Logica Trianguli, 1, 1997, 73-91

FITCH´S PROBLEM AND THE KNOWABILITY PARADOX: LOGICAL AND PHILOSOPHICAL REMARKS*

CONCHA MARTINEZ, JOSE-MIGUEL SAGÜILLO, JAVIER VILANOVA

Abstract
Fitch´s problem and the "knowability paradox" involve a couple of argumentations that are to each other in the same relation as Cantor´s uncollected multitudes theorem and Russell´s paradox. The authors exhibit the logical nature of the theorem and of the paradox and show their philosophical import, both from an anti-realist and from a realist perspective. In particular, the authors discuss an anti-realist solution to Fitch´s problem and provide an anti-realist interpretation of the problematic statement "It is knowable that r is known and yet unknown". Then, it is argued that the knowability paradox has a solution even if one adopts a realist point of view. The authors provide a solution that takes into account the ambiguity of the term 'knowability' by deploying a temporal possible world semantics for epistemic modalities.

Introduction

This paper puts forward an epistemic conception of paradoxes that takes seriously the fact that paradoxes are participant-relative to an individual´s or community´s ***state of beliefs***. Under this framework the paradox of knowability is analysed and its logical and philosophical underpinnings are examined. It is divided in four sections. The first provides a conceptual framework based on the ontic-epistemic-doxastic distinction. In this presentation logical objects and their properties find their natural place as a first step for the clarification of the term 'paradox', to be defined in the next section. Section two introduces a definition of 'paradox' based on the previous conceptual framework. Section three illustrates the historical fact that often the very same argumentation which is a theorem for a person, is also a paradox for somebody else. This special feature is illustrated in the case of Fitch´s problem and the knowability paradox; the discussion of this point

---

* This paper is funded by research projects XUGA 20506B96 of the Galician Government and DG PB95-0863 of the Spanish Government.

has been structured into three subsections which are entitled ***the truth-class theorem***, ***Fitch's problem*** and ***the knowability paradox***. Section four puts forward a solution to the paradox of knowability. It is shown that the argumentation involves a fallacy of ambiguity regarding the term 'knowability'. The ambiguity vanishes when the common-sense meaning of the word 'knowability' is brought to light by means of a temporal possible world semantics for epistemic modalities.

1. Logical objects and the ontic-epistemic-doxastic distinction

An argument in the present paper is a two-part system composed of a set of propositions **P** and a single proposition **c**. Strictly speaking either **P** logically implies **c** or **P** does not logically imply **c**. In the first case the argument is valid. In the second case, the argument is invalid. Validity and invalidity, under the present account, are intrinsic, non-epistemic properties of arguments.

Human beings have developed methods to get to know whether a given argument is valid or invalid. Argumentations are at the core of the procedures to discover the validity of a given argument. For example, we may deduce **c** from the set of premises **P** through a chain of reasoning **R**. That is to say, to develop an argumentation is a human activity that establishes that **P** implies **c** by means of a chain of reasoning that makes it evident to the agent/audience that the information-content of the conclusion was already contained in the information-content of the premise-set.

Now, every proposition is either true or false but not every proposition is either known to be true or known to be false. Truth and falsity are extrinsic properties of propositions in the sense that it is the world that is relevant for a proposition to be true or false. On the other hand, known to be true and known to be false are extrinsic properties of propositions in a double sense: elliptical reference is made both to an agent and to an extra-propositional reality.

A proposition that is neither known to be true nor known to be false is called a hypothesis. If we deduce by means of an argumentation a given hypothesis from premises all known to be true, then in addition to establish the validity of the argument involved we also prove the conclusion; i.e., we also establish the hypothesis to be true. This is the deductive method of settling a hypothesis on the basis of propositions that we already know. So a proof is a three-part system composed of a set of premises all known to be true, a conclusion and a chain of reasoning.

On the other hand, if we deduce a false proposition from a set of premises all known to be true augmented with a given hypothesis, we establish that the hypothesis is false. This is the hypothetical-deductive method of settling a hypothesis. (See [3], Corcoran 1989:18-19). So a deduction is a three-part system composed of a set of premises (without further specification as to our epistemic attitude towards them), a conclusion and a chain of reasoning.

The success of both methods is based on two logical principles:
1. The principle of *truth and consequence*: Every proposition implied by a true proposition is true.
2. The principle of *falsehood and consequence*: Every proposition that implies a false proposition is false.

It follows from this approach that the genus argumentation contains as species proof and deduction, and every proof is a deduction but not every deduction is a proof. Once again and more precisely, an argumentation is a three-part system composed of a set of premises **P**, a conclusion **c** and a chain of reasoning **R**. Subspecies of argumentation are proof and deduction, where if we prove a given proposition, *a fortiori*, we deduce that given proposition, but not conversely.

It is crucial to our present concern to notice that there are also argumentations that lead, not to knowledge of what is the case with respect to a given proposition, but rather to doubt and hence to suspend and to re-examine our previous states of belief. This is the case with paradoxical argumentations. As we shall see in the next section, a paradox is an argumentation involving beliefs (*stricto sensu*) of a given agent or a given community of thinkers with respect to the premise-set, the conclusion and the chain of reasoning of the argumentation involved.

2. Definition of "Paradox"

Now, a paradox is an argumentation that establishes a conclusion believed to be false (perhaps even contradictory). The proposition believed to be false is the conclusion of a chain of reasoning which is believed to be cogent and which is based on a set of premises all believed to be true. This is clearly a functional definition involving doxastic components.

The above use of the propositional attitude word “believed” makes clear the fact that a paradox is participant-relative. This participant-relative property should be spelled out with respect to a given audience´s state of belief concerning the premises, the conclusion and the chain of reasoning of the argumentation involved. Note that the audience may be an individual or a community. The fact that paradoxes are participant-relative indicates that an argumentation which is a paradox to a certain **X** at a certain time **T** may not be a paradox for that person at a later time.

No false proposition is implied by premises all true. Consider an argumentation **A** with premise-set **P**, chain of reasoning **R** and conclusion **c**. Suppose argumentation **A** is a paradox for a given community **X** at a given time **T**. Then at **T** at least one of the three following conditions is satisfied:

1. Conclusion-Condition: The conclusion believed by X to be false is actually true.
2. Premise-Condition: At least one of the premises believed by X to be true is actually false.
3. Derivation-Conclusion: The derivation of the conclusion from the premises, which is believed by X to be a cogent derivation, is actually fallacious.

It is clear that under the present account, there is no paradox in an ontic sense; i.e., there is no such a thing as a paradox *per se*. 'Paradox' is a relational term involving an argumentation, in the sense defined above, and an agent. This last point suggests that if we believe in the possibility of knowledge, then paradoxical argumentations are by nature *transient* logical objects in the sense that they involve a mistaken belief, as opposed to proofs which are *enduring* logical objects in the sense that they involve knowledge. So, the criterion for classifying argumentations that is suggested here is based on the propositional attitudes **X** may have with respect to the premises, the conclusion and the chain of reasoning of a given argumentation.

3. Theorem or paradox?

The previous definition of 'paradox' points out the fact that paradoxes are pragmatic objects relative to the state of beliefs of an agent. Moreover, it has been established by historical research that the set-theoretical argumentation known as **Russell's paradox**, was previously a theorem for Cantor himself. A few years before Russell, Cantor proved the **uncollected multitudes theorem**, which is an indirect proof that shows the falsity of the comprehension principle. Latter on, Russell had the paradox whose conclusion is a contradiction and whose premise is the comprehension principle. Fitch´s theorem and the knowability paradox follow the same pattern. First, a theorem was proved by Fitch [9] and later on a paradox was developed by Hart [10]. And in this case there is still a third element that has been denominated Fitch´s problem.

3.1. Theorem (Fitch 1963)

Let us start with what we will call from now on the truth-class theorem or Fitch´s theorem. As Fitch (1962) puts forward, this theorem has a very general character, and its meaning is very simple. It talks about any set of propositions C such that if all the propositions in the set are true, in symbols " p (p C → p)", and if the set is closed under conjunction elimination, in symbols " p q ((p q) C → p C q C)", then the proposition "r r C" is not a member of the class C. In symbols,

⊢ p (p C → p) p q ((p q) C → p C q C) → ¬ p ((p p C) C)

PROOF

- 1 p (p C → p)
- 2 p q ((p q) C → p C q C)
3 p (p p C) C
4 (r r C) C
5 (r r C) C →(r C (r C) C) E 2 2
6 r C (r C) C M.P. 5,4
7 (r C) C → r C E 1
8 (r C) C E 6
9 r C E→ 7,8
10 r C E 6
11 r C r C I 9,10
12 Contradiction
13 E 3, 4-12
14 ¬(p (p p C) C) I¬ 3-13

Example:
Preliminary Definitions:
Def. 1: In order for an object **x** to be a counter-example for a given proposition **p**, it is necessary and sufficient for it to make **p** false.
Def. 2: In order for a universal proposition to be false it is necessary and sufficient for it to have a counterexample.

(*) Every true proposition is known to be true.

According to def. 2 (*) is false since it has a counterexample. Consider any hypothesis (i.e., any proposition not known to be true and not known to be false). However, it is logically impossible to have a known counter-example.

Proof**:** Let r be a true proposition which is not known to be true. By def. 2, r is a counter-example for (*). Now, let us suppose that r is a known counter-example. Since knowledge distributes in conjunction, it follows that r is known and that it is known that r is not known. The elimination of knowledge give us that r is known and not known. Q.E.D.

In the context of the polemic realism-anti-realism, the truth-class theorem is used in two different ways to raise both what is known in the literature as "Fitch´s Problem" and what is known as "The paradox of knowability". These two uses of the theorem will be analysed in the following subparagraphs. In both cases the truth class involved in the aforementioned theorem is identified with the class of known propositions ("K"$\equiv_{df}$ " C"); in the following pages and to simplify notation, we will use "K" that should be read either as "it is known" (= "belonging to the class of known propositions") in the context of the knowability paradox, or as "it is verified that" in the context of Fitch´s problem.

3.2. "Fitch´s Problem"

The first use of the theorem intends to show how from an apparently "reasonable form of verificationism" ("every true proposition is knowable in principle", in symbols " p (p →◇Kp)", an obviously silly form of verificationism follows ("every true proposition is known", " p (Kp → p)". This conclusion obtains when the premise that says that if p is the case, then it is possible to know that p, (in symbols " p (p → Kp)"), is added to the set of premises involved in the truth-class theorem; in short, Fitch´s problem amounts to showing that the following argument is valid.

- 1 p (Kp → p)

- 2 p q (K(p q) → Kp Kq)
- 3 p (p →◇Kp)

⊢ p (p → Kp)

From the point of view of those who pose the problem, the validity of this argument implies that the anti-realist concept of truth is untenable. According to them, Fitch´s problem relies on a correct interpretation of the anti-realist thesis, and it plays the part of a *crucial* experiment. Hence, facing this result, the anti-realists should at least reconsider their thesis.

Let us check whether the argumentation that corresponds to Fitch´s Problem is valid:

PROOF:

- 1 p (Kp → p)	
- 2 p q (K(p q) → Kp Kq)	
- 3 p (p →◇Kp)	
4 p (p ¬ Kp)	
5 (r ¬Kr)	
6 ◇K(r ¬Kr)	M.P. 3, 5
7 ◇(Kr K¬Kr)	M.P. 2, 6
8 ◇(Kr ¬Kr)	M.P. 1,7
9	Contradiction
10	E 4, 5-9
11 ¬ p (p ¬Kp)	I¬ 3-10
12 p ¬(p ¬Kp)	Neg. , 11
13 p (p → ¬ ¬Kp)	Interdefinition →, 12
14 p (p → Kp)	Interchange, E. ¬¬, 13

From the point of view of the designers of the problem it is quite clear that the silly form of anti-realism follows, but the anti-realists, of course, do not agree.

Timothy Williamson (1992) shows that what is described in the literature as «Fitch´s Problem», is not a problem from the point of view of intuitionistic logic. Williamson claims that his result establishes that sophisticated anti-realism cannot be reduced to silly anti-realism. He analyses the anti-realist notion of *verifiability/knowability in principle* into a notion of possibility, ◇, and a non-mathematical operator: it is verified that, K. So, for any proposition p, “p is true iff it is possible that it is verified that p”, in symbols " $p\ (p \leftrightarrow \Diamond Kp)$". Then Williamson defines an axiomatic system of modal intuitionistic logic and shows that the silly form of verificationism does not obtain in such a system. That is to say, he shows that, $p\ (p \rightarrow Kp)$ does not follow from,

- 1 $p\ (Kp \rightarrow p)$
- 2 $p\ \ q\ (K(p\ \ q) \rightarrow Kp\ \ Kq)$
- 3 $p\ (p \leftrightarrow \Diamond Kp)$

if the logical constants are understood intuitionistically. For the sake of simplicity we will maintain the proof of $p(p \rightarrow Kp)$ that has been given above; the only difference will be that the intuitionistic version of the natural deduction system will be applied. In such a version, the formula that represents the trivial form of verificationism, $p(p \rightarrow Kp)$, does not follow. The question is that if the logical constants are understood intuitionistically the last of the inferences in the proof, from $p(p \rightarrow \neg\neg Kp)$ to $p(p \rightarrow Kp)$, does not obtain because the double negation elimination rule does not hold in intuitionistic logic.

Williamson himself points out several interesting questions that the given formalization of Fitch´s Problem in an intuitionistic system of modal logic raises; by analysing those issues, Williamson intends to show to what extent the intuitionistic system of modal logic he defines succeeds in capturing the creative subject theory (Dummett, 1977: 335-59). The creative subject theory is one of the most popular anti-realist ways of accounting for the meaning of mathematical statements.

1) One of the questions he raises has to do with a new objection that the realist might pose to the anti-realist; the realist could object that by the rule of contraposition and by the intuitionistic validity of " $p(\neg\neg\neg\ p \rightarrow \neg p)$", " $p(\neg Kp \rightarrow \neg p)$" clearly follows from " $p(p \rightarrow \neg\neg Kp)$". From a classical point of view this last theorem, " $p(\neg Kp \rightarrow \neg p)$", seems as strange as the one that expresses the so-called silly form of verificationism, but, it says something quite different if the logical constants are understood intuitionistically. From an intuitionistic point of view, the theorem means that the possibility of ever verifying "Kp" can only be eliminated if the possibility of ever verifying "p" is also eliminated. The important point here is that intuitionistic negation of "Kp" means that "$Kp \rightarrow$ " , and since "$Kp \rightarrow p$", "$\neg Kp$" obtains only if "$\neg p$" obtains.

2) But the most interesting difficulty that arises in Williamson´s proposal has to do with the interpretation of the verifiability operator “K”. “ p $(Kp \quad \neg Kp)$” is not a theorem of Williamson´s modal system (Williamson, 1992). This author argues that this is a positive aspect of the system since, according to the theory of the creative subject, this should be so. The theory of the creative subject is a mathematical theory that identifies the truth of a statement with *its verifiability/knowability in principle*. A statement is considered *verifiable in principle* if there exists a mathematical procedure that provides a mathematical solution but that *does not* imply that the mathematical solution can always be applied by a human being.

We consider that in order to correctly understand the theory there are three different properties of statements that have to be distinguished:

(a) A statement is undecidable if there is not, at a given stage, a mathematical procedure that can be applied in principle in order to establish whether a given statement is the case or not. At those stages in which there is no procedure defined, the given statement is neither true nor false; as Dummett says,

> “This is not because the ´not´ which occurs in ´... is not true´ or ´...was not true´ is non-constructive: we may reasonably view it as decidable whether or not a given statement has been proved at a given time. But though constructive, this is an empirical type of negation, not the negation that occurs in statements in intuitionistic mathematics.” (Dummett, 1977: 337).

An example of undecidable statement is Goldbach´s Conjecture.

(b) A statement is decidable if, at a given stage, there is a mathematical procedure that might or might not be effectively applied by a human being, but that from an ideal point of view could be applied. Dummett poses the following example:

> “For example, a statement that a certain large number is prime is decidable, and may, when we apply the decision procedure, turn out to be true. I was making the tacit assumption that it is already determinate how the decision procedure will turn out, because there is no room for any play in the process of applying it. Hence if it would turn out that the number is prime, the statement that it is prime is, on this definition I gave, true, even though we have at present no proof that it is, and may never

have one, though we possess what is in fact an effective means for constructing one." (Dummett, 1997)

(c) A statement is humanly-tractable if at a given stage there is a mathematical procedure that can be effectively applied by a human being, so that the statement is verified in fact. For example: "2+2=4" or "At the moment we have a job at the Department of Logic of the University of Santiago de Compostela".

So, according to the proponents of the theory of the creative subject the gap between the concept of truth and the concept of empirical verification, corresponds to the difference between decidable statements and what might be called humanly-tractable statements. In fact the three different sorts of statements described above are discriminated on the basis of two different concepts; these concepts are taken from logic-mathematics and computer science: the concept of Turing Machine (as a representative of all its theoretical equivalents) and the concept of computational tractability.

A Turing Machine is an ideal machine; a given mathematical statement is considered decidable if there is Turing machine that describes a method such that if that statement is the case, the Turing machine stops to assert that the statement is the case, and if that statement is not the case, the Turing machine stops to assert that it is not. A given statement is undecidable if the impossibility of constructing a Turing machine that stops to assert that something is the case or to assert that it is not, is established.

When mathematicians and logicians first tried to construct theorem-provers by implementing logical systems on computers, they intended to construct a theorem-prover that behaved like a Turing machine whenever a given statement, formulated in the language of the given logical system, was decidable. To their surprise it was not possible; there is a theoretical gap between what is decidable and what is computationally tractable. The existence of a decidable procedure is not a sufficient condition for the construction of a computationally tractable algorithm. The reason is that the concept of Turing machine is defined for an ideal infinite tape without time limitations, while the concept of computational tractability takes into account space and time in order to analyse the complexity of the different algorithms.

Now, the theory of the creative subject is ideal in the same sense in which the concept of decidability is, while the notion of *human*-tractability takes into account our epistemic limitations so it is the analogue of the notion of computational tractability. So, when Williamson asserts,

> "... K must sometimes give an undecidable result, even when applied to a decidable statement such as the one about the number of tennis balls in my garden on the 4th July 1990. Yet the intuitionistic theory of the creative subject assumes that what is known at a given stage is a decidable matter, and the assumption seems a reasonable one, at least to a first approximation" (Williamson 1992: 71-72),

he is asserting that "K" is not an empirical operator, while he starts his paper asserting that it *is* an empirical operator (Williamson, 1992: 64). We understand that the difficulty is due to the fact that at the beginning of the paper he identifies "K" with "it is verified" when he asserts that it is an empirical operator, while he asserts latter on that,

> "..for the intuitionist, a question has a right answer if we have a decision procedure for it, but we are not required to carry out that procedure in real time." (Williamson, 1992: 72).

We consider that, in order to solve this difficulty with the interpretation of "K", it could be said that "Kp" means that "It is decidable that p", and in that case the knowability principle, " $p(Kp \rightarrow p)$", does not hold, since it is not true that if a sentence is decidable then it is true; if a statement is decidable then it is true that it holds or it does not hold. This means that Fitch´s problem is not at all a problem because one of the premises is false. The class represented by "K" is not a truth class, so the truth-class theorem does not apply and Fitch´s problem does not arise. The problem remains how to extend this mathematical theory to empirical statements, but that is a different matter.

Our conclusion is then that Fitch´s problem is not a problem for the anti-realist and that that conclusion obtains on the basis of two completely different arguments:

1) On one hand, Williamson shows that the chain of reasoning that leads to the silly form of verificationism is not correct if classical logic is substituted by intuitionistic logic; nevertheless, from our point of view (the point of view of those who do not consider that classical logic should be substituted by intuitionistic logic) this solution to the problem is not wholly satisfactory.

2) On another hand, we consider that if the theory of the creative subject is formalized adequately, the meaning of "K" does not correspond to a truth-class, and one of the premises, " p(Kp → p)", is clearly false.

But both solutions rely on an anti-realist theory and we intend to show that even from a realist perspective there is not such a thing as Fitch´s problem or the knowability paradox.

3.3. The *knowability paradox (Hart 1979)*

In the polemic between realism and anti-realism, the truth-class theorem has been used to design what has been called *the knowability paradox*. In this case, a third statement is added to the three premises of Fitch´s theorem (3.1). The added premise says that there is a proposition which is true but unknown. When such a statement is incorporated to the premises, a contradiction follows. Please note that while in Fitch´s problem this formula is an assumption that is cancelled later on in the course of the deduction (line 10), in the knowability paradox the formula is a premise. Otherwise, the two deductions run parallel until the point in which a contradiction is deduced. Then, in the case we are considering (the knowability paradox), a contradiction obtains, while in the proof of Fitch´s problem (3.2) the negation of the assumed formula, " p(p ¬Kp)", follows by applying the negation introduction rule, and finally the derivation of the formula that expresses the silly form of verificationism, " p(p → Kp)", obtains.

PROOF:

- 1   p (Kp → p)
- 2   p q (K(p   q) → Kp   Kq)
- 3   p (p → ◇Kp)
- 4   p (p   ¬Kp)

5	(r ¬Kr)	
6	◇K(r ¬Kr)	M.P. 3, 5
7	◇(Kr K¬Kr)	M.P. 2, 6
8	◇(Kr ¬Kr)	M.P. 1,7
9		Contradiction
10		E 4, 5-9

It is important to note that the contradiction follows only if line 4 is considered as a true premise. So we arrive to the central point in the discussion between realism and anti-realism: for a certain sort of realist the knowability paradox in not such since one of the premises, " p (p → ◇Kp)", is false; but according to what has been said in 3.2, the anti-realist (or at least a certain kind of anti-realist) does not consider the content of line 4 as a premise but as an assumption. So, once we have come to this point, the question is why certain kind of a realist considers that line 4 is a premise for the anti-realist, while the anti-realist (or certain kind of an anti-realist) takes it to be just an assumption.

Well, the realist understands logical constants in classical terms, while the sort of anti-realist that accepts Williamson´s solution is interpreting them in intuitionistic terms. So, when the realist says that the anti-realists accept " p(p ¬Kp)" as a premise, the realist understands that the anti-realist claims that "there is a knowable proposition that is not known", where the realist interprets the existential quantifier classically. That means that, from the point of view of the realist, asserting " p(p ¬Kp)" does not commit him to exhibiting an example of a true but not known to be true formula. (We have posed precisely this case as an example of Fitch´s theorem in Section 1). But in intuitionistic logic the condition for correctly asserting an existential statement such as " p(p ¬Kp)" amounts to having a proof of anything that is a proof of "(r ¬Kr)" for any proposition "r", that is to say, the possibility of exhibiting a proposition that, in this case, should be correctly asserted but that is not known to be the case. But such a thing is a contradiction *in termini* and that is exactly what Fitch´s theorem puts forward according to the intuitionist.

We maintain that the realists who design the paradox are not being fair to the anti-realists; to assert such a thing does not mean that we have any special sympathy to the anti-realists thesis. What we intend to put forward is that:

1) the argumentation (the intended paradox) formulated in natural language does not sound paradoxical at all;
2) the problem comes up when the argumentation is translated, in a certain way, into a formal language.
3) So, in the next paragraph we provide what we understand is a more adequate translation of the natural language argumentation into a formal language.

4. Solving the paradox

Solving a paradox is likewise participant relative; the process of solving a paradox is the changing of the participant´s state of belief so a paradox is no longer a paradox for **X**. *Solving* a paradox amounts to forming a conclusive diagnosis concerning what is wrong in the argumentation according to conditions 1, 2 and 3 above, and to transforming and reclassifying it as one of the following:

A. A proof of the conclusion.
B. A proof of the negation of the conjunction of the premises.
C. An important step to an inconsistency proof.
D. A fallacy.

The process involved in solving a paradox can be described as follows: Suppose an argumentation **A** is a paradox for **X**. Let us assume that **X** checks the chain of reasoning and finds it to be cogent; i.e., **R** actually shows that **P** implies **C**. Since the conclusion **C** is believed to be false by **X**, then given the *law of falsehood and consequence* ("Every false proposition is implied by a false proposition") **X** may think that the problem, the mistaken belief, is with respect to at least one of the premises in the premise-set. So **X** undertakes the task of re-examining **P** in a way that **X** regains responsible belief in the premises. Once **X** does this, C is no longer believed to be false by **X**. So what at certain time was a paradox for **X** is now a proof of its conclusion and this realization is based on two processes:

a- Coming to know that the conclusion follows from the premises-set.
b- Coming to know the truth of the premises.

A different situation arises when **X** verifies the cogency of **R** and confirms that the conclusion is correctly believed to be false. Again, by the *law of falsehood and consequence* **X** realizes that the premise-set contains at least one false premise; this amounts to reclassifying the paradox as a proof of the negation of the conjunction of its premises.

A special subcase in this situation arises when the conclusion is believed contradictory. In many contexts the paradoxical argumentation involved here is called “antinomy”. By the *law of falsehood and consequence* **X** realizes that his/her premise-set is inconsistent and therefore the argumentation can in principle be transformed into an inconsistency proof for the theory.

Finally, it might turn out that **X** discovers that the problem lies in the chain of reasoning. In this case, at least one of the steps of immediate inference leading from the premises to the conclusion involves a mistake or even a gap in the reasoning. The difficulties here can be both complex and multifarious and under the present account, resolving the paradox amounts to reclassifying it as a fallacy of a certain kind.

Our diagnosis as to what went wrong in the knowability paradox focuses on the chain of reasoning. Particularly, a detailed analysis of the term 'knowability' shows that it is an ambiguous term. Moreover, in our view, the concept of knowability does not obtain an adequate treatment if it gets formalized only in terms of the possibility operator over an epistemic modality. We make this point clear by considering that the inconsistency of the premise-set of the paradoxical argumentation is not evident at all.

-1 Every known proposition is true.
-2 If the conjunction of two propositions is known, then each of its members is known.
-3 Some true proposition is not known.
-4 Every true proposition is knowable.
?
5 Let **r** be a true proposition which is unknown.
6 It is knowable that **r** is true and unknown.
7 It is knowable that **r** is true and it is knowable that **r** is unknown.
8 It is knowable that **r** is true and yet unknown.

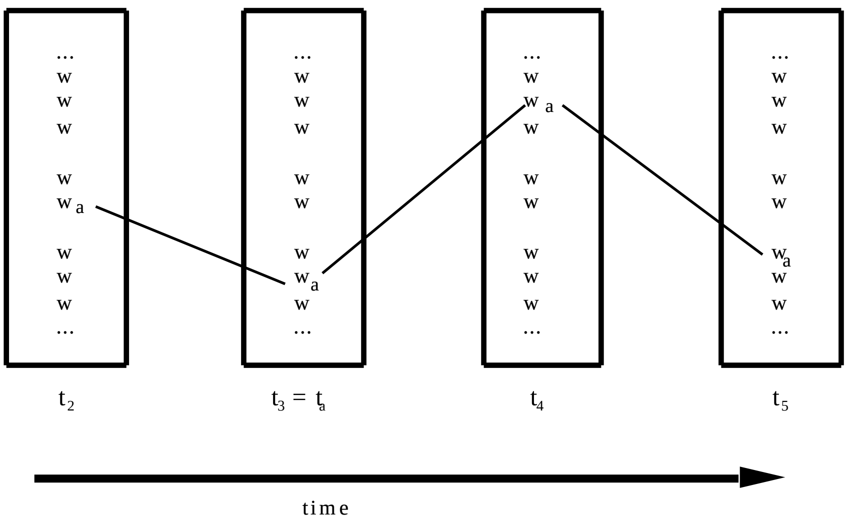
The difficulty in perceiving the problem in the ordinary English formulation of the argumentation is perhaps a symptom that indicates the ambiguity of the term 'knowability'. More precisely, it is our contention that the correct analysis of the term 'knowability' requires careful attention to its *temporal* meaning. Intuitively, to say in ordinary English that a proposition is knowable, means that for a given -unknown at the present time- proposition, there exists a future time in which that proposition is known. Specifically, if the proposition under consideration is that a given proposition "**r**" is not known at the present time, to say that that proposition is knowable means that there exists a future time in which it is known that **r** was not known at (back then) the present time. In symbols, we have that line 7 and 8 should read as follows:

5 $[\neg Kqt_a \ \& \quad t(t>t_a \quad Kqt)] \quad [\neg Kqt_a \quad t(t>t_a \quad K(\neg Kqt_a)t)]$

6 $[\neg Kqt_a \quad t[(t>t_a \quad Kqt) \quad \neg Kqt_a]$

A suitable formal semantics for our analysis of the term 'knowability' can be provided by a temporal modal frame for epistemic modalities according to the following definitions:

A *temporal epistemic frame* is a structure <W, T, <, $\{\approx_t\}_{t\ T}$, w_a, t_a>, such that W is a non empty set of possible worlds, T is a non empty set of temporal indexes, < is a linear order defined on T, $\{\approx_t\}_{t\ T}$ is a family of equivalence relations on W such that for every t T, $w_a \approx_{t\ T} w'$ iff w_a and w' share the same time slice, and where w_a and t_a are designed elements of W and T respectively. The idea is that, given a temporal index t, the set of worlds w' such that $w_a \approx_{t\ T} w'$, constitute the set of situations compatible with the state of our knowledge at the moment t. We can represent the formal semantics by using a diagram:



The truth-clause for the epistemic operator "K" is given as follows:

> "KA" is true at <w_a, t> iff "A" is true at <w',t> for all w' W.
> "A is knowable" is true at <$w_{a,}$ t> iff t > t_a such that for every w' W such that $w_a \approx_{t\ T} w'$, 'A' is true in <w',t>.

By using the epistemic temporal concepts just defined we can rewrite the argumentation under discussion and it is easy to see that no contradiction is deduced:

PROOF:

- 1 p t (Kpt → p)
- 2 p q t(K(p q)t → Kpt Kqt)
- 3 p (p → t>t_a (Kpt))
- 4 p t (p ¬Kpt_a)

5 (r ¬Krt_a)
6 t>t_a (K(r ¬Krt_a)t) M.P. 3, 5
7 t>t_a (Krt K(¬Krt_a)t) M.P. 2, 5
8 t>t_a (Krt ¬Krt_a) M.P. 1,7

Conclusion

We have provided a definition of paradox which makes it a species of the genus argumentation. In addition our definition points out the fact that paradoxes are relative to the state of beliefs of X in a given time T. In other words, a given argumentation that is a paradox for an individual or community of thinkers may not be a paradox for a different individual or community of thinkers. According to this sense of the term ‘paradox’ we have explained that, some anti-realists, such as Williamson, do not see themselves as facing any ´knowability paradox´ (as Hart poses it) though they recognise they have to face and solve, and they do, the so-called “Fitch´s problem”. Nevertheless, realists such as Hart consider that the knowability paradox is a paradox for the anti-realist. We have provided a solution to the knowability paradox which reflects on the fact that the belief on the cogency of the chain of reasoning is mistaken. Hence, the solution proposed consists in reclassifying the paradox as a fallacy of ambiguity concerning the term ‘knowability’. We propose a meaning for the term ‘knowability’, which is made clear by the definition of a temporal epistemic frame.

1) It is interesting to point out that the argumentation discussed is not a paradox -in the previous sense of paradox- for the realist logician, since realists believe there are true propositions which are not knowable. However, we believe our solution is also useful for the realist in a double sense: On one side there is a different sense of the term ‘paradox’ which is also common in the current literature and which could pertain to the realist view. If

'paradox' is defined as an argumentation whose premise-set is believed to be consistent, whose chain of reasoning if believed to be cogent and whose conclusion is believed to be contradictory, we may suggest that the realist has a paradox in this new sense (i.e., the set of premises is not obviously inconsistent to the realist eyes at first sight) and hence the solution proposed would also be most welcome for him.

2) On the other, since they claim that the knowability paradox constitutes a problem for the anti-realist (though we doubt there is any anti-realist who considers the knowability paradox itself as such) the provided solution to the paradox should be acceptable to the realist who considers the anti-realist has a problem. In such a case, what intended to be a crucial experiment against the anti-realist thesis is far from achieving its goal.

University of Santiago de Compostela
lflpcmav@usc.es, lflgsagu@usc.es, vilanova@usc.es

REFERENCES

[1] CORCORAN, J. (1973): "Gaps between logical theory and mathematical practice", Mario Bunge (ed.), *The Methodological Unity of Science*. Reidel Pub. Co. Dordrecht-Holland.

[2] CORCORAN, J. (1987): Review of Garciadiego, Alejandro. "The emergence of some of the nonlogical paradoxes of the theory of sets, 1903-1908", *Historia Mathematica*, 12, 4, pp. 337-351.

[3] CORCORAN, J. (1989): "Argumentations and logic", *Argumentations*, 3, pp. 17-43.

[4] COZZO, C. (1994): "What can we learn from the paradox of knowability", *Topoi*,13, pp. 71-78.

[5] CHURCH, A. (1956): *Introduction to Mathematical Logic*, Princeton U. P. Princeton, N. Jersey.

[6] DUMMETT, M. (1977): *Elements of Intuitionism*, Oxford University Press, London.

[7] DUMMETT, M (1997): "Is the concept of truth needed for semantics?", C. Martínez, M. U. Rivas & L. Villegas, (eds) , *Truth in perspective*, Avebury, forthcoming in 1997.

[8] EDGINGTON, D. (1985): "The Paradox of Knowability", *Mind* , 1985, pp. 557-68.

[9] FITCH, F.B.(1963): "A Logical Analysis of Some Value Concepts", *The Journal of Symbolic Logic*, 28, 2, pp. 135-142.

[10] HART, W. D. (1979): "The epistemology of abstract objects", *Proceedings of the Aristotelian Society*, vol. 53, pp. 153-65.

[11] KVANVIG, J. (1995): "The Knowability Paradox and the prospects for anti-realism", *Nous*, 29, 4, pp. 481-500.

[12] QUINE, W.O. (1962): "The ways of paradox", *The Ways of Paradox and Other Essays*, New York, Random House, pp. 3-20

[13] SAGÜILLO, J. M. (1994) "Paradoxical Argumentations", *Analysis and Evaluation*, Van Eemeren et al. eds., Amsterdam, pp. 13-22.

[14] TARSKI, A. (1944): "The semantic conception of truth", *Philosophy and Phenomenological Research,* 4, pp. 341-375.

[15] TARSKI, A. (1969): "Truth and Proof", *Scientific American,* 220,6, pp. 63-77.

[16] THOMASON, R.H.. (1984) : "Tense and modality", D. Gabbay and F. Guenthner (eds), *Handbook of Philosophical Logic* , vol II, pp. 136-166.

[17] VAN HEIJENOORT, J. (1967): "Logical Paradoxes", *Encyclopaedia of Philosophy*, 5, New York. Macmillan Pub. Co., pp. 55-51.

[18] WILLIAMSON, T. (1982): "On the paradox of knowability", *Mind*, 96, pp. 256-261.

[19] WILLIAMSON, T. (1990) : "Two incomplete anti-realist modal epistemic logic", *Journal of Symbolic Logic* , 55, pp. 297-314.

[20] WILLIAMSON, T., (1992): "On intuitionistic modal epistemic logic", *Journal of Philosophical Logic* , 21, pp. 63-89.