Course "The Ontology of Predication"
ENS, Fall 2008/9 - Handout 9

1. Qualified Comprehension Principle

- Now that we have become a bit familiar with impossible worlds, let us make the second move towards MMM have advertised above. This is due mainly to Priest [2005]. It consists in admitting a comprehension principle for objects in unrestricted form, by *parameterising* it to worlds: **given *any* condition $\alpha[x]$, some object is described by it. However, it has its characterizing properties, not necessarily at *this* world, but at *others*** – at the worlds that make the characterization true:

(QCP) For any condition $\alpha[x]$ with free variable x, some object satisfies $\alpha[x]$ *at some world*.

- This may help with the ontology of fiction. Fictional entities are the target of intentional states and cognitive representations. Hence Priest's justification of the move:

"Cognitive agents represent the world to themselves in certain ways. These may not, in fact, be accurate representations of this world, but they may, none the less, be accurate representations of a *different* world. For example, if I imagine Sherlock Holmes, I represent the situation much as Victorian London (so, in particular, for example, there are no airplanes); but where there is a detective that lives in Baker St, and so on. The way I represent the world to be is not an accurate representation of our world. But our world could have been like that; there *is* a world that is like that ."(Priest [2005]: 84).

To be more precise, there are *many* such worlds, since representations are incomplete with respect to many details.

- Objects picked by a description, therefore, may *always* satisfy the relevant condition, $\alpha[x]$. According to Priest, we don't need to isolate a subset of "nuclear" or "characterizing" properties. Also {goldenness, mountainhood, existence} works fine now, for we need not assume that an object so characterized, that is, an existent golden mountain, has its characterizing properties at the actual world. As far as we know, no golden mountain inhabits the actual world, but golden mountains are available at the worlds at which the stories we can tell on existent golden mountains hold.

- A Meinongian (possible and impossible) worlds semantics can be framed as a constant **domain structure**. This is quite natural for Meinongians. People assume variable domains in ordinary modal semantics to account for the idea that different things may exist at different worlds. But in a Meinongian framework the domain of each world is simply the *totality* of objects: that some object o exists at world w_1, but not at world w_2, is accounted for by having o satisfy the existence predicate $E!$ at w_1 and not at w_2, and all the epicycles of Kripke variable domain semantics are left behind. *Simplex sigillum veri*.

- However, this entails a kind of ***realist* attitude towards nonexistent objects**: they are out there independently of our intentional activity. If one has an anti-realist attitude towards fictional objects, maintaining that they are actually **created by the author's say-so**, one might need to complicate things a bit and switch to a variable domain structure. I'll come back to this later.

- We can stick to the idea that fictional objects also have the properties *entailed* by those they are explicitly characterized as having, given some suitable notion of entailment. Since in some forms of Meinongianism (such as Routley's [1980], [1982]) to exist is to have causal properties, and/or to be located in space and time, **one may take some properties or relations involving causal features to be existence-entailing** in one or more arguments (which properties or relations are existence-entailing may be a matter of ontological debate).

- This idea, as far as we know, is due mainly to Linsky and Zalta [1994] – actually, they talk of *concreteness*-entailing properties, but it is easy to provide a rough-and-ready translation from their modal theory to MMM.

- For instance, if *x* kisses *y*, then both *x* and *y* must exist, and if *x* thinks about *y, x* must exist although *y* need not (at least at possible worlds – one may probably fantasize on impossible nonexistent kissers and thinkers). This accounts for the intuitive idea that Sherlock Holmes, being a nonexistent object *at* this world, cannot actually kiss anyone here, nor can he entertain any thought here, although he can be thought of (by Doyle, or by any reader of the Doyle stories). However, if some Doyle novel features Holmes kissing Watson, then Holmes does kiss Watson at the worlds at which the characterization is true and, at those worlds (or at least, at the possible ones among them), Holmes does exist.

2. Replies to the Russellian objections

- By admitting impossible worlds, we can account for inconsistent objects, such as Quine's round square cupola of Berkeley College, that is, an object which, given that if something is a square then that thing is not round, is round and not round, $Rx \wedge \neg Rx$: we simply admit inconsistent worlds that realize contradictions.

- Notice that we do not need to admit *true* contradictions, or even possibly true ones. We can stick to the intuition that Quine's cupola is an impossible, therefore not possibly existing, object: if we tell a story about some rock climber who climbs on the round square cupola of Berkeley College, supposing that "*x* climbs on *y*" is a predicate which is existence-entailing in both arguments, then the round square is round, square, and existent – not at this world, nor at any possible world, but at those impossible worlds described by our story, and which realize the characterization "(existing) round square cupola of Berkeley College".

- Also the second Russellian objection can be answered straightforwardly: we can consider the condition "x is an existent golden mountain", and the object so characterized does have all its characterizing properties according to the QCP. What cannot be guaranteed *a priori* is that the object has those properties at the *actual* world; therefore, the QCP does not allow one to prove the existence of anything whatsoever.

3. Realism and anti-realism on nonexistents

- Various theories of fictional entities treat fictional entities as ***existent*** **and** ***abstract*** **objects**. Some versions claim that such objects are literally *created* by the say-so of the authors. For instance, according to Amie Thomasson [1999] fictional entities are artefacts, just like chairs and tables but, unlike chairs and tables, are abstract: Doyle created Sherlock Holmes; Sherlock Holmes did not exist before Doyle created him, and would not have existed, had Doyle not created him.

- This will not do in the Meinongian framework, if by "to create" one means "to bring into existence": for how could Doyle bring Holmes into existence, given that Holmes is a nonexistent object? After all,

(1) Holmes doesn't exist

is, intuitively, simply *true* (at this world, to be sure).

- Exactly *what* Did Doyle do, then, according to the Meinongian? I like the following explanation. Doyle provided a long and articulated description, contained in his novels. The description characterized the object which was baptized "Sherlock Holmes" by Doyle, and Sherlock Holmes has his characterizing properties at the worlds that make Doyle's description true. At those worlds, Sherlock Holmes exists, too (at least, the possible ones), for he certainly does many existence-entailing things.

- According to this kind of Meinongianism, which we may dub *realist*, Sherlock Holmes is just an object in the domain of the totality of objects; at some worlds, including the actual one, he does not exist. At some others, he does. Doyle was the first to call the object at issue "Holmes" – that is, to give to the object the name by means of which we refer to it. This counts as an extended, non-causal baptism. In Zalta's words:

"Instead of pointing and mentioning the relevant name, the author *tells a story.* I suggest that the act of storytelling is a kind of extended baptism, and is a speech act more similar to definition than to assertion. A story is required to baptize a nonexistent object as a fictional character. The author doesn't really establish or determine the reference of the name or names used, except in a derivative sense." (Zalta E. [2003], "Referring to Fictional Characters", *Dialectica* 57: 243-54.

- Once a fictional object has been so baptized, other speakers may pick the same fictional object and publicly refer to it. If we stick to the Meinongian idea that nonexistent objects are *given* via descriptions, this might entail that Doyle was continuously changing the subject as he kept on writing stories on Holmes: by providing different, more detailed characterizations, he was speaking of continuously different objects.

- However, one might resist this counterintuitive multiplication of objects and stick to our intuition that the Holmes which appears in *The Hound of the Baskervilles* it *the same* as the Holmes which figures in *The Sign of the Four*. By writing more and more novels on Holmes, Doyle obviously *restricted* the set of worlds which made the relevant description true: the more details he gave us on Holmes, the fewer worlds satisfied the relevant description, but it was still the same Holmes that was characterized by the novels.

- Another option for Meinongianism is to go *anti*-realist and claim that **fictional objects are actually created by the author's say-so, but in a non-existence-conferring sense**: maybe one can create nonexistents! This can be achieved by switching back to *variable* domains, and providing the following account of "creation". We may allow that the domain of objects at a world is not fixed once and for all. At least a part of the domain of the objects at a world *w supervenes* on the properties, features and activities of the objects that *exist at* that world *w*.

- Now we can claim that Sherlock Holmes is in the domain of the objects of the actual world, though he keeps being a nonexistent object here, because Doyle actually created him: it is because of Doyle's storytelling activity that Holmes is available at this world as something we can refer to and as an object in the domain of the (Meinongian) quantifiers. In this sense, Holmes *supervenes* on Doyle's intentional activities: these activities have resulted in the expansion of the domain of objects.

4. Formalities

- After presenting the two pillars of MMM informally, I am now going to make things more precise via a little bit of formal machinery. The following formal semantics comes basically from Priest's (2005) account, but with a few adjustments.

- First, we take the usual first-order language with a set of individual constants, *n*-place predicates (with a distinguished one-place predicate, *E!*), individual variables, the standard connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, two quantifiers, Λ and Σ, and the usual rules for well-formedness.

- Next, we add to the standard language an **intensional *representation* operator, Ψ:** if α is any well-formed formula, '$\Psi\alpha$' is a well-formed formula, to be read: "It is represented (in such-and-such fiction) that α".

- The primitive, Meinongian quantifiers Λ and Σ (to be read as "for all" and "for some"), as we know, are taken as existentially neutral: one can quantify on, and talk in general of, nonexistents. Existence is taken as a perfectly normal first-order property expressed by the predicate *E!*, employed to provide explicit existential commitment and to define the existentially loaded quantifiers. "All existing things are such that…" is:

$\forall x\alpha[x] =_{df} \Lambda x(E!x \rightarrow \alpha[x])$;

and "There exists something such that..." is:

$\exists x\alpha[x] =_{df} \Sigma x(E!x \wedge \alpha[x])$.

- An interpretation of the language is a sextuple <P, I, @, D, *R*, *v*>, where P is the set of possible worlds and I is the set of impossible worlds. P and I disjoint and W = P $\cup$ I is the totality of worlds. @ $\in$ P is the distinguished actual world. (This would not be needed if all we wanted from the semantics were accounts of validity and logical consequence; but @ is to do other jobs). D is a non-empty set of objects, *R* is a binary relation on the whole set W. The semantics is framed as a constant domain structure to keep things simple. Finally, *v* is the interpretation function assigning denotations to the non-logical symbols as follows:

If *c* is an individual constant, then $v(c) \in$ D;
If *P* is an *n*-place predicate and $w \in$ W, then $v(P, w)$ is a pair $<v+(P, w), v-(P, w)>$, with $v+(P, w) \subseteq D^n$, $v-(P, w) \subseteq D^n$.

If *P* is an *n*-place predicate, *v* assigns to it an extension $v+(P, w)$ and an anti-extension $v-(P, w)$ relative to worlds. Intuitively, the extension of *P* at *w* is the set of *n*-tuples of which *P* is true there, and the anti-extension is the set of *n*-tuples of which it is false.

- For possible worlds, we require that the two be exclusive and exhaustive for any *P* – let us call this the **Classicality Condition**:

(CC) If $w \in$ P, then $v+(P, w) \cap v-(P, w) = \varnothing$
$v+(P, w) \cup v-(P, w) = D^n$

This reflects the idea that possible worlds must be consistent and maximal: for any predicate *P*, if *w* is a possible world, *P* is either true or false of the relevant object(s) at *w*, but not both. Nevertheless, truth and falsity conditions are spelt separately (in a way familiar from various kinds of non-bivalent semantics), for things may go differently at impossible worlds.

- In order to evaluate quantified sentences we need assignments of denotations to the variables, the usual way: if *a* is an assignment (a map from the variables to D), then v_a is the suitably parameterized interpretation:

If *c* is an individual constant, then $v_a(c) = v(c)$;
If *x* is a variable, then $v_a(x) = a(x)$.

- Next, we read "$w \vDash^+_a \alpha$" as "α is true at world *w* with respect to assignment *a*", and "$w \vDash^-_a \alpha$" as "α is false at world *w* with respect to assignment *a*" (I will omit the subscript when dealing with closed formulas, for which different assignments, as usual, make no difference). Then we have, for atomic formulas:

$w \vDash^+_a Pt_1 \ldots t_n$ iff $< v_a(t_1), \ldots, v_a(t_n)> \in v+(P, w)$
$w \vDash^-_a Pt_1 \ldots t_n$ iff $< v_a(t_1), \ldots, v_a(t_n)> \in v-(P, w)$

- For negation:

$w \vDash^+_a \neg\alpha$ iff $w \vDash^-_a \alpha$
$w \vDash^-_a \neg\alpha$ iff $w \vDash^+_a \alpha$

- Since extensions and anti-extensions are exclusive and exhaustive at possible worlds, if $w \in$ P we have that $w \vDash^+_a \neg\alpha$ iff it is not the case that $w \vDash^+_a \alpha$; so thanks to the CC at possible worlds negation works "homophonically", the classical way. And since @ $\in$ P, that is, the actual world is possible, there are no truth value 'gluts' or 'gaps' at it (sentences that are both true and false, or neither): truth *simpliciter*, truth at @, behaves in an orthodox way with respect to negation.

- Conjunction, disjunction, the quantifiers and the (strict) conditional get the usual clauses at *possible* worlds. For all $w \in$ P,

$w \vDash^+_a \alpha \wedge \beta$ iff $w \vDash^+_a \alpha$ and $w \vDash^+_a \beta$
$w \vDash^-_a \alpha \wedge \beta$ iff $w \vDash^-_a \alpha$ or $w \vDash^-_a \beta$

$w \vDash^+_a \alpha \vee \beta$ iff $w \vDash^+_a \alpha$ or $w \vDash^+_a \beta$
$w \vDash^-_a \alpha \vee \beta$ iff $w \vDash^-_a \alpha$ and $w \vDash^-_a \beta$

$w \vDash^+_a \alpha \rightarrow \beta$ iff for all worlds w_1 (if $w_1 \vDash^+_a \alpha$ then $w_1 \vDash^+_a \beta$).
$w \vDash^-_a \alpha \rightarrow \beta$ iff for some world w_1 ($w_1 \vDash^+_a \alpha$ and $w_1 \vDash^-_a \beta$).

$w \vDash^+_a \Lambda x\alpha$ iff for all d $\in$ D, $w \vDash^+_{a(x/d)} \alpha$
$w \vDash^-_a \Lambda x\alpha$ iff for some d $\in$ D, $w \vDash^-_{a(x/d)} \alpha$

$w \vDash^+_a \Sigma x\alpha$ iff for some d $\in$ D, $w \vDash^+_{a(x/d)} \alpha$
$w \vDash^-_a \Sigma x\alpha$ iff for all d $\in$ D, $w \vDash^-_{a(x/d)} \alpha$

Where "$a(x/\mathrm{d})$" denotes the assignment which is the same as a, except that it assigns to x the value d.

- This means that the standard logical operators behave in an orthodox fashion (except for the fact that they are given truth and falsity conditions separately) at possible worlds. Instead, they behave anarchically at impossible worlds: here the **truth values are not determined recursively**. The interpretation function *v* treats formulas formed by means of them as ***atomic***, assigning extensions and anti-extensions directly. The idea was introduced by Rantala in order to provide a semantics for intentional operators making logical omniscience fail. So at points in I the truth values of conjunctions, disjunctions, and quantifiers, are not assigned recursively, but directly determined by *v*.

- Now, the existence-entailing features of some properties mentioned above can be accounted for in the machinery by adding some formal constraints. As advertised there, we will assume that, if an *n*-place predicate *P* is existence-entailing in its *i*th place, it is such at all possible worlds:

If $w \in \mathrm{P}$, then if $<\mathrm{d}_1, \ldots, \mathrm{d}_i, \ldots, \mathrm{d}_n> \in v+(P, w)$, then $\mathrm{d}_i \in v+(E!, w)$.

Priest (2005, 60) does not take a stand on this; but it seems to me that existence-entailments may be regarded as something similar to a kind of meaning postulates, fixing the semantics of some predicates and, in particular, their internal connections to the predicate "exists". Meaning postulates are usually taken as necessary truths, holding at all possible worlds. So if Santa Claus thinks about Pegasus at *w*, then Santa Claus exists at *w* though Pegasus need not; if Holmes kicks Moriarty at *w*, then both exist at *w*; and this holds whenever *w* is a possible world. What happens at impossible worlds, of course, is another story; nonexistent things, for instance, may think or kick at some impossible world.

- Finally, **worlds are allowed to access impossible worlds when the truth conditions for the representation operator Ψ are at issue**. For any $w, w_1 \in \mathrm{W}$:

$w \vDash^{+}_{a} \Psi\alpha$ iff for all $w_1 \in \mathrm{W}$ such that $wR\, w_1$, $w_1 \vDash^{+}_{a} \alpha$
$w \vDash^{-}_{a} \Psi\alpha$ iff for some $w_1 \in \mathrm{W}$ such that wRw_1, $w_1 \vDash^{-}_{a} \alpha$

- These are the key clauses, and their intuitive explanation goes as follows. The semantics of the representation operator is just a restatement of the usual (restricted) binary accessibility semantics for modal operators of ordinary modal logics. The idea is that wRw_1, that is, there is a ***representational accessibility*** (*R*-accessibility) from *w* to w_1, if and only if, **at w_1, things are as they are represented to be** (within such-and-such a story, tale, fictional work, etc.) **at *w***; or, equivalently: w_1 is *R*-accessible from *w*, just in case w_1 realizes the way things are characterized, or described – that is, represented – to be (within such-and-such a story) at *w*. For instance, take the worst nightmares one may have (at @); then a world *w* such that @*Rw* is a nasty world at which those worst nightmares come true.

- Now the reason why impossible worlds are accessible when evaluating formulas of the form '$\Psi\alpha$', is precisely that we can form inconsistent descriptions more or less of any kind: $\alpha[x]$ may be "*x* is a round square", "*x* is round and not round", "*x* is red or blue but it is not the case that *x* is red and it is not the case that *x* is blue", etc. We may represent (imagine, think of, tell stories about, etc.) objects having any (set of) properties, however extravagant. And our QCP tells us that for any condition $\alpha[x]$, there will be worlds at which some object satisfies $\alpha[x]$.

- Next, we need a definition of logical validity and logical consequence. If S is a set of formulas:

$\mathrm{S} \vDash \alpha$ iff for every interpretation <P, I, @, D, *R*, *v*>, and assignment *a*, if $@ \vDash^{+}_{a} \beta$ for all $\beta \in \mathrm{S}$, then $@ \vDash^{+}_{a} \alpha$.

- For logical validity, we have just that $\vDash \alpha$ is $\varnothing \vDash \alpha$. So logical consequence is truth preservation at the base world, @, in all interpretations, and logical validity is truth at @ in all interpretations. In fact, one may also define logical consequence as truth preservation at all possible worlds in all interpretations: the semantics sketched contains nothing to differentiate @ from any other world *w* in this respect, insofar as *w* $\in$ P, that is, *w* is possible (we are into what follows logically from what, therefore, is what follows from what at the worlds where logic is *not* different).

- The representation operator, despite having its semantics provided by an accessibility relation involving such anarchic guys as impossible worlds, must somehow be regimented. How to spell out the regimentation in detail is a very subtle issue (as shown by the classic discussion in Lewis [1978], "Truth in Fiction", *American Philosophical Quarterly* 15: 37-46) and, luckily, one we need not get into here. I will limit myself to some general remarks.

5. Regimenting the representation operator

- Fictional representations induced by fictional discourse must be **closed under some notion of logical consequence**. This is shown by the fact that we can argue about how things are in the relevant situation(s) and, in the process of arguing, we draw inferences. Some reasoning on what follows from what within a fiction is correct, some is not, also when the fiction at issue involves inconsistencies, even intentional ones.

- The premises of such reasoning sometimes are not explicitly stated within the fiction, but are imported as default assumptions from actuality, background information, shared knowledge, etc. For instance, if within a Doyle story it is said that Holmes was in London at a certain time, we can infer that he couldn't have been in Australia the day before, even when this isn't explicitly stated in the story. This is correct because Holmes' London is Victorian London; lacking information to the contrary, we can therefore import into our reasoning the actual truth that during the Victorian age there were no aeroplanes or other means that could allow one to move from Australia to London in a day. The set of worlds accessible by means of the representational relation must be somehow be appropriately constrained on the basis of the background information against which we read the fiction.

- How does the QCP square with all this? In naïve and nuclear Meinongianism any condition (or any nuclear condition) characterized some object, but with no proviso on worlds at which the object had the relevant properties. In particular, for any (nuclear) condition $\alpha[x]$, calling *o* an object characterized by the condition, we could have (using our newly introduced notation) @ $\vDash^{+} \alpha[o]$, and therefore, @ $\vDash^{+} \Sigma x\alpha[x]$. In our modally qualified framework, if an object *o* is characterized by $\alpha[x]$, then in general we have that @ $\vDash^{+} \Psi\alpha[o]$. Given the semantics of the representation operator, this entails only that, for all worlds *w* that realize the way things are described as being at @, that is, for all *w* such that @*Rw*, $w \vDash^{+} \alpha[o]$.

- As prescribed by the QCP, **objects do have the features they are described as having, but not *at* @: they have them only at the *R*-accessible worlds** *w* (which may be possible or impossible ones) **that realize such descriptions**. Sherlock Holmes is described (at @), by Doyle and by the Sherlock Holmes fans and Doyle readers engaged in internal fictional discourse, as being a detective and (let us assume) kicking Moriarty. He does not have these properties at @, though – in particular, because being a detective and kicking someone appear to be existence-entailing properties, whereas, at @, Holmes does not exist. He has them, though, at the worlds that realize the Doyle stories.