Ontological Commitment and Aboutness

“How thing are”, graduate course, spring term 2011
Philipp Keller, philipp.keller@unige.ch

Ontological commitment

To say that the name $x$ denotes a given object $a$ is the same as to stipulate that the object $a$ [...] satisfies a sentential function of a particular type. In colloquial language it would be a function which consists of three parts in the following order: a variable, the word 'is' and the given name $x$.

(Tarski 1956: 194)

A first stab at ontological commitment:

(OC) A truthbearer $p$ is ontologically committed to an entity $e$ iff $e$ has to exist for $p$ to be true.

Unfortunately, things are not that simple: whatever the truthbearer in question is (the sentence “$p$”, or the thought or proposition it expresses), it certainly has to exist in order to bear truth. We do not, however, want to say that we are committed to what bears the truth of what we are saying. Intuitively, we want to be committed to what these truths are about, not just to what these truths themselves are. Can we therefore say that we are committed to those entities the existence of which is entailed by the truth of what we say, where entailment of existence is taken to be entailment of a statement to that effect?

(OC') A truthbearer $p$ is ontologically committed to an entity $e$ iff $p$ entails “$e$ exists”.

Can we then say that the entities we are committed to are those that make true what we are saying? That which makes some truthbearer true is often called its “truth-maker” (Mulligan et al. 1984: 287, following Russell’s use of the verb), its “ontological ground” (Armstrong 1991: 190, following Bergmann), or its “existential grounding” (Lewis 2001: 279). Truthmaking is then often characterised as the converse of ontological commitment:

(TME) An entity $e$ is a truthmaker for a truthbearer $p$ iff “$e$ exists” entails $p$.

According to this analysis, truthmaking and ontological commitment are converse relations, and both are spelt out in terms of entailment. As intuitive as this first analysis may seem, I think it is deeply mistaken and so for a number of reasons:

1. entailment is not cross-categorical
2. (OC) is analytic, (TME) is synthetic.
3. aboutness is left out of the picture

According to Quine’s criterion, what a sentence commits us to is determined by what its variables range over and hence by how it is formalised. The connection between formalisation and commitment, however, may be in one of two directions:

(i) Formalisation uncovers ontological commitment: the commitment of a sentence is determined by its logical form.

(ii) Formalisation is constrained by ontological commitment: sentences have an 'ontological form' that a correct formalisation has to confirm.

Quine privileges the first direction, while I think that the second is more important: we cannot simply read of our commitments from the logical forms of the sentences we accept — rather, we are guided in our formalisation efforts themselves by a sense of which commitments are acceptable. But how, one may ask, is it even possible that formalisation is constrained by, rather than constrains, ontological commitment? After all we do not, one may think, have an independent grasp of 'ontological', as opposed to logical form. But perhaps we do.

Referential indeterminacy:

English general and singular terms, identity, quantification, and the whole bag of ontological tricks may be correlated with elements of the native language in any of various mutually incompatible ways, each compatible with all possible linguistic data, and none preferable to another save as favored by a rationalization of the native language that is simple and natural to us. (Quine 1957: 4-5)

If we identify our knowledge of the world with knowledge of only its structure (in so far it is describable by purely general quantified Ramsey-sentences) — if "of the external world we know its structure and nothing more" (Newman 1928: 142)—, then our theories admit of different, incompatible but equally 'good' interpretations even if the domain of their interpretation is fixed. The problem arises, as Newman realises, from Russell's claim that we know the structure $W$ of the world in virtue of a relation $R$ about which we only know that exists — the problem then is that "any collection of things can be organised so as to have the structure $W$, provided there are the right number of them" (Newman 1928: 144):

Thus, on this [Russell's 1927] view, only cardinality questions are open to discovery! Every other claim about the world that can be known at all can be known a priori as a logical consequence of the existence of a set of $\alpha$-many objects. (Demopoulos and Friedman 1985: 627)

We must come up with additional constraints on admissible interpretations. Putnam considers an appeal to a causal theory of reference by the metaphysical realist: but this would be, he answers, just the addition of more theory to $T_1$ (Putnam 1978: 126) and therefore not help, as long as "reference" is not "glued to one definite relation with metaphysical glue" (Putnam 1980: 477). I think that it is this 'just more theory' reply to realists may plausibly be rejected:

Constraint $C$ is to be imposed by accepting $C$-theory, according to Putnam. But $C$-theory is just more theory, more grist for the mill; and more theory will go the way of all theory. To which I reply: $C$ is not to be imposed just by accepting $C$-theory. That is a misunderstanding of what $C$ is. The constraint is not that an intended interpretation must somehow make our account of $C$ come out true. The constraint is that an intended interpretation must conform to $C$ itself. (Lewis 1984: 62)

The intrinsic structure of the world, according to Lewis' proposal building on Armstrong's scientific realism about universals, comes from more or less natural properties, the 'real joints' in nature:

---

2 They add: "(The relevant theorem from set theory or second-order logic is the proposition that every set $A$ determines a full structure, that is, one which contains every relation (in extension) of every arity on $A$; such a structure forms the basis for a (standard) model for the language of second- (or higher-) order logic.)" (Demopoulos and Friedman 1985: 628)
Among all the countless things and classes that there are, most are miscellaneous, gerrymandered, ill-demarcated. Only an elite minority are carved at the joints, so that their boundaries are established by objective sameness and difference in nature. Only these elite things and classes are eligible to serve as referents. (Lewis 1984: 65)

Just structure, however, is not enough: the structure of the world must be specific, the elite properties and relations sparse (Merrill 1980: 77). These elite properties and relations will then act as ‘reference-magnets’, constrain admissible interpretations and make for “an objective inequalitarianism of classifications” (Lewis 1984: 67).

Putnam’s paradox shows that Quine’s insistence on the domain of quantification puts the cart before the horse: In advance of a criterion of ontological commitment, the very idea of choosing a domain of quantification to interpret our predicates in and have our variables ranging over does not make any sense. We cannot pick the domain over which the variables in a theory range before we know what its ontological commitments are and we cannot know what the commitments are before we have a domain in which to interpret the theory. Framed this way, I think, it can be easily seen that Lewis’ response to Putnam’s (and Quine’s) trivialisation problems is on the right track: we have to take the idea seriously that domains are not chosen by us, but given independently of and prior to our interpreting our theories. Putnam’s model-theoretic argument is based on the assumption that “we [i.e. our intentions] interpret our languages or nothing does” (Putnam 1980: 482) and can therefore be taken as reductio of that assumption. As Lewis (1984) argues, some referents, things and classes of things, are more eligible than others — intended interpretations of our language must maximise eligibility of referents overall (Lewis 1984: 65). Their eligibility does not depend on whether or how we conceive of them; it is itself a natural (and hence, highly eligible) feature of the things we quantify over.

**Truth as a formal concept**

The truth predicate has the disquotationalist feature enshrined in Tarski’s famous ‘convention T’:

\[(T) \quad \text{“}p\text{” is true iff } p\text{.}\]

But (T), while it is trivially true, still raises a lot of questions:

1. The truth of (T), for some substitution instance of “p”, imposes rather tight constraints on this substitution instance: “p” has to be a truth-evaluable, non-indexical and non-ambiguous sentence of English. How can we capture these restrictions on the implicit universal quantifier in (T)?

2. The truth of (T), for some substitution instance of “p”, also imposes constraints on ““p””, the name of the substitution instance. If instances of (T) are supposed not just to be true, but to be a priori or logical truths, then not just any name will do.

3. It is in virtue of (T) that the truth-predicate fulfills its primary purpose of capturing generalities. But the instances of (T) by themselves cannot do that; what else do we need?

4. What does (T) tell us about truth – the concept, the predicate and the property of truth?

5. If (T) tells us something about truth, does it tell us everything about it?

To say that some statement that … is true, is to say that … This much is beyond dispute. But if our initial platitude is to inform us about the concept of truth, much care is needed in its precisification. “The statement that …”, e.g., should not be explicated by “the statement having the truth-condition that …” or “the statement that is true iff …”, for this would make our account depend on a previously understood notion of truth.
How else, however, are we to understand the that-clause? One way is to adopt what Hartry Field calls a “quotational reading”: to say that $S$ (as we actually use it) means that neutrinos have mass, where $S$ is a sentence of our language, is simply to say that $S$ means the same as our sentence “Neutrinos have mass” (as we actually use it). Meaning-the-same-as is cognitive equivalence: the rules of our language or the norms we accept license the immediate inference of either of them from the other. To extend this account of “$S$ means that $p$” to other languages, we have to give an account of the factors involved in good translation without relying on a prior account of truth-conditions. We may want to restate

\[(EP)\] To say that the statement that ... is true is equivalent to saying that ...

as

\[(EP')\] For each English sentence that one substitutes for “$p$”, “the proposition that $p$ is true” is equivalent to $p$.

If the equivalence is at least modal, \( (EP') \) entails (instances of) the schema:

\[(EP')\] Necessarily, the proposition that $p$ is true if and only if $p$.

To get

\[(T')\] The sentence “$p$” expresses a true proposition iff $p$.

we need another premise, namely

\[(EXP)\] The sentence “$p$” expresses the proposition that $p$.

By the introduction of the technical term “proposition”, the precisification of the original, platitudinous equivalence \( (EP) \) leaves us with a claim about the logic of “…is true” \((EP')\) and a claim about how sentences relate to propositions \((EXP)\). This may seem innocuous enough: do we not have an independent grasp of what propositions are, namely the objects of so-called ‘propositional attitudes’ like believing, hoping and fearing? Even if this were true, however,\(^3\) there is still some mismatch between what is hoped and believed and what is expressed by sentences and quantified over in \((EXP)\).

Both \( (EP) \) and \( (EP') \) are a theory of truth only for propositions given in the form the proposition that $p$, where an English sentence is substituted for ‘$p$’. We not only do yet not have an account of the form

\[(T\leftrightarrow)\] For any proposition $x$, $x$ is true iff $C(x)$

not even one of the form

\[(T\ast)\] For any proposition $x$, if $x$ is expressible in English then $x$ is true iff $C(x)$

but just a collection of claims of the form:

\[(T\ast)\] For any proposition $x$, if $x$ is the proposition that ... then $x$ is true iff ...

\(^3\)As a theory about such attitudes, propositionalism carries a lot of commitments: that to stand in such an attitude is to stand in a relation to an object, for example, that such objects are expressed by (rather than e.g. identical to) linguistic items, that they have the 'completeness' required of them to be autonomously truth-evaluable, that different bearers of attitudes are related to the same thing iff the component clauses of true ascriptions of these attitudes express the same thing and so on.
This means that many accounts of the relata of so-called 'propositional attitudes' will not help us to understand the role propositions play in (T').

This is an example of a broader tension: the cash-value of talk of propositions in (T') is to 'stabilise' the application of the truth-predicate, to identify that in virtue of which the truth-predicate is correctly applied to items as diverse as statements, beliefs and sentences. This is captured in (EXP) and the theoretical role "proposition" plays in it is aptly characterised by what Schiffer (2003: 12-15) calls the "face-value theory about propositions":

(i) Propositions are the referents of singular terms.
(ii) Propositions are abstract.
(iii) Propositions are mind- and language-independent entities.
(iv) Propositions have truth-conditions and they have them both essentially and absolutely.

These characteristics are in tension with the role of propositions as the objects (and not just the contents) of beliefs, however. If believing that Zed is dead is standing in a relation to the referent of "that Zed is dead" (i) and this referent is both essentially and absolutely such that it is true iff Zed is dead (iv), then we may wonder how the belief could have turned from false to true on the occasion of the historical event that was Zed's death. Nothing, it seems, about the proposition changed at that time, hence nothing about the 'propositional part' of the belief. And still it went from false to true.

Apart from the unclarity of "proposition", there are, I think, two other reasons to stay with the non-committal (T) rather than (T'). The first is methodological, the second logical.

The basic puzzle: intrinsic, but relational representation

Suppose I produce the following shape:

Zed is dead.

By fixing your gaze on it, you learn that Zed is dead. How is this possible? How is it that by fixing our gaze on some marks of chalk we learn something about people and things in no (obvious) connection to us? This, I take it, is the primordial philosophical question about representation.

It is ordinarily assumed that the representation question divides into two sub-questions:

(i) how is it possible that the marks of chalk mean what they do?
(ii) how is it possible that by understanding what they mean, we learn something about the distant past?

It is only if you accept this sub-division, that the notion of a proposition starts playing an (apparently) explanatory role:

(i*) the marks of chalk express the proposition that p;
(ii*) the proposition that p is true iff (or: in virtue of the fact that) Zed is dead.

The notion of a proposition is a means to divide our initial question into these two sub-question. Deflationism about truth then becomes an attractive position: it becomes almost irresistibly plausible to say that (ii*) holds just because the proposition in question is what it is — saying that it is true that Zed is dead is just saying that Zed is dead. (i*), though it encodes a contingent fact about language, seems utterly trivial — all you need to see its truth is some competence of English. Our initial puzzlement thus evaporates into two trivialities and we wonder what made us think our question was interesting in the first place.

4 (EP') and (EP*) do not tell us, for example, what it is for propositions described as sets of possible worlds or as an equivalence class of sentences to be true.
Just a little below the surface of the two trivialities lie two deep mysteries, which seem completely intractable if approached in these terms:

(i) How is it possible that our activities as language-using intentional agents bestow mind-independent and abstract entities with powers of representation?
(ii) How is it possible that relations of aboutness and of truthmaking hold between these abstract entities and things in the world?

I think the sub-division of our original question was a mistake and that we should cut the middle-man out: propositions, understood as abstract objects of belief, do not serve a useful explanatory function. We should not construe belief in terms of content, but rather explain content in terms of belief.

Many things may be said to have content, but most of them do so indirectly: they have content in virtue, for example, of having been produced in a certain way or with certain intentions, or of standing in some relation to other things that have content. The most important such relation is that of some things expressing other things. It is in virtue of expressing my beliefs that my utterances have content, and – subject to certain constraints – the beliefs expressed determine what content they have.

Most contentful things thus have their content extrinsically: they mean what they do in virtue of other things having a certain content. At some point, however, the bucket must stop: if there are any representational properties at all, some things must have them intrinsically. Because they are representational, however, they are relational even when exemplified intrinsically: they represent something other than themselves, creating a relation between their bearer and the things they make their bearer be about.

According to what Chisholm (1952: 56) calls “Brentano’s Thesis” – that intentionality is the mark of the mental – representational properties are extrinsic, but non-relational. They are extrinsic, because they are signs, but non-relational, because they are characterised by “intentional inexistence”: psychological states may exist even in the absence of what they are about. Perhaps surprisingly, intrinsic but relational representation and extrinsic but non-relational intentionality are compatible.

Representational properties like meaning that, representing a to be F or thinking of a as F are intrinsically exemplified by some thing x iff x exemplifies the property independently of how matters stand with respect to other things than x – no further properties have to be exemplified for other things for my thought, e.g., to represent a to be F. That some representational properties are exemplified intrinsically by some things follows from the following argument:

(i) Some things have representational properties.
(ii) If something exemplifies a representational property extrinsically, it does so in virtue of a relation that bestows it with this representational property.
(iii) In order for something to bestow something else with a representational property, the first thing needs to exemplify this representational property itself.
(iv) The transmission of representational powers can neither go on forever, nor go in circle: it must be started by something.
(v) A thing that has a representational property that is not bestowed upon it by something else exemplifies it intrinsically.

Even when they are exemplified intrinsically, however, representational properties are still relational: they connect their bearers to the things they are about. If my thought, for example, represents a to be F, it stands in the relation of aboutness to a and in the predication relation to the universal F. It is in virtue of these relations that my thought can stand in for a’s being F, and be in some sense further to be specified a substitute of this external fact.

Different accounts of this relation of standing-in have been proposed, from Aristotle’s ‘being-a-token-of’ – “It is not
The representationality of some properties has to be sharply distinguished from their intentionality. A property of something is intentional iff it is taken to be about something else than itself. It is so taken to be if we attribute to it conditions under which it may be said to be correct. Correctness conditions specify the intentional content, but being conditions do not themselves require this content to be satisfied. If I am looking for the Holy Grail, for example, my activity is directed towards, and rationalisable only with respect to the Holy Grail, which, or so let’s assume, does not exist. I am intentionally directed towards the Holy Grail, without standing in a relation to it: there is nothing, after all, for me to stand in a relation to.6

Because they are outward-directed, and cannot be accounted for without reference to their intentional objects, intentional states are extrinsic: they are what they are in virtue of participating in a complex process, which not only involves their objects, their bearer, but also a process of interpretation or understanding.7

The arguments in favour of objects of belief

The most important argument in favour of a relational construal of belief is that we need it to explain the similarity between John's believing that Sam is Maria's wife and Sam's believing that snow is white. If this is just two states of John and Sam respectively, the argument goes, how are we to explain that they make both of them believers? This argument is no good: by the same token, you could say that the similarity between my running at ten kilometres per hour and your running at fifteen kilometres per hour is an argument for construing running as a relation between a runner and a speed. The same example rebuts another argument in favour of a relational account of belief: while it is true that if someone is a believer, then s/he believes something, the soundness of existential generalisation is not conclusive evidence of relationality. It is equally necessary that if you run, you run at a certain speed – but, again, this does nothing to show that running is a relation.

More explicitly, it has been taken to be an argument in favour of the acceptance of objects of belief that we can infer both (4) and (5) from (1-3):

1. Huey believes that snow is white.
2. Dewey believes that grass is green.
3. Louie believes that grass is green.
4. Both Huey and Dewey believe something.
5. There is something that both Dewey and Louie believe.

More generally, we can infer (7) from (6) – allegedly showing that believing is a relation to some object of belief:

6. Huey is a believer.
7. There is something that Huey believes.

Both these arguments are inconclusive, however. For we can also, and with equal right, both infer (4*) and (5*) from (1*-3*) and infer (7*) from (6*):

1*. Huey runs at 8 km/h.
2*. Dewey runs at 10 km/h.

possible to converse by bringing in the objects themselves, but instead of the objects we use words as tokens*, Sophisti Elenchi 1, 164a6-8 – to the scholastic modes of objective existence.

6It is only as an analysis of intentionality, not of representationality, I think, that Aristotle’s theory of thoughts being likenesses of objects has any plausibility.

7This has been made particularly clear by Charles W. Morris, one of the founders of the theory of signs: “The properties of being a sign, a designatum, an interpreter, or an interpretant are relational properties which things take on by participating in the functional process of semiosis.” (1939: 82)
(3*) Louie runs at 10 km/h.
(4*) Both Huey and Dewey run at a certain speed.
(5*) There is some speed at which both Dewey and Louie run.
(6*) Huey is a runner.
(7*) There is some speed at which Huey runs.

But running is not standing in a relation to a speed. What makes the inferences valid, is a certain
metaphysical necessity: that every running occurs at a certain speed means that running always allows
for adverbial modification with a determinate of the determinable speed. This does not, however,
mean that running has an internal object which is a speed.

In the same way, it seems to me, it may very well be true that if someone believes, there is always
something that is believed, even though believing is not a relation to what is believed.

A reason to think that believing that \( p \) is not standing in a relation to (the proposition) that \( p \) (though
perhaps itâ￿s standing in a relation to the objects (the proposition) that \( p \) is about):

Truthmaker maximalists hold that truth is a derelativisation of the relation of being made true,
though not itself relational. The things made true are plausibly taken to be sentence-tokens or in-
scriptions, that are made true in some specific way: The inscription “Huey ate the cookies on the
couchâ￿â￿â￿â￿, e.g., is made true by some cookie-eating event as an inscription of English and as
specifying the location of either Huey or the cookies.

Analogously, an adverbial theory of belief might hold that belief is a derelativisation of a broader
and conceptually prior notion of acceptance, in which sentence-tokens stand to believers in certain
ways. Believing that \( p \) is a matter of there being a sentence-token that is accepted in a certain way
(i.e. the way in which, if it were true, would be made true).

Two immediate problems arise:

1. The famous translation argument, going back to Frege and expanded by Church (1950) and
   Bigelow (1980) against sententialist and paratactic theories of belief ascription seems applica-
   ble: (i) if the relata of acceptance are sentence-tokens, (i) and
   8 Tick glaubt, dass Schnee weiss ist.
   involve relations to different sentences. Moreover, (ii) the acceptance of “Snow is white” by
   Huey can only underlie the truth of (i) if we presuppose that, in it, “snow” means snow rather
   than, say, grass. But the claim is not that
   9 Huey accepts some sentence-inscription which means that snow is white.
   is synonymous with (i), or conceptually equivalent with it, or an explication of what is implicit
   in (i). (9) rather specifies some truthmaker for (i): it is a metaphysical analysis of (i) and their
   correlation a substantial thesis.

2. The adverbial theory seems committed to (i) ascribing beliefs only to language-using creatures
   and (ii) only in contents for which some inscription may be found. Against (ii), the adverbialist
   may claim that the accepted inscription is produced in (i). To avoid (i), the relation of accep-
   tance has to be construed as passive (Cohen 1992) and as a more general attitude than belief as
   ordinarily construed, which "comes apart from belief in cases where one is warranted in acting
   on the assumption that \( p \) or taking it for granted that \( p \) or trusting that \( p \) for reasons that do
   not bear on the likely truth of \( p \)" (Wright 2004: 177). So construed, we may say that (intelligent)
dogs accept sentence-tokens and accordingly have beliefs.
References


