Lectures on Universal Logic

Lecture 9 – Semantic Closure in Universal Logic

The (accidental) occurrence of inconsistencies is certainly not enough to believe in true contradictions. The main motivation for strong paraconsistency (dialetheism) is *universality* as a feature of language and cognition.

Universality means that we are aiming at – and supposedly capable of – a theory of, say, language *in general*, that is not just of this or that language or languages of this or that formal structure. And this theory is expressed in language, so that at least some language can be its own meta-language (with respect to all interesting properties of that language, semantics included).

Universality means as well that we use fundamental concepts like 'set' unrestricted (i.e. not only for sets of some sort, given in sets of another kind). That means there is a universal set corresponding to 'set'.

Syntactic Closure and Self-Reference

A language can talk consistently about its own *syntax*. This presupposes expressive resources to name expressions/terms of the language and to represent syntactic properties. Once a language contains a basic system of arithmetic (the system **Q**) this is feasible.

Within such a language L one can give a general structural description of what is a well-formed expression, what expressions are generalizations, and one can even define what counts as a derivation or a proof in that language. *Being provable in L* as a syntactic property of a formula can be defined within language L. Even if there was a hierarchy of language levels in L, given that the levels have some index there are formula that consistently talk about all such indices or talk about indices that are above the level such a formula is on. Syntactic hierarchies in contrast to semantic hierarchies are not strictly downwards, even in standard logic.

Semantic Closure Defined

A language L is semantically closed iff L is able to talk about its own semantics. The meanings of the terms of L can be given within L then.

If a language is semantically closed it can not only talk about its own expressions (by suitable names or quotation marks), but it can also apply semantic properties to these terms, and even to the terms that express semantic properties.

Given that a language L has names for its terms, one can build the name for an open formula "F( )", and one can then insert that very name into the open formula itself, given rise to the self-application of formulas and properties.

Examples are equivalents to more intuitive expressions like

(1) Sentence number (1) is well-formed.

(2) This sentence contains five words.

(3) The third sentence on this page is true.

The sentences in the examples above are often called "self-referential". There are, however different types of self-referentiality:

(a) *temporal back reference*: x refers to itself in that phase 2 of (enduring) x refers to phase 1 [Example: a talk];

(b) *partial self-referentiality*: x refers by part y of x to part z of x [Example: a name in a language L referring to another term of L];

(c) *partially mediated self-reference*: x refers by some part y of x to x as a whole and thereby also to the part y referring [Example: the expression "language" refers to the whole and thereby to itself as well];

(d) *performative self-reference*: some x refers in a process to itself such that the reference is guaranteed successful [Example: the pronoun "I" taken to mean "the one speaking now"];

(e) *total self-referentiality*: some x is structured in a way that it can only exist as this self-referential structure, and has no parts not being part of that structure [Example: traditional model of the *Ego* ("*the* I")].

Type (c) is the one we are concerned with here: Some formula refers to itself by a description/name occurring in that formula.

The Liar

Given semantic closure and self-referentiality, a sentence/formula expressing some semantic property can ascribe that property to itself. (Such a sentence/formula is called a ‘fixed point’.) An infamous example is the Liar:

(1) Sentence (1) is false.

The Liar is a fixed point for the predicate "( ) is false", or – in some other version – of "( ) is not true".

Now consider (1): If (1) is true, then (1) is false, because the general term in (1) "( ) is false" should apply to the singular term "sentence 1". If (1) is false, then (1) is true, because (1) is just saying that it is false. So we get:

True(1) ≡ False(1)

or given that we have a two-valued logic where "false" is just the opposite of "true":

True(1) ≡ ¬ True(1)

So (1) is an *antinomy*: a sentence A where we have a proof for A and ¬A.

Usually this means we also have a proof of A ∧ ¬A [cf. Lectures 3 and 5].

As an example, let’s walk through the two proofs in case of the Liar:

[Proof of ¬A]. If (1) is true, then (1) is false (by simple properties of being true). This is inconsistent, i.e. the assumption of (1) being true leads to inconsistency, so by the rule of Negation-Introduction (¬I) we get that (1) cannot be true

[Proof of A]. If (1) is false, then (1) is true, just saying that it is false. This is once again inconsistent, i.e. the assumption of (1) being false leads to inconsistency so by (¬I) we get that (1) cannot be false.

[Proof of A ∧ ¬A]. Taking the two proofs together by Conjunction-Introduction (∧I) we get: (1) is true and (1) is false.

Given standard logic from "True(1) ≡ ¬True(1)" we can easily derive an explicit inconsistency "True(1) ∧ ¬True(1)":

1.<1> True(1) ≡ ¬True(1)	PREM
2.<1> (True(1) ⊃ ¬True(1)) ∧ (¬True(1) ⊃ True(1))	Df. ≡,1
3.<1> True(1) ⊃ ¬True(1)	(∧E),2
4.<1> ¬True(1) ⊃ True(1)	(∧E),2
5.<5> True(1)	PREM
6.<1,5> ¬True(1)	(⊃E), 3, 5
7.<1> ¬True(1)	(¬I), 5, 6
8.<8> ¬True(1)	PREM
9.<1,8> True(1)	(⊃E), 4, 8
10. <1> ¬ ¬True(1)	(¬I), 8, 9

11.<1> True(1)   (¬E), 10

12.<1> True(1) ∧ ¬True(1)   (∧I), 7, 11 ■

Proofs are given here often in standard Natural Deduction fashion. In (5) a premise is introduced for reduction, and discarded (written "5") given the contradiction between (5) and (6). The same for (8).

The Liar is a case of bad semantic self-referentiality. There are also harmless semantically self-referential sentences, like:

(2) Sentence (2) is true.

Consider (2):

If (2) is true, then the general term "( ) is true" applies to the singular term "sentence (2)" and so (2) is true. No contradiction occurs.

If (2) is false, then (2) is not true, i.e. (in two-valued logic) (2) is false. No contradiction occurs. It seems that we can either assume (2) to be true or to be false. (2) is underdetermined in that we cannot reason from the falsehood of one of these assumptions to the truth of its opposite.

So by *fiat* we say that (2) is true, and call it "the truth-sayer".

There are more harmless semantically self-referential sentences, like

(3) Sentence (3) is meaningful.

The Liar is the basic case of bad semantic self-referentiality. It bears its badness on its sleeve. It is not hidden. Within a language that is semantically closed semantic self-reference may be hidden, however.

It need not be the case that one sentence refers to itself directly (as in the case of the Liar). There may be situations of cross-reference the result of which is a Liar-like situation. Consider:

(4) Sentence (5) is true.

(5) Sentence (4) is false.

Given that (5) refers to (4) the self-attribution of falsehood in (4) is delayed, so to speak, through (5). We get a version of the Liar. Now, however, we may not recognize it so easily. Suppose today you only utter a single statement

(6) What the pope declares today is true.

and *as a contingent matter of fact* he only says that day:

(7) Everything N.N. says today is false.

The Dialetheist's Agenda

The Liar is an antinomy. Dialetheism claims that it cannot be prevented, since a natural language is a semantically closed language. Since, furthermore, the antinomy can be proved, it has to be true.

So, the dialetheist has to show three things:

(I) The contradictions can be proven in a sound non-standard logic,
if we use a semantically closed language.

(II) We have to use a semantically closed language.

(III) There is no satisfactory alternative to accepting the antinomies.
[i.e. the attempts to prevent them either fail or have consequences
worse than dialetheism]

Convention (T)

(ad I)

To reason from the provability of an antinomy to its truth we need at least one correct/sound paraconsistent logical system. Besides that the reasoning about the Liar employs some properties of negation and Tarski's convention

(T) "p" is true (in L) if and only if p.

ascribing truth to the name of a sentence/statement is true just in case we can use that sentence truly, i.e. it is true. This convention is not tied to a specific theory of truth, since it may be taken either as a minimal condition, a statement of correspondence or a disquotational analysis. Whatever truth is, the adequacy of (T) seems to be beyond doubt. It is neither tied to the underlying logic being 2- or many-valued.

(T) formulates a part of our ordinary understanding of truth. It should be part of the semantics of a semantically closed language.

Relating Convention (T) and the Liar we have for the Liar:

(8) "(1) is false" is true (in L) if and only if (1) is false.

and given that the Liar just is the sentence "(1) is false" and falsity is not truth we get:

(9) (1) is true (in L) if and only if (1) is not true.

There is room for debate how to take the "if and only if" in (T). So far, we have taken it as a simple bi-conditional [≡]. In many systems of paraconsistent logics it will be a stronger bi-conditional (like [↔], maybe disallowing for some detachments allowed for [≡]).

Negation

True contradictions are said to be sentences such that A and ¬A are true. How is that compatible with our concept of negation? How can we justify that [¬] behaving thus still expresses *negation*?

An intuitive acceptable concept of negation should support:

(MN) (i) If A is false, then the negation of A is true.

(ii) If A is true, then the negation of A is false.

These conditions seem to express the idea that the negation of a sentence A expresses the opposite of A.

Note that the two conditions are not expressed as bi-conditionals. So they do not give us *Tertium Non Datur* [nothing requires A to have a value at all in (i) or (ii)].

If "false" and "not true" are the same, then the two conditions just give us the ordinary truth table:

A	¬A
T	F
F	T

Take the Liar again, using the instance (9) of convention (T) and the two conditions on negation we can reason thus:

• Given (9) if "True(1)" is true, then "¬True(1)" is true. If "True(1)" is false, then because of (i) "¬True(1)" is true, and so, because of (9), "True(1)" is true. So: *in both cases* are "True(1)" and "¬True(1)" *both* true. That is meant by true contradictions.

• Both of them, however, are also *false* (not-true) at the same time: If "True(1)" is false, then because of (9) "¬True(1)" is false. If "True(1)" is true, then because of (ii) "¬True(1)" is false, and so is "True(1)" given (9). That is also meant by true contradictions. Contradictions are true in as much as they are true and false at the same time.

When does some connective [¬] codify negation? What does the concept of negation entail anyway?

The negation in many paraconsistent logics do separate falsity from non-truth. Nevertheless there are paraconsistent logics (e.g. **LP**) the negation of which satisfies the conditions (i) and (ii). These negations (respectively the connective [¬]) can claim to capture our intuitive concept of negation. They do so at least to a significant degree; there may be other requirements on negation, but which should these be?

Some claim that *ex contradictione qoudlibet* is essential to a real negation. Then certainly many paraconsistent logics have no negation at all. But *ex contradictione qoudlibet* obviously is highly *counter*intuitive.

Some claim that certain logical truths define 'real' negation, maybe *Tertium Non Datur* or negation-introduction:

$$A \supset \neg A \vdash (\neg A)$$

or maybe contraposition:

$$A \supset B, \neg B \vdash (\neg A)$$

All this, however, is controversial. Some of these theorems hold in some systems of paraconsistent logic.

Are Natural Languages Closed?

(ad II) in the dialetheist's agenda:

The natural languages we use seem to be semantically closed. We can express the antinomies in these languages. Natural languages have the resources of naming and corresponding self-reference. Semantic properties can be expressed in natural languages. *Prima facie* natural languages, therefore, are the very paradigm of universal languages.

If somebody wants to deny this, he has the burden of proof. Our ordinary conception of our language had to be seriously mistaken then! Usually the arguments against the semantic closure are based just on the antinomies.

One may further claim that neither do we speak nor do we need semantically closed languages in any science.

I will argue against this in (ad III).

Independent from that philosophy cannot restrict itself to non-universal languages. The language of philosophy has to be semantically closed. Philosophy does not want to deal only with the structure or conditions of talking in some specific language or languages of some kind, but aims at a theory of the basic structures and conditions of having a language *in general*. This requires the corresponding resources to express the universal claims.

There may not be a hierarchy of languages so that we always talk in a *last* meta-language the semantic properties of which cannot be made clear, except in a further turn of the screw (a new meta-language ...).

Universal theories of meaning, truth, knowledge etc. were not to have if we can talk only from some meta-language "down" to some distinct object-language. A general statement like

(K) Knowledge is true belief.

would be not well-*formed*.

But these are the very theories that philosophy is after. And notwithstanding their lip-service to hierarchy solutions of the antinomies most philosophers propose their *general* theories of meaning, truth, belief, reference, knowledge etc. They are right to do the latter, since we have such universal concepts.

There seems to be *no* crucial difference between formal languages and natural languages with respect to the properties being of interest here (i.e. semantic and structural properties), although formal languages have no native speakers, mostly no pragmatics, no socio-linguistics ...

We can investigate and formalize the logical structures of any natural languages. That is one of the central tenets of logic and formalization.

We not only talk about properties of all (natural) languages, it seems even incoherent whether there could be two completely incommensurable languages. Such a system could never be identified as a language at all. Our concept of language, therefore, involves unity and universality. There has to be a set of properties defining what a language is. These properties are preserved in change or translation.

Elucidating these properties and making them explicit from our intuitive understanding of language(s) is the traditional understanding of (transcendental) philosophy (of language).

Without semantic closure we would not be *able* to elucidate a concept that we seem to have!

Assessing the Alternatives

So I take it that we need semantic closure. No-one, but dialetheism seems to be able to deliver it. [ad (III) of the dialetheist's agenda]:

There are two major alternatives to paraconsistent treatments of the antinomies:

(IIIa) many-valued semantics or truth-value gaps,

(IIIb) the hierarchy of semantic meta-languages (in Tarskian style).

Both fail, but for different reasons.

(IIIa)-type solutions solve some antinomies, but the linguistic resources they employ in their formal framework are sufficient to generate new versions of the antinomies.
(IIIb)-type solutions result in an outrageous pragmatic self-contradiction.
Let us turn to (IIIa) first:
It has been claimed that the problems with the antinomies show that that the crucial sentence has no truth value whatsoever or some further value besides "true" and "false".
One may criticize this move as *ad hoc*: a further truth value is assumed just to avoid the antinomies. Since there are more general arguments to the failure of bivalence (arguments from referential failure, vagueness, indeterminacy) a proponent of this view can defend himself as merely applying a framework that is around to solve other problems with bivalence to the problem of the antinomies.
Antinomies, however, are no problem tied to bivalence. If the Liar sentence (λ) is taken as neither true nor false the reasoning concerning sentence (1) does not go through, that is right. The problem is that the *linguistic framework* employed to solve this antinomy is sufficiently rich to allow for new versions of the old antinomies, like the Liar.
What is a "linguistic framework"?
I take a linguistic framework to consist of both the language defined as well as its meta-linguistically expressed distinctions and semantics. Three-valuedness, for example, is part of some linguistic framework, given the interpretation of formulas in some semantics.
Antinomies can be avoided by having no logic at all. Keeping our power to infer from premises usually some distinctions are introduced to argue that the antinomies lack some crucial property or are on the wrong side of some crucial border.
A general *hypothesis*:

> A linguistic framework rich enough to avoid some the antinomies, generates its own versions of them.

This hypothesis is rather vague, but it can be illustrated on several versions to deal with antinomies.
With respect to three-valued approaches, one can introduce a *strengthened* Liar, (λ‘). If the (old) Liar (λ) is neither true nor false, it is not true. A three-valued framework has two values (say "false" and "undetermined") opposite to "true". But now we can *say* that (λ) is not true. This gives us a strengthened Liar:

(λ‘) (λ‘) is not true.

We argue again by cases:

(a) If (λ‘) is true, then (λ‘) is not true (by convention (T), see above)

(b) If (λ‘) is false, then (λ‘) is not true, because falsity is opposed to truth, and being non-true is what (λ‘) says of itself, so (λ‘) is true.

(c) If (λ‘) is undetermined, then (λ‘) is not true, so (λ‘) is true.

Since:

(P3) Not-True(A) ≡ False(A) ∨ Undetermined(A)

we get by Dilemma [that both alternatives (b), (c) of the right side of (P3) applied to (λ‘) imply "True(λ‘)"] a *new antinomy*:

(10) True(λ‘) ≡ Not-True(λ‘)

A framework involving three truth values involves the implicit or explicit validity of some principle like (P3). Even if "Not-True" is not *introduced* as a truth-value within the semantics, the semantic framework has *the expressive power* to introduce this notion. What we do here is to re-introduce a bifurcation within the realm of the truth-values. "Not-True" works like the notion of falsity in ordinary bivalent semantics. (This forced bifurcation is always available.)

The strengthened Liar, therefore, might reside on the language level at which we can express "Not-True", and given the framework there has to be *some* such level. The only way out would be to resort to some strict separation of language levels [i.e. to type (IIIb)-solutions we argue against below].

The trick introducing the strengthened Liar is independent from the number of truth-values present. Thus (IIIa)-type solutions won‘t work!

No Levels

What is wrong with the hierarchy solution [ad (IIIb)]?

The main problem is not that at first sight we cannot find these levels in ordinary language. That is a problem, since the assumption of these levels would mean that there is some hidden syntactic structure with no corresponding surface structure, although it is decisive for truth! Such an analysis would be a major revision of our understanding of our language!

Nonetheless there are even deeper problems in store:

The main problem is that if there were these levels and if the theory was true, the very statement of the theory and its given explanation (speaking *in general* about *all* language levels) would be impossible. The hierarchy conception says we are always talking from some level in the hierarchy, and *at the same time* makes a general statement to the effect that constructing the semantics of a language (level) we just go one level up.

Let's look at the details:

Suppose we had a hierarchy of truth-predicates. Each of them has an index *i*. Ascribing truth would be like:

(11) "p" is true-at-level-n (in L).

Now if indexes are numbers, we are able to talk about them in a sufficiently rich language, able to talk about its own *syntax*. So we can built for each sentence A the expression "the-level-of-A".

Consider then the supposedly semantic available sentence:

(12) (12) is not true-at-level-of-(12).

This sentence has to occur *somewhere* in the hierarchy on pains of the inexpressibility of some truths. Let us call this level "true-12"; then "true-at-level-of-(12)" is identical to "true-12". Convention (T) applies at every level of the hierarchy for sentences of that level:

(13) True-12(A) if and only if A.

Applying (13) to (12) as a sentence at level 12 we get:

(14) True-12(12) if and only if (12).

And this yields a new antinomy:

(15) True-12(12) if and only if not true-12(12).

Being able to talk about the indexes re-introduces antinomies in the context of semantic vocabulary. The only solution to resist the argument above is to give up the assumption that (12) occurs somewhere in the hierarchy, has a level. Given that (12) could not even be formulated.

Then, however, no sentence that talks in this general fashion about indexes is possible. The following sentence would be impossible too:

(16) The truth predicate of a level n is defined at level n+1.

This sentence, however, can be expressed in natural language, which are said to be captured. And (16) better be expressible if the theory of language levels is to be expressed at all! – If sentences like (16) were impossible the whole theory of (IIIb) would be *inexpressible itself*! Being unable to make statements about the hierarchy in general means that we are unable to understand the basic semantic concepts at all:

Either somebody who understands terms like "meaning", "denotation", "truth" etc. is able to express the building principles of definitions at some level, but this is impossible as we have just seen with sentence (16).

Or he has to have available the *whole completed hierarchy* in his mind, but this surely is not possible given that the human mind is a finite representational system and the hierarchy being infinite.

So the hierarchy conception leaves us hanging in the air concerning our being able to understand semantic concepts at all!

The hierarchy conception says we are always talking from some level in the hierarchy, and *at the same time* makes a general statement to the effect that constructing the semantics of a language (level) we just go one level up.

This is not just a contradiction. A contradiction one might suppose is what a dialetheist is to accept anyway. The situation is worse. If the theory is true – and, of course, as an adherent of the theory you believe it to be true – something is impossible *to do*, but you just do it! Given the theory the adherents are doing something impossible in the strict sense. It is not just a miracle. Suppose you believe in God: God in His glory might choose to perform miracles to authorise His revelations. Even He, however, cannot do what is logically impossible (as the Scholastics argued). He cannot make the square round. He cannot make the hierarchy solution work. What is done by the adherent of the hierarchy solution is completely mysterious.

I call this fact the MYSTERY.

Doing the impossible seems to me even more bizarre than claiming that some contradictions are true. Even if you don't think so, the failure of the hierarchy solution is obvious: Dialetheism provides semantic closure – so it claims – at the cost of true contradictions; the hierarchy model implies either the MYSTERY or a true contradiction *and* gets nothing for that, only the absence of semantic closure!

Tarski himself has not applied his hierarchy model to natural language, he even seems to concede that "truth" for natural languages is contradictory. One should, however, try if one is a Tarskian about formal languages to apply the construction to natural languages to have universal theories about language. The problem with the MYSTERY is that this *extended* Tarskian approach does not work.

The *particularist* Tarskian approach of only defining truth for some languages at least involves natural language statements like:

(17) We define the semantics of a language in its meta-language.

Now if natural language is ultimately considered inconsistent the last meta-language of this approach is inconsistent, and what then should these adherents of consistency think about the success of its workings?

Dialetheism avoids MYSTERY at the price of accepting some true contradictions. One may introduce some methodological principles:

(MT) Whatever implies MYSTERY cannot be true.

This surely is somewhat vague, but the idea should be clear. If some argument leads us back to the hierarchy (i.e. MYSTERY), there has to be something wrong with it. This in a way is *ad hoc*, but one can consider corresponding discussions as a context of discovery which logical principles of some paraconsistent systems are still too strong. To restrict (MT) we need another methodological principle:

(MR) Avoid MYSTERY at minimal costs.

This is an extension of a more general idea in constructing paraconsistent logics. The revised logic should come as close as possible to classical logic to keep as many of the theorems we like as we can get.

Liars without Semantic Expressions

As a note to the foregoing it should be mentioned that semantic closure is a *sufficient* condition for antinomies, but that it isn't a necessary condition in that sense that one can have forms of the Liar which do not use semantic concepts like "truth", "satisfies", "denotes", but *propositional quantification* and negation.

Let's take "λ" to be another name for sentence (1):

(1) $(\forall p)((\lambda \equiv p) \supset \neg p)$

Then – by the definition of "λ" – we have:

(2) $\lambda \equiv (\forall p)((\lambda \equiv p) \supset \neg p)$

From this an antinomy ensues: Inserting (1) for "p" in (1) we can detach with (2) and have the negation of (1) as well as (1).

Propositional quantification is also antinomic if coupled with truth *operators* (not just a truth *predicate*).