Comments on *Many-Valued Semantics for Vague Counterfactuals* by Marco Cerami and Pere Pardo

ONDREJ MAJER

As the title suggests, the authors propose a framework for analyzing vague counterfactuals in the framework of fuzzy logics. Although both subjects have been studied for quite a long time, Cerami and Pardo present one of the first attempts combining them (this topic is also independently studied by Běhounek and Majer).

1 Counterfactual conditionals

Many philosophers have attempted to provide logical analysis of counterfactuals (or natural language conditionals in general), most prominently Nelson Goodman, Robert Stalnaker and David Lewis. The most widely accepted solution currently is the Lewis-Stalnaker semantics, which is also the starting point for the authors.

There are several specifications of counterfactual conditionals based on grammatical form (subjunctive) and logical form (false antecedent). The authors specify the notion of counterfactual as conditional sentences whose antecedent is assumed to be false. This is certainly in accordance with the generally accepted convention, but it should be pointed out that there is no universally accepted criterion (sometimes we also use the subjunctive form for the situations which are not "contrary to the fact"). We also might not be sure about truth/falsity of the antecedent, as Lewis himself points out, and hence an adequate logical analysis of counterfactuals to also cover this case (the antecedent happens to be true after all).

The authors start with an exposition of the Lewis' sphere semantics, which is based on a standard possible-worlds semantics used in modal logics. Lewis assumes that a universe of possible worlds is equipped (on the top of existing structures like accessibility relation)with an ordering representing the similarity of a given world to the actual world (or a world which plays the role of the actual world). Formally, the similarity ordering can be represented by a system of spheres—a system of subsets of the universe which are nested (linearly ordered by inclusion), closed under unions and centered around the actual world. Although the sphere semantics is the one mostly referred to, Lewis provides several formalizations of the similarity relation in the terms of comparative possibility, selection functions (mentioned by the authors), degrees of similarity and some other notions and proves that each of them is equivalent to (some version of) the sphere semantics.

Lewis starts with the definition of a counterfactual saying that the counterfactual $A \mathrel{\Box\!\!\rightarrow} C$ is true at a given world, iff either there are no A-worlds (worlds in which A is true) or there is an AC-world, which is closer to the actual world than any $A\neg C$-world.

Another definition also discussed by Lewis employs the notion of the closest antecedent permitting sphere:

$A \;\Box\!\!\rightarrow C$ is true at a given world iff either there are no A-worlds or the material conditional $A \to C$ holds in the closest sphere containing an A-world.

This idea was formalized by Robert Stalnaker using selection functions and for many (including the authors) it seems to be more intuitively appealing than the Lewisian definition. However, the definition is less general as it relies on an assumption that many authors (including Lewis) consider restrictive and not very well justified. It is called the Limit assumption (LA) and requires that if a formula A is true in some world, then there always exists a closest sphere containing an A-world. In other words there is no non-well founded sequence of A-permitting spheres.

Cerami and Pardo argue in favor of the Limit assumption and claim that the case of non-well founded sequences "... is not often found in the classical framework, ... ", but this does not seem to be quite justified. One can find sentences from everyday communication which do not satisfy LA. Consider the sentence "If I were (strictly) taller than 2 meters, I would play basketball" (or any other sentence using strict ordering)—even if we reason classically, there is no closest world in which I am strictly taller than 2 m. From this point of view it seems quite natural to avoid the Limit assumption and the price to be paid—not using the notion of the closest antecedent-permitting sphere—does not seem to be too high.

In the discussion about the Limit assumption the authors say: "Another possibility discussed by Lewis is that the intersection of a set of spheres is empty." This might be misleading as Lewis allows an empty intersection just for a set of spheres, each of which is A-permitting (contains some A-world) for a particular formula A. The intersection of all spheres is always non-empty (centering condition), it either contains just the actual world (Strong centering according to Lewis) or it might contain some other worlds as well (Weak centering).

Lewis provides an axiomatization for several counterfactual logics. Different properties of the similarity relation are characterized by different axioms, the only exception is the Limit assumption. It is usually referred to in the literature as 'the Lewis' system of counterfactuals', having in mind his basic system VC but, it should be stressed, that this is not the only one.

The authors intentionally concentrate on the semantics and neither provide an axiomatization of vague counterfactuals nor discuss the relation of their semantic framework to Lewis' axiomatic system.

When discussing Lewis' approach to possible worlds the authors say: "Lewis refuses the idea that a possible world is merely a propositional evaluation of sentences, as in the tradition of frame-based semantics for Modal Logic. Nevertheless, he makes use of such a notion of possible world when he defines the syntactical calculus of counterfactuals."

Both claims are disputable. In my opinion possible worlds semantics for modal logic neither speaks nor has to speak about propositional evaluations—a possible world is a primitive entity ('truthmaker') at which a formula takes a truth value (either bivalent or multivalued) and is not straightforwardly identified with a propositional evaluation. It can perfectly happen in a modal frame that two worlds agree on all values, but they are still two different possible worlds. Strictly speaking Lewis does not use the notion of evaluation in the definition of syntactical calculus. In the completeness proof he builds canonical models over maximal consistent sets of sentences, but this is a purely technical use.

2 Vague counterfactuals

After introducing the necessary apparatus from fuzzy logic, the authors proceed to the definition of a counterfactual. In agreement with the standard view, they assume that the antecedent of a counterfactual is not true, which can be paraphrased in a many-valued setup as 'not completely true'. They provide several generalizations of the classical Lewisian approach based on this reading.

As in the classical case (with Limit assumption), the truth of a vague counterfactual $\phi \;\Box\!\!\rightarrow \psi$ depends on the behavior of the conditional $\phi \rightarrow \psi$ at the worlds, at which ϕ is (in some degree) true and which are closest to the actual world. Depending on the truth condition for the antecedent they provide three analyses:

1-*semantics* the closest worlds in which the antecedent is fully true

r*-*semantics the closest worlds in which the antecedent is true in the degree r

more than actual semantics the closest worlds in which the antecedent is more true, than in the actual world

The truth degree of the counterfactual in question is defined as the minimum of truth degrees of the conditional $\phi \rightarrow \psi$ in the closest worlds satisfying the corresponding condition.

1-semantics is a natural and intuitive generalization of the classical case; this is unlike the r-semantics, which seems to be of rather technical importance. The corresponding technical notion of an r-cut is used in mathematical fuzzy logic, but does not seem to have natural counterpart in natural language (we usually do not say "If I were at least 2/3-happy, life would be great."). I understood it rather as an auxiliary notion necessary for the *more than actual* semantics, defined as a limit of values of the counterfactual in question under r-semantics with r approaching the actual value.

The authors prove, that each of the proposed semantics correspond to the Lewisian one (more precisely to the system with Strong centering and Limit assumption). In particular, if the evaluation of propositions is bivalent, then the value of a counterfactual under the r-semantics (for the $r = 1$) coincides with the value in the sphere semantics.

3 Comments

Limit assumption From a formal point of view non well-founded sequences of truth values are a serious problem in the semantics of fuzzy logic (cf. the definition of the universal quantifier in predicate fuzzy logic and the problem of safe structures) so the Limit assumption seems to be a quite natural way to avoid them. However, the motivation for formalizing counterfactuals comes from natural language and its everyday use and from this point of view the Limit assumption is non-natural and restrictive, even in the classical bivalent case (see the discussion above). It would be interesting to observe if the Limit assumption is a necessary condition for the presented approach or if it is possible to avoid it (and still exclude the possibility that the truth value of some counterfactuals is undefined).

Crisp vs. vague approach to antecedent In the 1-semantics a counterfactual is evaluated with respect to the worlds, where the antecedent is strictly true (similarly true in the degree r in the r-semantics). On the other hand, one of the advantages of fuzzy solutions is that they are 'robust' in the sense that they take into account not only objects which (strictly) satisfy a certain condition (have a certain property), but also objects satisfying it 'roughly'. In the case of counterfactuals, not only the worlds where the antecedent is strictly true should count, but (to some extent) also the worlds in which it is almost true. Imagine an antecedent that is strictly true only in one world in the minimal sphere but in which there are several worlds which are true in the degree 0.99. Intuitively, they should also have some influence in determining the value of the vague counterfactual in question. Of course, a question then arises as to how to determine the dependence of the value of a vague counterfactual on the degree in which the antecedent is true. Similar reasoning can be applied to the minimality condition (strictly minimal vs. 'closed enough' to minimal).

The fuzzy paradigm seems to be reflected by *more than actual semantics*, where the 'close enough' worlds influence the value of a counterfactual via the notion of limit.

Comparisons with Lewis' approach The authors do compare their semantics with the sphere semantics, but it might be interesting to make a deeper comparison with the Lewisian approach.

Formal representation of counterfactuals is motivated by everyday use of conditionals, so an adequate formalization (either classical or many-valued) should reflect some common linguistic intuitions. Lewis thoroughly discusses the examples of use of conditional sentences, where some standard patterns of reasoning (valid for the material conditional) do not hold—he calls them counterfactual fallacies. His default examples are weakening (from $A \to C$ infer $A \wedge B \to C$) , contraposition (from $A \to C$ infer $\neg A \to \neg C$) and transitivity (from $A \to B$ and $B \to C$ infer $A \to C$). Although the authors give some examples which illustrate their own solution, it would interesting to show how their system deals with Lewis' fallacies.

This is closely related to the last point—the authors focus on the semantics of counterfactuals and show the correspondence to the sphere semantics It would be interesting to know if their semantics corresponds to the Lewisian axiomatic system.

Ondrej Majer
Institute of Philosophy
Academy of Sciences of the Czech Republic
Jilská 1
110 00 Prague 1, Czech Republic
Email: majer@flu.cas.cz