses of complex compositional behaviour across the vast range of natural languages. My task is to situate inferentialist semantics with respect to this tradition.

» «

In selecting truth-conditional semantics for my target, I do not mean to ignore or to downplay the many *other* issues in semantics, concerning how it is that the basic lexical items get their semantic values, how to understand context sensitivity, ambiguity, and much else besides [6]. Those are also important issues, but they are orthogonal to my central issue concerning the relationship between proof theory and model theory, and what this might say, more generally, about the relationship between representational semantics and inferentialist semantics. ¶ My focus here is to consider the power of truth-conditional semantics, to revisit what kinds of commitments are involved in adopting it, and what it can do for us, in the light of inferentialist semantics for quantified modal logic. Exactly what kind of commitment is involved in the kind of 'world' talk that is in use by linguists.

For example, when John seeks a *unicorn*, is he seeking a mythical horned horselike creature, or is he after a privately held startup with an obscenely high valuation? Is the difference between the two kind of 'unicorn' a matter of two different lexical entries, or contextual variation concerning the one vocabulary item?

Linguists have noticed that their commitments concerning the entities and worlds of their semantic theories are not quite the same as the commitments of metaphysicians. Here is Barbara Partee:

> A non-absolutist picture seems to fit linguistic semantics better than an absolutist one, where by absolutist I mean the position that there is one single maximal set (or class, if it's too big to be a set) of possible worlds. If a philosopher could find arguments that in the best metaphysical theory there is indeed a maximal set, I suspect that would for the linguist be further confirmation that his enterprise is not metaphysics, and I would doubt that such a maximal set would ever figure in a natural language semantics. As various people have noted, *possible worlds* are really not so different in this respect from *entities*: every model-theoretic semantic theory I'm familiar with takes entities to be among the primitives—but puzzles about the identity conditions of individuals and about whether there is a maximal set of all of them are just as problematic, and it is just as questionable whether semantic theory has to depend on settling such questions. [. . .] it is the structure provided by the possible worlds theory that does the work, not the choice of particular possible worlds, if the latter makes sense at all. [8, p. 118]

How *should* we understand the semanticists commitment to a domain of possible worlds and to a domain of entities in their semantic theories? How *does* the structure provided by possible worlds theory do the work, rather than the choice of 'particular possible worlds', as Partee puts it? ¶ My own theoretical orientation when answering this question is owed to my second inspiration in these talks, Nuel Belnap. He writes:

Nuel Belnap (1930–2024)

> … in the tolerant spirit of Carnap, we believe that one is likely to want a *variety* of complementary (noncompeting) pre-semantic analyses—and most especially, a variety of pre-semantic treatments of one and the same 'language.' One does not have to 'believe in alternative logics' to repudiate the sort of absolutism that comes not from logic itself, but from narrow-gauge metaphysics or epistemology … although Carnap's beneficent influence is legendary, it seems worth repeating the lesson: There can and should be multiple useful, productive, insightful and pertinent analyses of the *same* target. Pre-semantics therefore emphasizes the usefulness of thinking in terms of a *variety* of pre-semantic systems. [1, p. 1]

Belnap's salutary advice is all-too-often ignored. Of *course* we can benefit from a number of non-competing analyses of the one target. Different accounts of the one phenomenon will foreground distinct features. One can be quite satisfied with the traditional truth-conditional semantics for for intensional type theories, while being curious about the perspective that an inferentialist analysis of the same phenomenon might afford us. ¶ I will offer an inferentialist *pre*-semantics for this truth-conditional semantic framework. The aim is to vindicate Barbara Partee's judgement, to the effect that possible worlds models serve as an incredibly useful structuring device, but it need not follow that linguists (or any of us) must fix on some antecedently given, metaphysically privileged class of possible worlds. The inferentialist pre-semantics will help us give a different account of the function of ‘possible world’ talk in truth-conditional semantics. ¶ Furthermore, this inferentialist pre-semantics will have as a corollary, an account of how the *intension* of a natural language expression (given as a function of type $\langle s\alpha \rangle$ in the truth-conditional semantics) might correspond to something that a language user might have the capacity to *grasp*, to a greater or lesser extent, bringing *representational* and *use-based* semantic accounts closer together. ¶ To bring our reflections full-circle, I will then show how the defining rules for the logical expressions count as giving the *meanings* of those expressions in just this sense.

3.2 Taking a God's Eye View

First, recall that *positions* in our inferentialist account of the kind of language use salient for modalising are structured into distinct zones. One zone distinguished as the ‘actual’ zone, in that the commitments recorded there are taken to apply to how things actually are. The other zones record claims that are endorsed as *possible*, which are considered as *alternatives*. A position might look like *this*:

$$[_{@}\ p, q, \cancel{r} \mid p, s, \cancel{r} \mid r, \cancel{p}, \cancel{q}\ ]$$

As we saw in the presentation of natural deduction proofs for this modal calculus, we could *label* the alternate zones, as a notational variant. Let’s do that:

$$[_{@}\ p, q, \cancel{r} \mid_{a} p, s, \cancel{r} \mid_{b} r, \cancel{p}, \cancel{q}\ ]$$

So, our syntax adds, to the language of formulas, a category of zone tags. We can then add to our conceptual arsenal the capacity to use these labels in further judgements. It makes sense to say that *in zone b* r *is granted*, or for short ‘$r : b$’. We can think of this as a claim which holds in *any* zone at all, and introduce it to our vocabulary with this defining rule:

$$\frac{\mathcal{C} \succ A \cdot i}{\overline{\mathcal{C} \succ A : i \cdot j}}\ {:}\ Df$$

To make the claim that A holds at i, in zone j means no more and no less than to make the claim that A in zone i. ¶ This addition to our vocabulary is conservatively extending and uniquely defining, and it has the effect of taking the ‘anthropological’ perspective of talking *about* the claims we make and the zones in which we make them. We take a ‘God’s eye view’ of the different scenarios we consider, to *globally* talk of those scenarios.

» «

Once this addition is made, it is an immediate consequence that these tagged formulas may be included as components of larger formulas. It makes sense to conjoin, disjoin, negate and use first-order quantifiers *on* these formulas, and for these formulas themselves to be tagged. As a result, we are moving toward a fully two-sorted first-order logic, with two basic sorts. Formulas can contain *entity* terms and variables (the original singular terms) and zone tags.

We can also apply modal operators to tagged formulas, but this seems hardly worth the effort, since if the tagged formula $A : i$ is present in one zone, it follows in all others, so all zone-tagged formulas are necessary, if true.

It is worth pausing at this point: adding tagged formulas looks like it comes very close to giving us singular terms referring to *possible worlds*. Such a reading is, of course, *permissible*, but it is in no way *required*. The commitments we have made thus far do not mean that we must treat our tags are referring devices. There is no logical compulsion (other than force of habit) to take these tags to denote possible worlds, where these are understood as some kind of *thing*. Our language already contains a sort (the singular terms) that ranges over things, and there is no requirement that for each tag, we should take it to be referring to any of those things.

In the language under discussion here it is consistent to be an *entity monist* (of necessity, there is only one *thing*) but to take there to be different possibilities. In limit positions modelling such commitments, there is only one entity, but there is more than one *zone*. If we take zone tags to be referring expressions, they cannot refer to *objects*, since there are not enough objects to do the job. This does not settle the issue of whether world-talk should or should not be construed as object-talk, but that the two-type framework in focus here does not force that conclusion upon us.

» «

With the addition of explicit zone tags *in formulas*, it is one small step to allow for *quantification* into zone index position. Let’s add zone variables (why not use ‘w’?), and with the obvious defining rules for universal and existential quantification into tag position, $\forall w(A : w)$ is equivalent to $\Box A$, and $\exists w(A : w)$ is equivalent to $\Diamond A$. ¶ We are, piece by piece, expanding the expressive capacities of our vocabulary, using only conservatively extending and uniquely defined logical concepts—here, exploiting the labels identifying the different zones for counterfactual supposition—and the result is a richer vocabulary, still inferentially defined, that describes *exactly* the same class of models that we motivated on inferentialist lines. We have moved from a zone-internal vocabulary making claims that differ from one zone to the next, to a zone-neutral *external* vocabulary. This shift incurs no greater ontological commitment than was made previously, but the resulting two-sorted first-order language with object quantification and zone quantification brings us closer to the vocabulary of Ty_2.

3.3 GOING UP THE LADDER

Our next step at enriching the vocabulary, inferentially, is to add λ abstraction, and to enrich the family of types, beyond having terms for zones (type s), singular terms (type e), sentences (type t), and predicates (type $\langle et\rangle$). The extension to abstraction, and to higher types is easy to motivate inferentially. We already have variables of type e, so it is easy to motivate abstraction into e position. If our language has the one-place predicates F and G, it is not hard to conceive of an object's being *both* F and G. We can form the sentence $Fx \wedge Gx$, and this is one step from forming the *complex predicate* $\lambda x(Fx \wedge Gx)$. Something has this feature if and only if it is both F and G. Predicate abstraction is governed by a straightforward defining rule:

$$\frac{\mathcal{C} \succ A^x_b \cdot i}{\mathcal{C} \succ (\lambda x A)b \cdot i}\ \lambda_{\langle et\rangle} Df$$

To apply a complex predicate to a singular term, substitute the variable bound by the abstraction operator with the term to which it is applied.

What goes for predicate abstraction can go for other types, too. For example, once we have predicates, it is natural to abstract into *predicate* position. Given a predicate variable P (of type $\langle et\rangle$), and a singular term a (type e) the expression $\lambda P\, Pa$ has type $\langle\langle et\rangle t\rangle$. It expects a predicate as an input, and returns a sentence, to the effect that this is a feature borne by the item a. We can generalise: $\lambda P(Pa \vee Pb)$ expresses the higher order property of properties had by *either* a or b. In general, given any type α, we can define the abstraction operator of type $\langle\alpha t\rangle$ generalising the defining rule above:

$$\frac{\mathcal{C} \succ A^P_B \cdot i}{\mathcal{C} \succ (\lambda P\, A)B \cdot i}\ \lambda_{\langle\alpha t\rangle} Df$$

where A has type t, and P and B have type α. ¶ The definition of abstraction at higher types is straightforward on an inferentialist view. Introducing abstraction terms is simply repackaging information that was otherwise available in the prior vocabulary. Indeed, the conservativity and uniqueness of abstraction at all types is straightforward to show, given that formulas with λ-terms can be immediately rewritten without them. ¶ The *power* of abstraction is unleashed, however, when you combine it with *quantification* at all type levels.

Matters are more subtle when we wish to define abstractions of types other than $\langle\alpha t\rangle$ which result in sentences. Gallin defines the general reduction rule for λ terms axiomatising a type-general *identity* relation, where $A \equiv B$ is of type t whenever A and B have the same type. It is more general to impose the defining rule:

$$\frac{\mathcal{C} \succ C\,A^P_B \cdot i}{\mathcal{C} \succ C\,(\lambda P\, A)B \cdot i}\ \lambda_{\langle\alpha\beta\rangle} Df$$

where A has type β, D has type $\langle\beta t\rangle$, and P and B have type α.

Given variables at each type level, it is natural to *generalise* into those variable positions, with exactly the same rules we had previously given, for the existential and universal quantifiers.

$$\frac{\mathcal{C} \succ A \cdot i}{\mathcal{C} \succ \forall P\, A \cdot i}\ \forall_\alpha Df \qquad \frac{\mathcal{C}, A \succ C \cdot i}{\mathcal{C}, \exists P\, A \succ C \cdot i}\ \exists_\alpha Df$$

Here, the variable P has type α, and is required to not occur free in the context $\mathcal{C}$ (or the tag i). ¶ These rules will have their desired effect only in the presence of an underlying principle of SUBSTITUTION, according to which the variable P is inferentially general among terms of type α. With the appropriate substitution guarantees in place, we will be able to make inference steps from $\forall P\, A$ to A^P_B, and from A^P_B to $\exists P\, A$, for any B of type α, as one would hope. ¶ As before, the addition of quantifiers at every type level is uniquely defining, and conservatively extending. However, the conservative extension result comes with significant caveats. ¶ First, the richer structure available when quantifying into higher types (even in the predicate type $\langle et\rangle$) means that quantification is *impredicative*. This means that our usual argument to conservative extension is blocked. The appropriate *left* rule for the universal quantifier has this form:

The 'usual argument' goes like this: take the invertible defining rules, and provide equivalent left-right rules which satisfy the subformula property (anything above the inference line is also present below the line). Then show that the only remaining rule in your calculus (the *Cut* rule) may be eliminated, and hence, any derivable sequent can be derived without *Cut*, and so, has a properly *analytic* derivation using only the vocabulary in the endsequent.

$$\frac{\mathcal{C}, A^P_B \succ C \cdot i}{\mathcal{C}, \forall P\, A \succ C \cdot i}\ \forall_\alpha L$$

where B is any item of type α. Since this type might involve terms of great logical complexity, it might even contain the formula $\forall P\, A$ as a *subformula*, and hence, there is absolutely no guarantee of the subformula property. (Consider $\forall P(Pa)$, where P has type $\langle et\rangle$. $\lambda x(\forall P(Pa) \land Fx)$ also has type $\langle et\rangle$ and so, substituting for the quantifier we get $\lambda x(\forall P(Pa) \land Fx)\, a$, which has $\forall P(Pa)$ as a subformula.) ¶ The rules are impredicative, as they stand. It should not be surprising if they behave differently to the first-order quantifiers.

Nonetheless, as a family, the λ and quantifier rules at each type are conservatively extending over the original two-sorted first-order modal logic. It is not too hard to see why. Any model of the two-sorted first-order modal logic has two domains, a zone domain D_s, and an entity domain D_e (which has a distinguished subset D_e^w for each zone w, of the objects that are taken, in zone w, to exist). For uniformity, we will take there to be a special domain $D_t = \{0, 1\}$ of truth values. A domain $D_{\langle\alpha\beta\rangle}$ for type $\langle\alpha\beta\rangle$ is a set of functions from D_α to D_β. *Which* functions? The easiest choice, in one sense, is to say *all* such functions. These are the domains that our variables of each type range over, and the standard recursive truth conditions for formulas apply to our enriched language, and they agree, systematically with the truth conditions for our formulas in the original two-sorted first-order language, which was *itself* a conservative extension of the prior modal language without world labels. ¶ Our original simple modal logic, with its modest ontological commitments, sits inside a richer vocabulary, which can be modelled on exactly the same structures. No positions open in the old vocabulary are closed off by appeal to higher-order reasoning. The elaborate edifice of higher types is a metalanguage that the semanticist might use to explicate the compositional features of the languages we wish to understand.

The choice of 'all' such functions is *easy*, but it does involve some serious mathematical and logical baggage. If our entity domain D_e is infinite, the domains $D_{\langle et\rangle}$, $D_{\langle\langle et\rangle t\rangle}$, $D_{\langle\langle\langle et\rangle t\rangle t\rangle}$, … climb up the hierarchy of infinite cardinals, and the logic of such 'standard' models is beholden to the commitments of the underlying set theory in which it is formulated. Nothing like this dependence obtains for first-order theories.

Although the 'standard models' of Ty_2 are forbiddingly mathematically rich, and implicated in issues in set theory that seem quite far away from concerns in natural language semantics, these models need play no more than a modelling role, used to assure us that the higher type vocabulary conservatively extends our first-order commitments. There is good evidence that the inferences that actually do the *work*, in formal linguistics, are given in the defining rules for abstraction and the quantifiers [4].

3.4 WHAT THIS MEANS MEANING MIGHT MEAN

With all this, I have given an inferentialist perspective on a rich and expressive type theory. This language is a structuring tool for the *theorists* to describe and explain compositional patterns that are implicit in how we *use* our vocabulary in our natural languages. There is no requirement that these logical constants at higher types are the kinds of things we find simple or natural as language users. They can be *implicit* in the patterns in our linguistic and conceptual behaviour.

There is no requirement that the everyday language user, competent in the use of the indefinite article 'a', has to understand that is a term of type $\langle\langle et\rangle\langle\langle et\rangle t\rangle\rangle$ or that it means $\lambda P\lambda P'\exists x(Px \land P'x)$.

The basic materials in the models for Ty_2 (the ground-level domains of entities, zones and truth values) can be given the same interpretation that the inferentialist gave for models of a first-order modal language. The inhabitants of the entity domain are the possibilia (where commitment to possibilia is explained in terms of commitment to what is possible concerning what exists), and the inhabitants of the *zone* domain simply label the different commitments concerning what is possible. The rest of the edifice built atop these domains is superstructure, classifying this base in various ways. ¶ The *functional* domains of higher types provide a space in which the linguist can describe the capacities and commitments of language users. An item of type $\langle\alpha t\rangle$ corresponds to a *distinction* between items of type α. So, a *predicate* (type $\langle et\rangle$) corresponds to a distinction between entities (between those that have the feature and those that don't).

I could spend time explaining how appeal to the domain D_t of *truth values* does not incur any special ontological commitment beyond that involved in ruling some things *in* and some things *out*, but everyone seems to have already *got* that point, and no-one is particularly exercised by ontological commitment to D_t.

An item of type $\langle\langle et\rangle t\rangle$ corresponds to a distinction between those first-level distinctions, and so on. ¶ Given any of type α, we have a corresponding *intension* of type $\langle s\alpha\rangle$. Something of this type is not something of type α alone, but a choice of something of type α, in each different zone. An item of type $\langle se\rangle$ is a selection of an object, from zone to zone. To grasp a concept of this type is to be able to apply it under counterfactual supposition.

There is much more to be said, here, but at the very least, these types begin to look like things that can have cognitive and communicative significance. We can *grasp* the meaning of some part of speech (or the corresponding concept), to a lesser or greater extent, as we are able to work with at concept more narrowly or broadly. The particular capacities displayed, of selecting items, categorising them in various ways, categorising those categorisations—and doing so not only in the here-and-now, but also under counterfactual suppositions, as we apply our reasoning and our imagination—seem like just the kinds of cognitive and communicative capacities that are important for creatures like us. ¶ The inferentialist pre-semantics discussed here at least gives a *hint* of how an intensional type theory like Ty_2 provides a useful structuring vocabulary for those capacities, and how it does so without semanticists having to defer to metaphysicians for license to reason in this way.

¶ The truth-conditional semantics is *vindicated* by its inferentialist pre-semantics.

» «

Let me return full circle. In the first two lectures, I took the defining rule for to have this form:

$$\frac{\mathcal{C}, A \succ B \cdot i}{\mathcal{C} \succ A \to B \cdot i}\;{\to}Df$$

I have made the zone tag explicit, as this rule applies in each zone.

When viewed from Ty_2, the material conditional is a lexical item of type $\langle t\langle tt\rangle\rangle$, as it conjoins two sentences. Items of type $\langle t\langle tt\rangle\rangle$ are interpreted in models of Ty_2 as binary truth functions. ¶ Now, the one binary truth function can be presented in many different ways. The lexical entry for material conditional '$\to$' is constrained only by the traditional boolean valuation condition. To grasp a concept with this interpretation, it suffices to affirm $A \to B$ when A is denied or B is affirmed, and to deny $A \to B$ when A is affirmed and B is denied. This is, of course, altogether too strong a constraint, because we might affirm A and deny B and the issue of whether $A \to B$ holds simply does not arise for us. Second, even when the issue does arise, we might find it hard to interpret. ¶ Competence with a lexical item with an interpretation like this can be acquired in many ways. One is to take $A \to B$ to be a shorthand for $\neg A \lor B$. Another would be to take it as shorthand for $\neg(A \land \neg B)$. Yet another would be to give it the defining rule ${\to}Df$ given above. Each will deliver the same pattern of verdicts, when applied correctly, and each would have the same *intension*, as the interpretation is kept fixed under subjunctive suppositions. If a language user has a concept that appears to have the same type, and the pattern of behaviour in the use of that concept robustly differs in extension from this (or it differs in *intension*, as displayed under suitable zone shifts), then this is good evidence that our interlocutor means something other than '$\to$' by this concept.

It can also be lifted to operate on intensions, taking type $\langle\langle st\rangle\langle\langle st\rangle\langle st\rangle\rangle\rangle$.

$v(A \to B) = 0$ iff $v(A) = 1$ and $v(B) = 0$.

QUICK: is $((p \to q) \to p) \to p$ true?

However extension and intension does not exhaust meaning, at least as least as far as inferential behaviour is concerned. While $A \to B$ and $\neg A \lor B$ agree on extension and on intension, they are given different definitions. They come to say the same thing by different routes. The rules introducing a concept that is given by definition are *basic* for that concept, and the distinction between basic and derived inferences is important for giving an account of proof, of understanding, and giving an account of how we might acquire concepts. So, if these issues seem important to *semantics*, we our theories should have something to say about them. Truth-conditional semantic theories, insofar as they represent meaning by extension and intension, focus on the *result* of interpretation, drawing our attention away from the *process*. ¶ Supplementing your truth-conditional semantics with an inferentialist *pre*-semantics gives you the means to keep your representational theory while giving you a new insight into its founda-

If I ask my logic students to show that $A \to B$ and $\neg A \lor B$ are equivalent, there is some work that I am asking them to *do*. I am not asking them to show that $A \to B$ and $A \to B$ are equivalent, which is a much simpler task.

tions, and a whole host of logical and conceptual tools to address these sorts of questions.

WHERE TO, FROM HERE?

Let me end with gestures to the future in three different directions.

First, this is just one effort at giving a Belnap-inspired alternative pre-semantic analysis of a familiar semantic system. I would like to see more! Take your favoured foundational theories, and see if they can be understood and analysed in different, non-competing ways. I expect that we will learn things.

Proof theoretical approaches are a natural home for *hyperintensional* distinctions. Even though A and B might be logically equivalent, a proof of A is not necessarily a proof of B. Truth-conditional semantics tends to flatten out distinctions between logically equivalent statements, while they might have different *semantic*, *epistemic*, and *metaphysical* significance. ¶ An inferentialist pre-semantics should give scope for modelling natural and motivated hyperintensional distinctions, alongside other approaches to hyperintensionality.

If the proof of the equivalence of A and B is complex and difficult to find, then of course there will be proofs of A that are not also proofs of B.

This entire investigation was motivated by first paying attention to the prevailing contours of *logic*, as a foundational discipline with its own treasures and insights. I have sought to make use of both its proof theoretic and model theoretic techniques, because doing so gives us more to work with, as philosophers, and exploring those connections with other foundational issues—in this case, in semantics—gives us new insights which, in turn, means we return to those logical techniques with greater understanding.

REFERENCES

[1] NUEL BELNAP. "Under Carnap's Lamp: Flat Pre-semantics". *Studia Logica*, 80(1):1–28, 2005.

[2] ELIZABETH COPPOCK AND LUCAS CHAMPOLLION. "Invitation to Formal Semantics". Draft available at `https://eecoppock.info/bootcamp/semantics-boot-camp.pdf`, 2024.

[3] DANIEL GALLIN. *Intensional and Higher-Order Modal Logic: With Applications to Montague Semantics*. North-Holland Pub. Co., Amsterdam; Oxford, 1975.

[4] MICHAEL HAHN AND FRANK RICHTER. "Henkin semantics for reasoning with natural language". *Journal of Language Modelling*, 3(2):513–568, Feb. 2016.

[5] DAVID LEWIS. "General semantics". *Synthese*, 22(1):18–67, 12 1970.

[6] RYAN M. NEFDT. *The Philosophy of Theoretical Linguistics: a Contemporary Outlook*. Cambridge University Press, 2024.

[7] BARBARA H. PARTEE, editor. *Montague Grammar*. Academic Press, 1976.

[8] BARBARA H. PARTEE. "Possible worlds in model-theoretic semantics: a linguistic perspective". In STURE ALLEN, editor, *Possible Worlds in Humanities, Arts, and Sciences: Proceedings of Nobel Symposium 65*, pages 93–123. Walter de Gruyter, Berlin, 1988.

[9] BARBARA H. PARTEE. *Compositionality in Formal Semantics: Selected Papers by Barbara H. Partee*. Explorations in Semantics. Wiley-Blackwell, 2004.

[10] RICHMOND THOMASON, editor. *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press, New Haven, 1974.

[11] TIMOTHY WILLIAMSON. "How did we get here from there? The transformation of analytic philosophy". *Belgrade Philosophical Annual*, 27:7–37, 2014.