The Property Paradox in (not so plain) English*

Philippe Schlenker
(Institut Jean-Nicod, CNRS; New York University)

Final version in:

Gutzmann, D., Köpping, J. and Meier, C. (eds), *Approaches to Meaning. Composition, Values, and Interpretation,* pp. 203-219, Current Research in the Semantics/Pragmatics Interface 32, Brill.

Abstract: It is well-known that untyped systems with property-denoting expressions that can appear in argument and predicate positions are prone to paradox. We argue that a version of the 'property paradox' can be stated with simple grammatical resources in English, and that a trivalent solution should be given to it, akin to that given to paradoxes of truth.

1 The Property Paradox in the 'Naïve Theory of Properties'

Reference to properties is widely used in semantic analyses of natural language (e.g. Chierchia and Turner 1988, Zimmermann 1993, and Moltmann 2004, among others). But as is summarized in Field 2004, the 'naïve theory of properties' is prone to a 'property' version of Russell's paradox:

According to the naïve theory of properties, for every predicate $\Theta(x)$ there is a corresponding property $\lambda x\Theta(x)$. Moreover, this property $\lambda x\Theta(x)$ is instantiated by an object o if and only if $\Theta(o)$. More generally, the naïve theory involves the following 'naïve comprehension schema':

NC. $\forall u1 \ldots \forall un \exists y[Property\ (y)\ \& \forall x(x \text{ instantiates } y \leftrightarrow \Theta\ (x, u_1 \ldots u_n))]$.

This naïve theory of properties has many virtues, but it seems to have been shattered by (the property version of) Russell's paradox. 'Seems to' have been shattered? There is no doubt that it *was* shattered, if we presuppose full classical logic. Let us use the symbol ∈ to mean 'instantiates'. The Russell paradox involves the Russell property *R* corresponding to the predicate 'does not instantiate itself'. So according to the naïve theory, $\forall x[x \in R \leftrightarrow \neg(x \in x)]$. Therefore in particular,
(*) $R \in R \leftrightarrow \neg(R \in R)$.' (Field 2004)

Field's enterprise is to find a logic that is strong enough to capture part of the naïve theory of properties, yet not so strong that it generates paradoxes. This is certainly a worthy logical task; but if one is interested in the semantics of natural language, one should first ask whether paradoxes can or cannot be constructed from fragments of English that only include 'property talk'. Obviously, if we gave ourselves the word *property* and made the necessary assumptions to ensure that for every formula, there exists a corresponding property, we could express in plain English the property paradox. All we would need to do is to express with English words the formulas that are mentioned in the paragraph by Field that we quoted at the outset. But this should certainly not be taken to show that the *grammar* of English suffices to generate the property paradox. Rather, one should conclude from a natural language rendering of Field's reasoning that certain assumptions about properties lead to paradox. Under this view, then, *nothing in the grammar of English would have to be modified if we were to abandon the belief that for each formula there exists a property that corresponds to it.*

---

*Earlier versions of this manuscript benefited from critical comments by Barry Schein and Chris Tancredi; this version benefited from further comments by Jan Köpping and Graeme Forbes – many thanks to all. Unfortunately, remaining errors are mine. But fortunately, I can safely assume that Ede Zimmermann will find them all, and will probably go on to correct them too, as he regularly has in the past.
When I first met Ede at Sinn und Bedeutung 1998 in Leipzig, he asked me: "Bist du Elsässer?" – and I didn't quite understand the term (I had learned my German in Berlin, and in Hegel). We have had many deeper conversations since that time, on every conceivable topic in semantics, logic and the philosophy of language. They have been a constant intellectual treat and have provided me with long-term food for thought (not to mention the drinking). It is a pleasure to dedicate these musings to Ede – and to wish him many happy returns.
(The research leading to these results was supported in part (after May 1st, 2014) by the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement N°324115 FRONTSEM (PI: Schlenker). Research was conducted at Institut d'Etudes Cognitives (ENS), which is supported by grants ANR-10-IDEX-0001-02 PSL* and ANR-10-LABX-0087 IEC.)

Thus it might be reasonable to conclude that all is well with the grammar of English, though one might get into trouble if one insists on making misguided assumptions about ontology.

2 The Property Paradox in Natural Language

The situation is more complex, however. We will now see that quantifiers and pronouns, combined with plausible rules of inference (ones that have nothing to do with the word *property* or any other technical terminology), might suffice to generate the property paradox in English. However, some linguistic work is needed to come up with the paradoxical statements, which are very complex, for reasons outlined below. But before we embark on this construction, we need to be clear about the role played by quantifiers and pronouns *independently* of the issue of property talk. Thus we will start with some brief remarks about individual-denoting quantifiers and pronouns before applying our results to property talk.

2.1 Individual-denoting Quantifiers and Pronouns

Quine famously argued that 'to be is to be the value of a variable bound by an existential quantifier'. His criterion applies most clearly to a theory stated in First-Order Language with a standard syntax[1], but we can extend it to natural language – on condition, of course, that we identify plausible analogues of quantifiers and variables.

In the case of reference to individuals, the problem is not hard to solve. *Something* is a plausible counterpart of the existential quantifier. For variables, one may either use the pronoun *it* or the relative pronoun *which*. A particularly pedantic translation of (1)a is afforded by (1)b, where every variable is translated by a pronoun.

(1) a. $\forall x_1 \exists x_2 P(x_1, x_2)$
b. Everything1 is such that something$_2$ is such that it$_1$ protects it$_2$.
c. Everything$_1$ is such that there is something$_2$ which$_2$ it$_1$ protects.
d. Everything protects something.

Of course (1)d is by far the most idiomatic translation of (1)a, but it has the disadvantage of not making the variables explicit (many contemporary syntacticians believe that even in this case a particular sort of variables, called 'traces', are syntactically but not phonologically present. They further believe that there exists a level of syntactic representation in which quantifiers roughly figure in the syntactic positions that they would have in First-Order Logic. If so, (1)d has a representation such as the following: *Everything$_1$ something$_2$ t_1 protects t_2*).

Interestingly, there are idiosyncracies of natural language that make the translation of certain First-Order formulas slightly more complicated. Suppose we wanted to translate (2)a using the recipes we used in (1). We would end up with translations that are deviant or do not have the desired meaning:

(2) a. $\exists x_1 P(x_1, x_1)$
b. #Something$_1$ is such that it$_1$ protects it$_1$.
c. #There is something$_1$ which$_1$ protects it$_1$.

The source of the problem lies in a grammatical constraint called 'Condition B', which (among others) prohibits an object pronoun from coreferring with the subject of its predicate (e.g. Chomsky 1981). Of course in this case there is an easy fix, which is to use the reflexive pronoun *itself* instead of *it:*

(3) a. $\exists x_1 P(x_1, x_1)$
b. Something$_1$ is such that it$_1$ protects itself$_1$.
c. There is something$_1$ which$_1$ protects itself$_1$.

Still, even if reflexive pronouns were not available we could find a more devious way of expressing (3)a while circumventing Condition B. The idea is to start, not with (3)a, but with (4)a, which is equivalent to it. We then translate this sentence into English and avoid the Condition B violation:

(4) a. $\exists x_1 \exists x_2 (x_2 = x_1 \ \& \ P(x_1, x_2))$
b. Something$_1$ is such that something$_1$ is such that it$_1$ is identical to it$_2$ and it$_1$ protects it$_2$.
c. There is something$_1$ which$_1$ is identical to something$_2$ which$_2$ protects it$_1$.

[1] As observed in Quine 1995, the usual criterion of 'ontological commitment' must be modified if one uses a system such as Quine's own 'Predicate Functor Logic', which is variable-free.

The idea, then, is to introduce an additional quantifier and a predicate of identity to circumvent the prohibition against coreferential arguments[2]. This piece of trickery will come in handy in our discussion of property talk, where we won't be able to appeal to reflexive pronouns to circumvent Condition B.

2.2 Property-denoting Quantifiers and Pronouns

Let us now turn to properties. First, it seems that we can quantify over them:

(5) John is clever. Therefore there is something that John is – namely clever.

Furthermore, pronouns sometimes appear to refer back to properties. This is somewhat easier to show in French or in Italian than in English, because these languages allow regular pronouns to play this role (like other pronouns in French and Italian, they appear in the pre-verbal ('clitic') position; unlike individual-denoting pronouns, they always take 'default' gender, i.e. they always appear in the masculine even if the predicates they refer back to exhibits feminine features, as is the case below). In English, the pronoun *it* doesn't work well in this case, but the relative pronoun *which* does (we may marginally use the demonstrative *that* to obtain the desired meaning, but the resulting sentences are still a bit awkward):

(6) a. Anne est travailleuse et sa fille le sera aussi (French)
Anne is hardworking-feminine and her daughter it-masc will-be too
'Anne is hardworking and her daughter will be too'
b. Ann is hardworking, which her daughter is too
c. ? Ann is hardworking, and her daughter is that too

So far we have only demonstrated that natural language has quantifiers and pronouns that may range over properties. The key to produce a paradox, however, is to show that the underlying system is not fully typed, and that *different occurrences of the same property-denoting variable can appear in predicate and in argument positions*. If this were not the case, we could assume that the system is 'implicitly' typed, in the sense that even though the *morphology* does not display any distinction between property-denoting pronouns that appear in argument and in predicate positions, syntactic or semantic rules still prohibit the same variable from simultaneously appearing in both types of position.

But (7)b suggests that this is probably not the case in English (I write 'probably' because some speakers might find (7)b a bit marginal); and I find (7)b' rather acceptable in French. It seems that certain parts of the language might not be typed.

(7) a. John is something$_1$ which$_1$ is very important [, namely clever].
b. (?) There is something$_1$ which$_1$ John is and which$_1$ is very important [, namely clever].
b'. Il y a quelque chose que Jean est, à savoir sage, et qui est fort important dans la vie.
It there have some thing that Jean is, to know wise, and that is very important in life.
'There is something that Jean is, namely wise, which is very important in life.'
c. $\exists x_1\ (x_1(j)\ \&\ important(x_1))$

The key formal observation, made apparent in (7)b-c, is that the same variable (namely *which*$_1$ in (7)b and x_1 in (7)c) occurs both in a predicate position (*which*$_1$ *John is* in (7)b and $x_1(j)$ in (7)c) and in an argument position (*which*$_1$ *is important* in (7)b and $important(x_1)$ in (7)c).

At this point a clarification is in order. Jan Köpping (p.c.) suggests that instantiated counterparts of (7)b fail, as shown in (8)a; this might cast doubt on the idea that property talk in English is entirely untyped.

(8) a. John is clever. *Clever is important.
b. C(j) and important(j)

But we do not need to claim that the system is entirely untyped. In order to obtain a derivation of the property paradox with the grammatical resources of ordinary English, all we need to do is to find a *subpart* of the language which is 'sufficiently' untyped to give rise to the paradox. As far as we can tell, sentences like (7)b-b' are rather acceptable whereas instantiated counterparts like (8)a aren't, but we won't need the latter to derive the paradox.

Although in the preceding cases we started from properties that were denoted by atomic predicates, the same constructions can be extended to properties that are denoted by complex formulas. Thus in the examples in (9) it is possible (though admittedly not required) to understand the predicate-denoting pronoun as referring to the conjunctive property of being intelligent and hard-

[2] It can be ascertained independently that (i) pronouns bound by different quantifiers do not in general count as 'coreferential', and that in any event (ii) what counts for purposes of Condition B is presupposed rather than asserted identity.

working (note that in French the singular form of the pronoun must be used; the plural pronoun is simply ungrammatical in this case).

(9) a. Anne est (à la fois) intelligente et travailleuse, et sa fille le (*les) sera aussi.
*Anne is (both) intelligent and hard-working, and her daughter it (*them) will-be too.*
'Ann is both intelligent and hard-working, and her daughter will be too'
b. Ann is (both) intelligent and hard-working, which her daughter is too.

Furthermore, in this case as well different occurrences of the same property-denoting variable may occur in predicate and in argument positions:

(10) a. Ann is both intelligent and hard-working, which$_1$ is useful but which$_1$ unfortunately her daughter isn't.
b. Ann is something which$_1$ is useful but which$_1$ unfortunately her daughter isn't (namely (both) intelligent and hard-working).

In order to represent these sentences insightfully, we use in our discussion a notation in which complex predicates can be formed by an operation of λ-abstraction. In (11)a we represent *Ann* and *her daughter* with *a* and *d* respectively, and we assume that x_1 is a free variable which denotes the same thing as $[\lambda x_2\ (intelligent(x_2) \wedge hard\text{-}working(x_2))]$. On this assumption, (11)b follows from (11)a.

(11) a. $[\lambda x_2\ (\text{intelligent}(x_2) \wedge \text{hard-working}(x_2))](a) \wedge \text{useful}(x_1) \wedge \neg x_1(d)$
b. $\exists x_1(x_1(a) \wedge \text{useful}(x_1) \wedge \neg x_1(d))$

If our judgments are correct (and they could of course be challenged), these expressive possibilities will now allow us to generate the property paradox in plain English, without any recourse to technical terms such as *property*.

2.3 The Property Paradox in English

Let us start from the inference in (12)a to (12)b, which is plausibly valid:

(12) a. This is green. Therefore:
b. There is something$_1$ [, namely green,] which$_1$ each thing$_2$ is if and only if it$_2$ is green.

This inference has the following form (we represent *this* by a free variable or a constant y):

(13) a. $g(y)$
b. $\exists x_1\ (\forall x_2(x_1(x_2) \leftrightarrow g(x_2))$

In this case the formula that triggers the inference involves an atomic predicate. But it would seem that the inference is also intuitively valid when the initial formula involves a complex predicate. Let us start with a simple example, in which the pronoun *which*, represented in (14)b as a free variable, appears to refer to the property denoted by $\lambda x_2\ (green(x_2) \wedge tasty(x_2))$; we use *a* to stand for *this apple*, and *a'* to stand for *that apple*.

(14) a. This apple is green and tasty, which that apple isn't.
b. $[\lambda x_2\ (\text{green}(x_2) \wedge \text{tasty}(x_2))](a) \wedge \neg x_1(a')$

(On this construal, the denotation of *a'* may validate the second conjunct by not being green or not being tasty or both.)

It seems reasonable to infer (15) from (14), where the part between square brackets may be omitted in (15)a and in (15)b:

(15) a. There is something$_1$ [, namely green and tasty / being both green and tasty,] which this apple is and which that apple isn't.
b. $\exists x_1\ (x_1(a) \wedge \neg x_1(a')\ \ [\wedge\ x_1 = \lambda x_2\ (\text{green}(x_2) \wedge \text{tasty}(x_2))])$

It would also seem reasonable to infer (16):

(16) a. There is something$_1$ which$_1$ each thing$_2$ is if and only if it$_2$ is (both) green and tasty.
b. $\exists x_1\ (\forall x_2\ (x_1(x_2) \leftrightarrow (g(x_2) \wedge t(x_2)))$

Let us now derive the paradox itself. We start from true statements such as those in (17):

(17) a. This apple is green, which (itself) is colorless.
b. This apple is concrete, which (itself) is abstract.

We may want to infer the sentence in (18), represented as in (19), where the variables x_1 and x_2 denote the same things as *concrete* and *green* respectively:

(18) a. This apple is concrete, which (itself) is not concrete.
b. This apple is green, which (itself) is not green.

(19) a. $\text{concrete}(a) \wedge \neg\text{concrete}(x_1)$
b. $\text{green}(a) \wedge \neg\text{green}(x_2)$

If indeed x_1 and x_2 denote the same things as *concrete* and *green* respectively, we may be tempted to find a natural language counterpart of the formulas in (20).

(20) a. $\text{concrete}(a) \wedge \neg x_1(x_1)$, where x_1 denotes the property of being concrete.
b. $\text{green}(a) \wedge \neg x_2(x_2)$, where x_1 denotes the property of being green.
c. $\exists x_1\, (x_1(a) \wedge \neg x_1(x_1))$

But the most natural candidates, given in (21), are ungrammatical or do not have the intended meaning:

(21) a. #This apple is concrete, which_1 it_1 isn't.
b. #This apple is green, which_2 it_2 isn't.
c. #This apple is something_2 which_2 it_2 isn't.

However this should come as no surprise: as was observed in our discussion of (2), Condition B prohibits an object pronoun from referring to the same thing as the closest subject. But this is precisely what the relative pronoun *which* is trying to do in (21)![3] Still, with some trickery, we can express the desired meaning. Thus instead of translating directly the sentences in (20), we will use the more complicated sentences in (22), whose natural language counterparts are given in (23):

(22) a. $\text{concrete}(a) \wedge \exists x_2\,(x_1 = x_2 \wedge \neg x_2(x_1))$ with x_1 = concrete
b. $\text{green}(a) \wedge \exists x_2\,(x_1 = x_2 \wedge \neg x_2(x_1))$ with x_1 = green
c. $\exists x_1\, (x_1(a) \wedge \exists x_2\,(x_1 = x_2 \wedge \neg x_2(x_1)))$

(23) a. This apple is concrete, *which_1 is identical to something_2 which_2 it_1 isn't.*
b. This apple is green, *which_1 is identical to something_2 which_2 it_1 isn't.*
c. This apple is something_1 *which_1 is identical to something_2 which_2 it_1 isn't.*

Since there is an ambiguity in English *be*, which could be understood as 'identity be' or 'predicative be', *we adopt the convention of using* 'be identical to' *in the first case and simple* 'be' *in the second.* The resulting sentences in (23) are undoubtedly complicated, ambiguous, and hard to understand. Our point is not that these are idiomatic or even comprehensible English, but just that (i) they are formed by means that do not appear to violate the grammar of English, and that (ii) if our earlier observations are correct, they should have a reading that corresponds to the formulas in (22) (that they are difficult to understand could come from the muliple ambiguities they give rise to, from their syntactic complexity, or from the multiple repetitions they involve). In particular, they satisfy Condition B; for instance, in (23)c the two property-denoting pronouns that occur in *which_2 it_1 isn't* are bound by different quantifiers, and therefore do not count as 'coreferential'.

We may now apply generalization illustrated in (13) and (16) to the italicized complex predicates in (23) – with the result in (24)a. The latter is the natural language translation of (24)b, which in turn is equivalent to (24)c.

(24) a. There is something_3 which_3 each thing_1 is if and only if it_1 is identical to something_2 which_2 it_1 isn't.
b. $\exists x_3\,(\forall x_1\,(x_3(x_1) \leftrightarrow \exists x_2\,(x_2 = x_1 \wedge \neg x_2(x_1))))$
c. $\exists x_3\,(\forall x_1(x_3(x_1) \leftrightarrow \neg x_1(x_1)))$

The property paradox arises when we ask about the witness of the statement of (24)a whether it itself has the relevant property. Due to the ability of natural language pronouns to pick out the witness of a preceding existential statement, as in (25), the paradox will arise when (24)a is followed by a relatively simple question, as in (26)b:

---

[3] In addition, the sentences in (21) exhibit a configuration in which (i) a *wh*-expression moves to the left of a coindexed pronoun, and (ii) the pronoun in question c-commands the trace of the *wh*-expression. Examples that satisfy (i) but not (ii) are deemed 'Weak Crossover' violations; those that satisfy (i) and (ii) are called 'Strong Crossover' violations. But as noted in Lasnik and Stowell 1991 (among others), with appositive relative clauses Weak Crossover effects just fail to arise in this configuration, as shown in (i).

(i) Gerald_i, who_i his_i mother loves t_i, is a nice guy. (Lasnik and Stowell 1991 p. 698)

(I personally find the following sentence rather odd: *Gerald_i, who_i he_i thinks nobody likes t_i, is a nice guy.* This might seem to go against Lasnik and Stowell's generalization. But there might be an orthogonal problem with this sentence, which involves the semantics of attitude reports: some authors have argued that De Se Logical Forms should be preferred whenever they can be used to report an attitude (e.g. Schlenker 2005). Here the trace of *who_i* must a be a De Re, non De Se term, which might explain the deviance of the sentence.)

(25) You bought something to eat. Was it good?

(26) There is $something_3$ $which_3$ each $thing_1$ is if and only if it_1 is identical to $something_2$ $which_2$ it_1 isn't.
Is it_3 identical to $something_2$ $which_2$ it_3 is?

The question has no non-self-contradictory answer. To say it in technical terms: the witness of the existential statement is self-applicable if and only if it is not self-applicable. Without using any technical vocabulary, we can state the paradox in (not quite) ordinary English:

(27) It_3 is identical to $something_2$ $which_2$ it_3 is if and only if it_3 is identical to $something_1$ $which_1$ it_3 isn't.

Needless to say, given the complexity of the sentence and the multiple ambiguities it gives rise to, more work would be needed to make sure that the crucial reading is indeed derivable by independently attested grammatical means. We leave this empirical investigation for future research.

3 Towards an Objectual Semantics

Let us now assume for the sake of argument that our analysis is on the right track, and that property-denoting pronouns and quantifiers make it possible to express the property paradox with the grammatical resources of ordinary English. There are two conceivable solutions to this problem, each of which departs in some way from speakers' immediate intuitions.
–We could deny that the patterns of inference that yield the paradox apply without restriction. In particular, we could deny that from *John is P* one can *always* infer: *There is something that John is, namely P*. If so, we could claim that there is no real object-language paradox here; the appearance of one might stem from the fact that speakers erroneously apply incorrect principles of reasoning that are not sanctioned by the grammar of their language.
–Alternatively, we could bite the bullet and say that property talk does make it possible to construct paradoxical statements in the object language.

Since the first line has been explored at length by others, we will briefly discuss ways to develop the second. It should be pointed out that there is a rich literature on this topic, some of which concerns models of the untyped lambda-calculus, and some of which is more linguistically oriented (see for instance Hindley and Seldin 1986 for the former, and Turner 1983, 1985 for the latter). Here we sketch a simple existence proof that some models with the desired properties can be constructed. Within a highly simplified language, we will take the denotation of property-denoting expressions to be... formulas of another language. This will be achieved by way of translation into a language with a satisfaction predicate and with quantifiers that range over formulas of that same language. In this way, we will translate $x_2(x_1)$ as $Sat(x_2, x_1)$, where the latter formula means that the object denoted by x_1 satisfies (i.e. makes true) the formula with one free variable denoted by x_2. The advantage of this reductive procedure is that it is straightforward to construct models for a language that contains its own satisfaction predicate. But it is clear that more sophisticated models should be investigated if the present line is to be developed seriously.

3.1 Syntax and Translation

Our base language is formed by the following syntactic rules (constants could be added, or treated *à la Quine* by analyzing *Socrates sleeps* as *there is something which Socratizes and which sleeps*):

(28) L (= the base language)

$$F := P^n_i (x_{k_1} \dots x_{k_n}) \mid \neg F \mid (F \wedge F') \mid (F \leftrightarrow F') \mid \exists x_k F$$

In order to model property talk, we extend L to a language L^P which includes formulas of the form $x(x')$, where x and x' are two variables:

(29) L^P (= L extended with property talk)

$$F := x_{k_1}(x_{k_2}) \mid P^n_i (x_{k_1} \dots x_{k_n}) \mid \neg F \mid (F \wedge F') \mid (F \leftrightarrow F') \mid \exists x_k F$$

The semantics of L^P will be obtained by way a translation into an extension L^+ of L which contains a satisfaction predicate *Sat*:

(30) L^+ (= L extended with a satisfaction predicate)

$$F := Sat(x_{k_1}, x_{k_2}) \mid P^n_i (x_{k_1} \dots x_{k_n}) \mid \neg F \mid (F \wedge F') \mid (F \leftrightarrow F') \mid \exists x_k F$$

It is intended that the interpretations we consider for L^+ satisfy the following condition (i.e. that they are 'fixed points' in the sense of Kripke 1975):

(31) *Sat(x, y)* is:
a. true just in case *x* denotes a formula with (exactly) one free variable which is made true by the denotation of *y*;

b. false just in case either (i) *x* denotes a formula with one free variable which is made false by the denotation of *y*, or (ii) *x* denotes something which is not a formula with one free variable;
c. neither true nor false in all other cases (hence in particular if *x* denotes a formula with one free variable but is made neither true nor false by the denotation of *y*).

Standard results guarantee that such interpretations can indeed be constructed (see for instance McGee 1992 p. 170).

The translation between L^P and L^+ is then straightforward: any formula *F* of L^P is translated into a formula *F** of L^+ obtained by replacing every occurrence of the form *x'(x)* with *Sat(x', x)* – as is stated more pedantically in (32):

(32) a. $[P^n_i (x_{k_1} ... x_{k_n})]^* = P^n_i (x_{k_1} ... x_{k_n})$
b. $[\neg F]^* = \neg F^*$
c. $[(F \wedge F')]^* = (F^* \wedge F'^*)$
d. $[(F \leftrightarrow F')]^* = (F^* \leftrightarrow F'^*)$
d. $[\exists x_k F]^* = \exists x_k F^*$
e. $[x_{k_1}(x_{k_2})]^* = Sat(x_{k_1}, x_{k_2})$

3.2 Semantics

Following one of the options laid out in Kripke 1975, we resort to the Strong Kleene System to give a semantics to L^+ and thus – indirectly – to L^P. The key clauses are copied below; the existential quantifier is treated as general disjunction (which in turn can be defined from negation and conjunction):

P	¬ P
1	0
0	1
#	#

P \ Q (P ∧ Q)	1	0	#
1	1	0	#
0	0	0	0
#	#	0	#

P \ Q (P ↔ Q)	1	0	#
1	1	0	#
0	0	1	#
#	#	#	#

For any interpretation function I and any assignment function s, for each formula G,

$[\![\exists x_i\, G]\!]^s = 1$ iff for at least one object d, $[\![G]\!]^{s[x_i \to d]} = 1$
$[\![\exists x_i\, G]\!]^s = 0$ iff for each object d, $[\![G]\!]^{s[x_i \to d]} = 0$.

As in Kripke's theory of truth, the domain of objects includes formulas of L^+, which is of course crucial to obtain paradoxical statements.

3.3 Examples

Let us illustrate the translation procedure with a few examples. In each case (a) is the English sentence to be analyzed, (b) is its translation in our property-friendly system, and (c) is the translation of (b) under *.

(i) Simple Examples

(33) Something is green.
b. $\exists x_1$ green(x_1)
c. $\exists x_1$ green(x_1)

(34) a. There is something$_1$ which$_1$ each thing$_2$ is if and only if it$_2$ is green.
b. $\exists x_1 \forall x_2 (x_1(x_2) \leftrightarrow$ green$(x_2))$
c. $\exists x_1 \forall x_2 (Sat(x_1, x_2) \leftrightarrow$ green$(x_2))$

(ii) A 'reflexive' example

Let us now consider a sentence with a reflexive flavor:

(35) a. Something$_1$ is identical to something$_2$ which$_2$ it$_1$ isn't.
b. $\exists x_1 \exists x_2 (x_1=x_2 \wedge \neg x_1(x_1))$
b'. $\exists x_1 \neg x_1(x_1)$
c. $\exists x_1 \neg Sat(x_1, x_1)$

The formula in (35)b is equivalent to that in (35)b', whose semantics is provided by (35)c. This sentence is trivially true given our conventions: if x_1 denotes something other than a formula with one free variable, by the condition in (31)b, $Sat(x_1, x_1)$ is true, and so is $x_1(x_1)$.

(iii) A paradoxical statement

Let us now consider a sentence which will turn out to be paradoxical according to the present system:

(36) a. There is something$_3$ which$_3$ each thing$_1$ is if and only if it$_1$ is identical to something$_2$ which$_2$ it$_1$ isn't.
b. $\exists x_3 \forall x_1 (x_3(x_1) \leftrightarrow \exists x_2 (x_2 = x_1 \wedge \neg x_2(x_1)))$
b'. $\exists x_3 \forall x_1(x_3(x_1) \leftrightarrow \neg x_1(x_1))$
c. $\exists x_3 \forall x_1(Sat(x_3, x_1) \leftrightarrow \neg Sat(x_1, x_1))$

We will work directly with the simplified formula in (36)b', whose semantics is provided by (36)c. We call G the formula $\forall x_1(x_3(x_1) \leftrightarrow \neg Sat(x_1, x_1))$.
–Clearly, no object assigned to x_3 can make G true, since for $x_1 = x_3$ G cannot be true.
–On the other hand, when x_1 denotes the formula $H := \neg Sat(x_1, x_1)$, the right-hand side of the biconditional can only have the indeterminate truth value. This is because in this case $Sat(x_1, x_1)$ is true just in case H satisfies $\neg Sat(x_1, x_1)$, which holds just in case $\neg Sat(x_1, x_1)$ is true (since x_1 denotes H); and similarly $Sat(x_1, x_1)$ is false just in case $\neg Sat(x_1, x_1)$ is false. Only the third case is coherent: $Sat(x_1, x_1)$ must have the indeterminate truth value.

3.4 An improvement

The semantics of (36) isn't quite what we had promised, however. We wanted, in essence, to guarantee that from:

(37) $\neg x_2(x_2)$ (for some value of x_2)

one could infer:

(38) $\exists x_3 \forall x_1(x_3(x_1) \leftrightarrow \neg x_1(x_1))$

But we have failed: (37) is often trivially true, but we have just shown that (38) (i.e. (36)b') has the indeterminate truth value.

What is the source of our predicament? The problem lies in our treatment of the biconditional, whose Strong Kleene semantics is exceedingly weak: as soon as one of its arguments is indeterminate, so is the entire biconditional. What we would like is an alternative biconditional (henceforth written as $\leftrightarrow^*$), which is true just in case its two arguments have the same truth value, and is false otherwise:

(39) $[\![G \leftrightarrow^* H]\!] = 1$ iff $[\![G]\!] = [\![H]\!]$; $[\![G \leftrightarrow^* H]\!] = 0$ iff $[\![G]\!] \neq [\![H]\!]$

Why couldn't we add $\leftrightarrow^*$ to a Strong Kleene Logic? Because in the general case it would make the system inconsistent. To see why, observe that $\leftrightarrow^*$ combined with Kleene's strong negation $\neg$ suffices to define the (so-called 'weak') negation $\neg^*$:

(40) a. $[\![\neg G]\!] = 1$ iff $[\![G]\!] = 0$; $[\![\neg G]\!] = 0$ iff $[\![G]\!] = 1$
b. $[\![\neg^* G]\!] = 1$ iff $[\![G]\!] = 0$ or $[\![G]\!] = \#$; $[\![\neg^* G]\!] = 0$ iff $[\![G]\!] = 1$
c. It follows that for any tautology T, $[\![\neg^* G]\!] = [\![\neg (G \leftrightarrow^* T)]\!]$

But the problem is that once $\neg^*$ is introduced in the language, paradoxes can be defined which cannot be treated within the Strong Kleene System. In particular, we can define a 'Strengthened Liar' λ which says: $\neg^* Tr(\lambda)$, where Tr is the truth predicate. It is immediate that λ can be neither true nor false, for the familiar reasons; but it also can't be indeterminate, because it would be something other than true and should thus be true given the semantics of $\neg^*$. The situation is no less dismal with the property paradox. We start by defining the property of not being self-applicable in terms of $\neg^*$:

(41) $[\lambda x\ \neg^* x(x)]$

The problem arises when we ask whether this property – call it π – is or isn't self-applicable; it is immediate that $\pi(\pi)$ has a classical truth value, and further more that $\pi(\pi)$ if and only if $\neg^*\pi(\pi)$.

But there is no reason to be deterred. Nothing prevents us from using $\leftrightarrow^*$ in an *extension* of our language, as long as it doesn't appear in the formulas the quantifiers range over (in other words, we first construct a Kripkean interpretation for the language L^P or L^+ without $\leftrightarrow^*$, whose quantifiers range over formulas of L^+; only then do we add $\leftrightarrow^*$ to the result to obtain a new language which *does not* contain its own satisfaction predicate, but a satisfaction predicate for L^+). It is then immediate that from the truth of any formula $F[x_1]$ of L^P with one free variable x, we can infer:

(42) $\exists x_3 \forall x_1(F[x_1] \leftrightarrow^* x_3(x_1))$

The reason is simple. Consider the translation $F^*[x_1]$ of $F[x_1]$ obtained (as in (32)) by replacing every occurrence of the form $x'(x)$ with $Sat(x', x)$. By the definition of our semantics for L^P, $F^*[x_1]$ has the same value as $F[x_1]$. Furthermore, $Sat(x_3, x_1)$ has the same value as $F^*[x_1]$ if x_3 denotes $F^*[x_1]$. Finally, by the definition of our semantics for L^P, $x_3(x_1)$ has the same value as $Sat(x_3, x_1)$, which is the value of $F[x_1]$ if x_3 denotes $F^*[x_1]$. This shows that (42) is true.

Since this result applies in full generality, it holds of (36) as well. Starting from a formula $\neg x_2(x_2)$, where x_2 does not denote a formula, we obtain a trivial truth, from which we can infer the desired result, thanks to (42):

(43) b'. $\exists x_3 \forall x_1(x_3(x_1) \leftrightarrow^* \neg x_1(x_1))$
c. $\exists x_3 \forall x_1(Sat(x_3, x_1) \leftrightarrow^* \neg Sat(x_1, x_1))$

On the other hand, it is interesting to observe the behavior of the formula (or rather of one formula) that witnesses the truth of this existential claim. In this case we have:

$F[x_1] = \neg x_1(x_1)$

$F^*[x_1] = \neg Sat(x_1, x_1)$

If x_3 denotes $F^*[x_1]$, i.e. $\neg Sat(x_1, x_1)$, the desired result is immediate: $Sat(x_3, x_1)$ has the same value as $\neg Sat(x_1, x_1)$. (Furthermore, when x_1 denotes the formula $\neg Sat(x_1, x_1)$, it is clear that the latter must have the indeterminate truth value, since the formula satisfies itself if and only if it doesn't; $F^*[x_1]$ can thus be thought of as a 'trivalent property', which is indeterminate when applied to itself.)

As mentioned, more sophisticated interpretations will not doubt have to be constructed to provide a serious trivalent semantics for property talk. We leave this for future research.

If our conclusions are correct (a big *if* indeed), it remains to ask *why* it is that paradoxes can be generated with property talk in natural language. I would like to answer: *why not*? There is no particular reason to assume that language should have been 'created' so as to eschew the property paradox – or other paradoxes, for that matter. In fact, from a naturalistic standpoint it is rather unsurprising that parts of the grammar should be untyped or badly typed (certainly the existence of object language paradoxes exerted a negligible pressure on the evolution of language, and had no reason to be avoided!). Our conclusions are perhaps surprising if one thinks that God created language, that he was a good logician, and that he cared about paradoxes. The present note casts doubt on this conjunction.

References

Chierchia, Gennaro and Turner, Raymond: 1988, Semantics and property theory. *Linguistics & Philosophy*, 11:261-302.

Chomsky, Noam: 1981, *Lectures on government and binding*. Dordrecht: Foris.

Field, Hartrey: 2004, The Consistency of the Naïve Property of Sets. *The Philosophical Quarterly*, 54, 214: 78-104

Hindley, J. Roger and Seldin, Jonathan P.: 1986, *Introduction to Combinators and λ-Calculus*. Cambridge University Press.

Kripke, Saul: 1975, Outline of a Theory of Truth. *Journal of Philosophy* 72: 690-716

Lasnik, Howard and Stowell, Timothy: 1991, Weakest Crossover. *Linguistic Inquiry*, Vol. 22, No. 4.

McGee, Vann: 1990, *Truth and Paradox,* Hackett.

Moltmann, Friederike: 2004, 'Properties and Kinds of Tropes: New Linguistic Facts and Old Philosophical Insights'. *Mind* 113, 1-43.
Quine, Willard Van: 1995, *From Stimulus to Science*. Harvard University Press.
Schlenker, Philippe: 2005, The Lazy Frenchman's Approach to the Subjunctive. In T. Geerts, I. van Ginneken & H. Jacobs (eds.), *Romance Languages and Linguistic Theory 2003*. Amsterdam: Benjamins, 269–309.
Turner, Raymond: 1983, Montague semantics, Nominalization & Scott's Domains. *Linguistics & Philosophy* 6: 259-288.
Turner, Raymond: 1985, Nominalizations and Scott's Domains II. *Notre Dame Journal of Formal Logic* 26, 4: 463-478
Zimmermann, Thomas Ede: 1993, On the proper treatment of opacity in certain verbs. *Natural Language Semantics*, 1, 149–179.