Moorean Sentences in Update Semantics[1]

Igor Sedlár

Comenius University, Bratislava
Slovak Academy of Science, Bratislava

Abstract: We outline a novel solution to Moore's paradox within the framework of update semantics, which explains Moorean absurdity in terms of non-cohesiveness. It is argued that, unlike the outlined solution, Gillies' treatment of the paradox within this framework is not satisfactory.

Keywords: epistemic modalities, Moore's paradox, update semantics.

The present paper undertakes to outline a new solution to Moore's paradox (hereafter MP) in the framework of update semantics (US). The paper is organised accordingly. The first section introduces MP and mentions some of the notable solutions. Several widely accepted requirements for a satisfactory solution to MP are stated. The second section discusses epistemic modalities and the basic ideas underlying US and serves as a concise introduction to the framework. The third section begins with a rigorous treatment of US and introduces Gillies' solution to MP, see Gillies (2001). The fourth section examines the main weakness of Gillies' solution and outlines a new solution based on US. The concluding section sketches possible directions of future work.

---

1 Versions of the paper have been read at Tilburg University and the Institute of Philosophy of the Czech Academy of Sciences. My gratitude for comments and suggestions belongs to the audiences at the respective workshops, to J. Podroužek, and to two anonymous referees of this journal. Work on this paper has been partially supported by the Jan Hus Educational Foundation and the Foundation of the P.J. Šafárik University in Košice.

1 Moorean sentences and Moore's paradox

Moorean sentences[2] are sentences of either of the two following forms:

(1) p, but I don't believe that p,
(2) Not p, but I believe that p.

Sentences of the form (1) usually go under the name of *omissive* Moorean sentences, whilst those of form (2) are known as *comissive* Moorean sentences. Hence "It is raining, but I don't believe that it is" serves as an example of an omissive Moorean sentence, while "It does not rain, but I believe that it is" is an example of a comissive Moorean sentence.

Sentences of both types have an astounding feature: they seem to be *consistent* in the sense of being capable of describing an obtaining fact (*i. e.* of being true), yet it is *absurd* to assert them. It might be the case that it is raining outside right now and I do not believe that it is (because I am too wrapped up in writing this paper and I don't mind the outside world, for example). However, my assertion that this is the case would surely strike you as somewhat odd if not plainly unintelligible. This is the heart of *Moore's paradox*: How can a consistent sentence be absurd in this manner?

Many different answers have been proposed. There is a family of *pragmatic-logical* answers, based on the idea that assertions of consistent Moorean sentences imply inconsistent propositions. Hintikka, for example, claims[3] that x's assertion of a sentence A of type (1) pragmatically implies the proposition that x believes that A is true. Hintikka then goes on to demonstrate that this proposition is inconsistent if we assume that belief meets certain plausible conditions.

There are also many *philosophical* answers, based upon various ideas stemming from epistemology, speech act theory, philosophy of mind, the theory of practical rationality, etc. For a sample of these, see Green and Williams (2007).

Solving MP amounts to explaining the nature of the absurdity of asserting Moorean sentences (*i.e.* to explain the nature of Moorean ab-

2 Our discussion in this section builds upon Green and Williams (2007) which is the standard collection of papers related to MP.

3 See Hintikka (1962).

surdity). Researchers active in this area have identified at least three requirements that every satisfactory solution has to meet:

1. The solution applies to omissive and comissive Moorean sentences alike. It applies to every sentence that exhibits the form of absurdity in question.
2. The solution does not assume a special kind of act concerning Moorean sentences, e.g. assertion.
3. The solution does not presuppose questionable principles concerning belief, e.g. some of the principles of (normal) doxastic logic.[4]

2 Epistemic modalities and update semantics

Yet a new kind of answer emerged recently, see Gillies (2001). Its main point is the idea that *Moorean sentences are in fact inconsistent*. The tension between consistency and absurdity is merely a chimera.

How come? The notion of Moorean sentences being consistent is simply the result of applying an inappropriate concept of consistency, *i.e.* truth-functional consistency. A sentence A is truth-functionally consistent iff it might be true (hence iff it is satisfiable, in an informal sense). However, assertions of Moorean sentences contain assertions concerning belief;[5] hence the consistency of these sentences has to be judged accordingly. What is the appropriate concept of consistency to be applied here?

The right concept is approached via the notion of *epistemic modality*. Consider the following sentence

(3) It is possible that Heart of Midlothian did not win today's match.

Suppose Mark, a Hearts' loyal fan, saw today's match and witnessed a glorious victory. Yet he has to admit that (3) is true in the sense of *metaphysical possibility*. If Hearts' opponents would have had more

4 A good example is positive introspection: if one believes that p, then one believes that one believes that p. It is argued that the principle does not hold for actual human belief. Consequently, any solution to MP that presupposes the principle is fallacious.

5 „I don't believe that p" in (1) and „I believe that p" in (2).

luck or if some of Hearts' key players would have been injured, then Hearts could have lost or the match could have ended with a tie. In other words, today's victory was not predestined.

On the other hand, there is an important sense in which (3) cannot be true from Mark's viewpoint. He was there, only slightly drunk, and saw the victory with his own eyes. The information available to him rules out the possibility that Hearts actually did not win the match. Thus (3) is not true in the sense of *epistemic possibility*, at least not from Mark's viewpoint.

The example makes it clear that the modality "it is possible that" in (3) can be read at least in two ways.[6] The first (metaphysical) considers a broader set of possible states of affairs – those that are metaphysically or logically admissible. The second (epistemic) considers a narrower set – only the possible states of affairs that are consistent with the information available to our individual are admissible.[7] A sentence *A* is possible if it is true at least in one admissible state of affairs, but the set of admissible states varies with respect to the two readings.

Despite this common feature, epistemic modalities are formalised in a slightly different manner. A widely accepted framework is that of *update semantics* (hereafter US).[8] The crucial ideas behind US are simple: (*i*) individuals have information about their surroundings, (*ii*) the body of information available to an individual is changing with incoming information, (*iii*) a key feature of propositions is the way they affect one's body of information in case one accepts the proposition as true.

Perhaps a short explanation why US is considered to be more suitable for dealing with epistemic modals than the Kripke possible-world framework might be appropriate. For example, it is argued that conjunctions involving epistemic modalities are not in general com-

---

6 The notion that the difference between metaphysical and epistemic modality is a difference of „reading" is not generally accepted. See Yalcin (2007). I employ it for the sake of brevity.

7 The individual-relativity of epistemic modalities is not uncontroversial, see von Fintel and Gillies (2007).

8 The key publications are Groenendijk – Stokhof – Veltman (1997) and Veltman (1996).

mutative. The following example is taken from Gillies (2001).[9] Suppose that I am waiting for Ann and Bill, and I know that they have already left their homes and are on their way. The doorbell rings and I say "It might be Ann… It's Bill", where the dots indicate a pause where I open the door. This is a natural and perfectly intelligible thing to say. But changing the order of the statements results in "It's Bill… It might be Ann", which is barely understandable.

Before going on to formal details, let me mention the general strategy behind Gillies' approach to MP. He suggests that we should read the modality "I believe that *p*" in Moorean sentences as a shorthand for "It is not *epistemically* possible that *not-p*". This means that we should analyse Moorean sentences by means of US. When we do this, we observe the simple fact that Moorean sentences are inconsistent with respect to US (US-inconsistent). Finally, and this is the main moral of the story, *Moorean absurdity simply is US-inconsistency*. Hence there is no tension between consistency and absurdity – there is no paradox at all. As will be apparent from a more detailed discussion in a moment, I do not share Gillies' conviction that Moorean absurdity amounts to (or simply boils down to) US-inconsistency.

3 Moore's paradox in US

Now I set the stage in a rigorous manner.[10] Let us begin with a Boolean language over a countable set of propositional variables *Pr* together with primitive connectives ~ ("not") and ∧ ("and"). We add a unary operator *M* ("it is possible that"). The set of formulas over this language is defined in the usual recursive way. A *possibility* is a subset of *Pr*, hence the set of all possibilities is the power set of *Pr*. An *information state* is a set of possibilities, *i.e*. a set of sets of propositional variables.

Possibilities correspond to possible states of affairs if we take propositional variables to stand for "simple facts". For example $\{p, q\}$ corresponds to the state of affairs where the simple facts *p* and *q* obtain,

---

9 Gillies notes that his examples are very much like those in Beaver (1993), which are in turn inspired by some of Veltman's examples.

10 The definitions of the technical notions of US are essentially those of Veltman (1996).

while the other simple facts do not.[11] Information states are sets of possibilities. This mirrors the intuition that our information about our surroundings is partial. For example, I looked into my drawer and so I know that there is no cake in there, but I am not sure whether this is true of my colleague's drawer. So my information about my surroundings settles the fact of my having a cake in my drawer, but does not settle it in the case of my colleague. If p stands for "My colleague has a cake in his drawer" and q for "I have a cake in my drawer", and p and q are the only propositional variables in the game,[12] then my information state is $\{\{p\}, \varnothing\}$. This corresponds to my inner monologue: "Whether I like it or not, I see that I don't have a cake in my drawer which means that q does not obtain. But I'm not sure whether my colleague does have one. So there are two possibilities: either she doesn't have a cake (darn it all) which means that both p and q do not obtain, or she has one and only p obtains".

An *interpretation* is a function that assigns to every formula X a unary operation on information states, denoted $[X]$. If s is an information state, then $s[X]$ denotes the result of applying $[X]$ to s. The function has to meet the following conditions:

1. $s[p]$, for a propositional variable p, is the set of possibilities in s that contain p,
2. $s[\sim X]$ is the complement of $s[X]$ relative to s,
3. $s[X \wedge Y]$ is $s[X][Y]$,
4. $s[M\,X]$ is s if $s[X]$ is not $\varnothing$, else it is $\varnothing$.

A formula X is *acceptable in a state s* iff $s[X]$ is not the empty set. X is *consistent* iff there is a state s in which it is acceptable. X is *accepted* in s iff $s[X] = s$.

The "meaning" of an interpreted formula in general is the way the formula changes the information state of an individual that accepts the formula as true. If I accept a propositional variable p as true, I accept the basic fact it stands for. This means that I no longer consider possible the states of affairs in which the fact does not obtain. I erase them from my list of possibilities. Hence, in the case of the cake-example, if I look into my colleague's drawer and I find a cake there,

11 Technically, possibilities correspond to propositional valuations.

12 I am hungry, so other basic facts do not mater right now.

then I accept p as true and my information state becomes $\{\{p\}\}$. After accepting a negation of a formula X I get rid of all the possibilities that would survive the addition of X. For example, if I wouldn't find a cake in my colleague's drawer, I would drop the possibility in which the fact p obtains. Conjunction is handled as operation composition. One important consequence of this is that is in not in general commutative. This means that accepting a conjunction $X \wedge Y$ proceeds by accepting the left conjunct X first (that changes the initial information state) and then accepting the right conjunct Y (that brings about another change). This is in accord with the abovementioned main reason for adopting US.

Claims to the effect that X is (epistemically) possible (*i.e.* formulas $M\ X$) work as *tests on information states*. Possibility, being understood epistemically, amounts to acceptability with respect to one's information state. X is epistemically possible (acceptable) for me if my information state allows me to accept X. This means that, after accepting X, I *would* still have at least one possibility to consider. In US the empty set corresponds to the *absurd state*, where *every* possibility is ruled out.[13] This is a state that does not match any possible state of affairs.[14] Now X is acceptable in s iff I wouldn't end up in the absurd state were I to accept X.

It is illuminating to return to the match-example. If p represents "Hearts won today", then Mark's information state contains only possibilities that contain p (he is certain that Hearts won). Were he to accept $\sim p$, he would have to get rid of all the possibilities that contain p, but then he would end up without any possibilities at all. But this is absurd. Hence $\sim p$ is not acceptable for him. This mirrors the fact that for Mark (3) is not epistemically possible. Mark cannot incorporate $\sim p$ into his picture of the world.

The unacceptability of $\sim p$ for Mark is caused by the specific nature of his actual information state. But it is plain that there are propositions Mark could not have possibly incorporated into any information state, be it what it may. A nice example of these *US-inconsistent* propositions is provided by any contradiction $p \wedge \sim p$. *Any* information state

13 The empty set plausibly represents the absurd state only if our stock of propositional variables is not limited, unlike the cake-example.

14 Note that $\{\varnothing\}$ is not empty.

touched by the operation $[p \wedge \sim p]$ turns into the empty set. The heart of Gillies' solution of MP is the fact that Moorean sentences share this feature: *Moorean sentences correspond to US-inconsistent formulas*.

Define a unary belief operator B ("it is believed that") as $\sim M \sim$. Hence $B\ X$ means "It is not (epistemically) possible that $\sim X$". The reader can easily check that $s[B\ X] = s$ if X is accepted in s (*i.e.* $s[X] = s$), else $s[B\ X]$ is the empty set. Now we have the following

Proposition 3.1. *Formulas* $p \wedge \sim B\ p$ *and* $\sim p \wedge B\ p$ *are US-inconsistent.*

Proof. We state the proof for $p \wedge \sim B\ p$, the second case is similar. Consider an arbitrary information state s. The state $s[p \wedge \sim B\ p]$ equals to $s[p][\sim B\ p]$, which turns into $s[p] - s[p][B\ p]$. If s or $s[p]$ is already empty, then the proposition holds trivially. It is clear that in the remaining case $s[p]$ contains only possibilities that include p (p is accepted in $s[p]$). Hence $s[p][B\ p]$ equals to $s[p]$ and $s[p] - s[p][B\ p]$ is the empty set. *Q.E.D.*

Proposition 3.1. suggests that Moorean sentences are actually inconsistent. Hence the easy way out seems to be to explain Moorean absurdity *as* US-inconsistency. This is the main idea of Gillies' approach.

However, this explanation encounters its own problems. Consider, for example, the sentence

(4) I don't believe that it is raining, but it is.

It is simply the example of an omissive Moorean sentence (see Section 1) with a different order of conjuncts. I claim that (4) is absurd in the Moorean sense. Yet the corresponding formula $\sim B\ p \wedge p$ is not US-inconsistent. To see this, consider the state $s = \{\{p, q\}, \{q\}\}$. $s[\sim B\ p \wedge p]$ equals to $s[\sim B\ p][p]$ and $s[\sim B\ p]$ equals to $s - s[B\ p]$. But since p is not included in every possibility in s, $s[B\ p]$ is the empty set and $s - s[B\ p]$ equals to s. Thus $s[\sim B\ p][p]$ amounts to $s[p]$, which is $\{\{p, q\}\}$. This means that there is an absurd Moorean formula that is not US-inconsistent. Moorean absurdity does not boil down to US-inconsistency. This line of criticism stems from Green's interesting paper Green (2002).

To be sure, Gillies claims that there is a non-absurd reading of (4). It suffices to imagine a person that doesn't know about the weather outside and correctly claims "I don't believe that it is raining…". The person then takes a look outside, sees the rain, and continues with "… but (now I know that) it is". Information input causes the utterance to be perfectly alright.

An immediate objection (also due to Green) is that uttering (4) without meanwhile receiving new information is still absurd. Gillies claims in response[15] that uttering (and even considering) (4) is *impossible* without meanwhile receiving new information. This is a consequence of

Proposition 3.2. *There is no non-empty information state s such that* $s[\sim B\ p \wedge p] = s$.

Proof. Consider an arbitrary s. Clearly p is either accepted in s or it is not. In the former case $s[\sim B\ p]$ is the empty set, hence $s[\sim B\ p \wedge p]$ is also the empty set. In the latter case $s[\sim B\ p]$ is s, hence $s[\sim B\ p \wedge p]$ is $s[p]$. But since p is not accepted in s, $s[p]$ is a proper subset of s. *Q.E.D.*

The proposition means that there is no information state that remains unchanged when touched by the operation $[\sim B\ p \wedge p]$. Gillies calls such formulas *non-cohesive* and infers that (4) cannot be "processed" without receiving new information. But if this is the case, the non-absurd reading of (4) is mandatory and the sentence cannot be absurd.

4 A new and simple solution

To my opinion the link between the non-cohesiveness of $\sim B\ p \wedge p$ and the claim that (4) cannot be absurd is very loose and a more elaborate argument is needed. First, it is not clear what "processing" a formula (or a sentence) means. I do not think that Gillies intends it to mean "accepting as true". But the update operations corresponding to formulas are designed to mirror the effects of *accepting* the corresponding statements *as true*. Hence no conclusions concerning "processing" can be drawn from technical results in US. Moreover, it is clear that (4) is absurd even if one does not "accept it as true". The whole point of MP is that it *cannot* be accepted as true. In short, I do not consider Gillies' response to Green's objections to be satisfactory.

[15] See Gillies (2002).

Yet I believe that the right solution is not far away. Gillies has tried to defend the non-absurdity of (4) by pointing out its non-cohesiveness, but I think that one can get the right picture by turning his strategy upside down: *non-cohesiveness is precisely the technical counterpart of Moorean absurdity*. This makes sense as soon as you realize that non-cohesive formulas are *precisely those that are not accepted in any non-absurd state*. If every non-empty state can be thought of as *rational*, then clearly non-cohesive formulas are not accepted in any rational state. They are not rationally acceptable.

Remember the belief operator B and the fact that X is accepted in s iff $B\ X$ is. Thus no rational information state permits to believe in a statement represented by a non-cohesive formula. In other words, one cannot rationally rule out the negation of any non-cohesive formula.

The technical underpinning of our solution's main idea is the following

Proposition 3.3. *Formulas* $p \wedge {\sim}B\ p$, ${\sim}B\ p \wedge p$, ${\sim}p \wedge B\ p$, *and* $B\ p \wedge {\sim}p$ *are non-cohesive.*

Proof. The non-cohesiveness of $p \wedge {\sim}B\ p$ and ${\sim}p \wedge B\ p$ follows from the clear fact that every US-inconsistent formula is non-cohesive and from Proposition 3.1. ${\sim}p \wedge B\ p$ is non-cohesive according to Proposition 3.2., and the proof of the fact that $B\ p \wedge {\sim}p$ is US-inconsistent and hence non-cohesive is left as an exercise to the reader. *Q.E.D.*

Observe that if Gillies' response to Green's claim that his explanation does not cover all absurd propositions is not satisfactory, then his proposed solution to MP does not meet all of the requirements stated in Section 1. Our solution, on the other hand, applies to Moorean sentences independently of the order of conjuncts, as witnessed by Proposition 3.3. Our approach is in accord with the remaining two requirements as well. As to the second one, our explanation is purely semantic and does not presuppose any pragmatic acts involving Moorean sentences. Hence it cannot favour some of these acts (e.g. assertions) over others. As to the third one, no problematic assumptions concerning the operator B are stated. It behaves like a simple test operator on information states.

5 Conclusion

To sum up, we have demonstrated that Gillies' proposed solution to MP is not satisfactory and we have outlined a new approach based on US. The main idea is to explain Moorean absurdity as non-cohesiveness. We have seen that this solution meets the main requirements for any successful solution to MP stated in the literature.

Yet some points deserve further investigation. The important idea that Moorean sentences might be true is not touched upon in the present framework. Gillies dismissed it as misleading and there is also a technical reason for this. There is no room for truth in US, so to speak. The framework is designed to formalise the effects of accepting information *as true* and has no means to tell "as true" from *true*. However, I think that there is at least a partial remedy for this. Suppose we work with a *poly-modal* version of US, containing belief operators B_i for every agent i form a given set of agents. Consider an agent x and his intuition that a Moorean sentence "p and x does not believe that p" might be true. Such a sentence corresponds to $p \wedge \sim B_x p$. The intuition amounts to the fact that there is an agent y such that $p \wedge \sim B_x p$ is accepted in y's current information state. In other words, some agent can rationally believe that p is true but x does not believe that it is. This approach employs a more elaborate notion of information state based on non-well-founded set theory, see Aczel (1988), and is neatly discussed in Gerbrandy's PhD. thesis (Gerbrandy 1998).

I have two more attractive problems to mention. First, are there US-consistent but non-cohesive formulas besides the formulas corresponding to Moorean sentences? Second, there is also the issue whether the class of non-cohesive formulas is definable only by means of US. In other words, is there a (normal) "static" modal doxastic logic L that proves $\sim B\ X$ iff X is non-cohesive?

Department of Logic and Methodology of Sciences
Faculty of Philosophy
Comenius University
Gondova 2
814 99 Bratislava
Slovak Republic
sedlar@fphil.uniba.sk

Institute of Philosophy
Slovak Academy of Science
Klemensova 19
813 64 Bratislava
Slovak Republic

REFERENCES

ACZEL, P. (1988): *Non-Well-Founded Sets*. CSLI Lecture Notes. Stanford: CSLI Publications.

BEAVER, D. (1993): Two birds and one stone. In: Kamp, H. (ed.): *Presupposition*. DYANA Report 2.2.A. Edinburgh: Center for Cognitive Science.

GERBRANDY, J. (1998): *Bisimulations on Planet Kripke*. PhD Thesis, ILLC. Amsterdam: University of Amsterdam.

GILLIES, A. (2001): A new solution to Moore's paradox. *Philosophical Studies* 105, No. 3.

GILLIES, A. (2002): Moore thoughts on update semantics. Central division meeting of the APA, Chicago.

GREEN, M. (2002): Moorean absurdity and update semantics. Central division meeting of the APA, Chicago.

GROENENDIJK, J. – STOKHOF, M. – VELTMAN, F. (1997): Coreference and modality. In: Lappin, S. (ed.): *The Handbook of Contemporary Semantic Theory*. Oxford: Blackwell.

HINTIKKA, J. (1962): *Knowledge and Belief*. Ithaca: Cornell UP.

VELTMAN, F. (1996): Defaults in update semantics. *The Journal of Philosophical Logic* 25, No. 3.

VON FINTEL, K. – GILLIES, A. (2007): An opinionated guide to epistemic modality. In: Gendler, T. S. – Hawthorne, J. (eds.): *Oxford Studies in Epistemology 2*. Oxford: Oxford UP.

WILLIAMS, J. – GREEN, M. (eds.) (2007): *Moore's Paradox: New Essays on Belief, Rationality, and the First Person*. Oxford: Oxford UP.

YALCIN, S. (2007): Epistemic modals. *Mind* 116, No. 464.