Many-Valued Semantics for Vague Counterfactuals

MARCO CERAMI AND PERE PARDO[1]

1 Introduction

This paper is an attempt to provide a formal semantics to counterfactual propositions that involve vague sentences as their antecedent or consequent. The initial work on counterfactuals, due to Lewis ([8]), was based on Classical Logic and thus it focused on crisp sentences. Lewis proposed a semantics based on the notion of possible worlds that captures alternative ways a world could be, including here our reality i.e., the world we live in (henceforth: the actual world, or our world). In Lewis' work, possible worlds can be seen as classical Boolean valuations, each valuation corresponding to a possible world.

Our aim is to generalize Lewis' system of spheres semantics based on Classical Propositional Logic, to a similar semantics based on many-valued logics. On the one hand, as long as the authors know, the claim that Lewis' systems of spheres provide a good semantics for counterfactual propositions does not seem to be controversial. On the other hand, however, many-valued logics are not universally accepted to be a good semantics for vague sentences (see [4] and [5] for an overview on different approaches to the subject). Even though this is not the right place to discuss whether many-valued logics are suitable tools for reasoning in presence of vagueness, we find necessary to spend a few lines in order to justify at least why our approach is based on this framework to formalize vague sentences.

Despite no general agreement exists on what a vague sentence is, there does not seem to be any problem in admitting that sentences involving predicates like *bald, young, near, tall* are vague sentences. A common feature of the above predicates seems to be the fact that they admit *borderline cases*, that make it not an easy matter to decide whether the sentences where they appear are either true or false. In our opinion the claim that every sentence has a truth-value (true or false) is a too strong claim indeed. Similarly for the claim that our inability to assign some truth-value (*true* or *false*) to some sentence must be a matter of *ignorance*, even when we happen to know the exact number of hair in a man's head, somebody's exact age and so on.

An important point in the defense of classical logic as the only tool for formalizing reasoning is usually that it is a very powerful tool. This is certainly true, but it is not a

[1]With partially support by Spanish MICINN projects FFI2008-03126-E/FILO (Eurocores-LogICCC Project LoMoReVI FP006), Consolider project AT (CSD2007-022), and ARINF TIN2009-14704-C03-03; by the Generalitat de Catalunya grant 2009-SGR-1434; and CSIC grant JAE Predoc 074. The authors wish to thank the two anonymous reviewers, Lluís Godo and Ondrej Majer for their helpful comments.

reason not to use, in some contexts, more refined tools: a chainsaw is a very powerful cutting tool, but this does not mean that somebody uses one to cut his nails.

On the other hand, t-norm based many-valued semantics is a robust and enough refining tool to handle vague propositions for several reasons. We report some of them:

1. It is simple and mathematically formalizable and this makes it suitable to be used in practical applications.

2. It does not presuppose ignorance, like epistemic theories on vagueness do, even when we have complete knowledge of the object under observation.

3. It gives an elegant solution to the famous Sorites paradox (see [7]) and this, for a theory, is a necessary condition to be a theory about vagueness.

4. Sentences that in other theories are considered as borderline cases can be handled like every other sentence.

In our context, then, vague sentences will be handled as many-valued sentences so we accept that there are sentences that can be neither (totally) true nor (totally) false.

As can be already seen from the classical definition of counterfactual, the antecedent of a counterfactual is false and not true at the same time. In a many-valued context where sentences can take intermediate values between true (1) and false (0), in contrast, a sentence can be not totally true (< 1) without being totally false (0). For this reason we prefer to generalize the definition of a counterfactual and define it as a conditional sentence whose antecedent is not 1-true. The present paper is an investigation on the consequences of this assumption with the aim of providing a semantics for vague counterfactuals.

As we will see later on, one immediate consequence of this new definition is a plurality of choices among possible definitions of truth-conditions for sentences like *If I were rich, I would be happy*. The requirement that the antecedent is not 1-true in the actual world forces us to choose among different sets of worlds where the antecedent is true enough to trigger counterfactual reasoning. This plurality of choices gives rise to different semantics depending on whether we demand that the antecedent is absolutely true (i.e. it has truth value 1) or relatively true (i.e. it has truth value >0) in other possible worlds (where, again, this antecedent must be more true than it actually is).

In the literature, the study of counterfactuals falls within the more general area of non-monotonic reasoning. Among related work, counterfactuals have been approached from distinct perspectives: uncertainty (see [10] for a probabilistic approach) and similarity (see [6]), but as far as we know, the present approach is the first study on counterfactuals from a fuzzy reasoning perspective. A main difference with the referred papers is that we focus on counterfactuals involving many-valued propositions, instead of just two-valued (or crisp) propositions.

The paper is structured as follows. In preliminary Section 2, we recall basic notions of Lewis' possible world semantics as well as the many-valued semantics based on the prominent t-norms: Łukasiewicz, Gödel and Product, with truth-values in the real unit interval $[0,1]$. In Section 3, we introduce vague counterfactuals and discuss general issues about them. We also present the basic language we will make use of.

The next sections introduce the different semantics for counterfactuals according to the three basic t-norms, as well as reducibility results among these semantics (within each basic t-norm) and interdefinability results in the Łukasiewicz case. In Section 4, we define a direct generalization of Lewis' semantics to the many-valued case, where antecedents are required to be 1-true, i.e. true to degree 1 within the real unit interval of truth-values $[0,1]$ generalizing the set $\{0,1\}$ for *false* and *true* in classical logic.

Then in Section 5 we consider a more general semantics requiring the weaker condition that the truth-value of the antecedent attains a fixed truth-degree higher than the actual one (its truth-value in our actual world). Finally, we present in Section 6 what we call *more than actual* semantics, requiring that the antecedent attains an arbitrary value higher than the actual one.

After these semantic results, we show in Section 7 the consequences of expanding the language with Baaz' projection operator Δ and truth-constants for degrees of truth (introduced in [1] and [9, 2, 3], respectively). In particular, we prove that these expansions may endow some semantics with the expressive power of any member of a more general class of semantics, of which the former is a particular instance. These results may be used to simplify the truth-conditions given by any semantics in the latter class.

2 Preliminaries

In this preliminary section, we introduce classical counterfactuals and t-norm based semantics for propositional fuzzy logics.

2.1 Counterfactuals and Lewis' semantics

As commonly understood in the literature, *counterfactuals* (like: if it were the case that φ it would be the case that ψ) are conditional sentences whose antecedent is assumed to be false. In contrast with classical material implication, where a false antecedent makes the implication vacuously true, a counterfactual with an actually false antecedent can still be false. Hence, material implication cannot model reasoning about how things would be like if the world was different.

The need to establish a semantics for counterfactual conditionals arises from the fact that there are contingent sentences. These are sentences that, though they might be actually false, they do not express impossible or contradictory facts. As such, we can think of a state of affairs, not much different from the actual one, where such sentences are contingently true. This can be done because their truth would not violate too much fundamental laws valid in the actual world. When these contingent sentences occur as antecedents (or consequents) of some material implication, their contingency makes the implication contingent as well. However, we may want to capture the intuitive correctness of conditional sentences between contingent facts: for instance, *If there were no carbon on Earth, there would be no life either.* The truth of this counterfactual is not established by their actual truth-values. But it can indeed be established on the basis of other considerations.

Two forms of counterfactuals occur as expressions in the natural language, here called *would*- and *might*-counterfactuals:

- would-counterfactuals: *if it were the case that* φ*, it would be the case that* ψ; this form expresses a necessity statement concerning any relevant world where φ is true; for the counterfactual to be true, ψ must hold at all these worlds.

- might-counterfactuals: *if it were the case that* φ*, it might be the case that* ψ; this is a possibility statement concerning relevant worlds where φ is true; a single such world making ψ true will account for this possibility thus making this counterfactual true.

One of the main attempts to provide a semantics for the truth-conditions of a counterfactual conditional has been David Lewis' [8]. In order to give a semantics for counterfactuals, Lewis imagines that we can think about *possible worlds*.[2] The possibility of talking about possible worlds arises from the fact that we can represent our world by means of contingent sentences which express facts surrounding us: "the sky is blue", "water is composed of oxygen and hydrogen" are examples of such sentences. A possible world is a consistent set of facts or propositions (also called a state of affairs), possibly with some of these facts not holding in our world (i.e. with sentences expressing these propositions being actually false). Mathematical or logical truths do not express contingent propositions (they hold across possible worlds -no matter how much different they are from our world), and hence these are not seen as expressing propositions that constitute possible worlds. We will denote the set of propositions that are believed to hold in our world by the expression *actual world*, while every other maximally consistent set of propositions will be considered as a possible world. The set of possible worlds will be denoted by W. The elements of W can be distributed into subsets of worlds which are more or less similar with respect to the world selected as the actual world. Lewis suggests to organize the set of possible worlds into *spheres*, or subsets of W. These spheres are totally ordered by inclusion, with outer spheres containing inner spheres. The actual world is located at the innermost sphere, also called the center. Outer spheres contain inner spheres plus those worlds that are different enough (w.r.t. the actual world) to be located in these inner spheres. This gives an account of the difference of possible worlds with respect to the actual world.

The definitions in this subsection are taken from Lewis [8].

DEFINITION 1 *A* system of spheres *on the set of possible worlds W, denoted by \$, is a subset of the power set of W which is totally ordered by inclusion, closed under arbitrary unions and non-empty intersections.*

It is possible to define several kinds of systems of spheres. In particular, Lewis considers the following additional properties of systems of spheres:

- *universal*: for every $w \in W$, there exists $S \in \$$ such that $w \in S$; i.e. every possible world is at least in one sphere. Equivalently, $\bigcup \$ = W$.

[2] Actually Lewis maintains that possible worlds do exist, but a long and tedious philosophical debate would be needed to define the meaning of the word *exist* in this context. We only will maintain that we can think about possible worlds and can use such idea for reasoning about counterfactuals or similar notions: in fact, conditional reasoning does not need the *real* existence of objects called "possible worlds".

- *centered*: there exists $w \in W$ such that $\{w\} = \bigcap\{S \in \$\}$; i.e. there is a unique world in the inner sphere. We will denote such a system by $\$^w$.

Intuitively, the elements of inner spheres are worlds not much different from the world(s) lying on the center of that system, while the elements of the outer spheres can be very different from the world(s) at the center.

The truth-value of a counterfactual in a system of spheres is defined in terms of the truth conditions of the formulas occurring in it. Boolean formulas are evaluated in a given world (taken as a propositional evaluation) in the usual way.

DEFINITION 2 *Consider the language* $L = \{\top, \bot, \neg, \wedge, \to\}$ *and the set* $\mathrm{Fm}(L)$*:*

$$\mathrm{Fm}(L) \quad := \quad \mathrm{Var} \cup \{\top, \bot\} \quad | \quad \neg\varphi \quad | \quad \varphi \wedge \psi \quad | \quad \varphi \to \psi$$

Let W be a set of worlds w. For each w, we define a propositional evaluation *e_w as a homomorphism* $e_w\colon \mathrm{Var} \to \{0,1\}$*, which assigns value 1 to each* $p \in \mathrm{Var}$ *iff p is true in world w. As usual, e_w can be inductively extended to every formula in* $\mathrm{Fm}(L)$*. We define the set of possible worlds satisfying φ, called φ-*worlds*,*

$$\begin{array}{rcl}
[\top] & = & W \\
[\bot] & = & \emptyset \\
[p] & = & \{w \in W \mid e_w(p) = 1\}, \textit{for each } p \in \mathrm{Var} \\
[\varphi \wedge \psi] & = & [\varphi] \cap [\psi] \\
[\varphi \to \psi] & = & (W \smallsetminus [\varphi]) \cup [\psi] \\
[\neg\varphi] & = & W \smallsetminus [\varphi]
\end{array}$$

The two former definitions are related by the equivalence: $e_w(\varphi) = 1 \iff w \in [\varphi]$*. We will consider the binary connective $\equiv$ as a connective defined by:*

$$\varphi \equiv \psi \quad := \quad (\varphi \to \psi) \wedge (\psi \to \varphi)$$

With such a semantics and notion of system of spheres, Lewis proceeds to define the truth conditions for counterfactuals as follows:

DEFINITION 3 *Consider the language* $L_{\Box\!\!\to} = \{\top, \bot, \neg, \wedge, \to, \Box\!\!\to, \Diamond\!\!\to\}$ *and the set* $\mathrm{Fm}(L_{\Box\!\!\to})$*:*

$$\mathrm{Fm}(L) \quad := \quad \mathrm{Fm}(L) \quad | \quad \varphi \Box\!\!\to \psi \quad | \quad \varphi \Diamond\!\!\to \psi$$

A would counterfactual $\varphi \Box\!\!\to \psi$ is true at world w (according to system $\$$*) iff either*

(1) no φ-world belongs to any sphere $S \in \$$, or

(2) some sphere $S \in \$$ does contain at least one φ-world, and $\varphi \to \psi$ holds at every world in S.

A might counterfactual $\varphi \Diamond\!\!\to \psi$ is true at world w (according to system $\$$*) iff both*

(1) some φ-world belongs to some sphere $S \in \$$, and

(2) every sphere $S \in \$$ that contains at least one φ-world contains at least one world where $\varphi \wedge \psi$ holds.

In this paper we will assume that systems of spheres are universal and centered. We will denote by $\$^w$ the system of spheres centered on the world w. Moreover, our actual world will be denoted by $w^\star$. Another possibility discussed by Lewis is that the intersection of a set of spheres is empty. This is the case when there is a non-well-founded decreasing sequence of spheres whose intersection is not an element of the sequence. Even if this case is not often found in the classical framework, Lewis proposes a new condition to exclude it. The *Limit Assumption* stipulates that, if a φ-world exists, then there always exists a closest sphere $S \in \$$ containing a φ-world, for any $\varphi \in \mathrm{Fm}(L_{\Box\!\!\rightarrow})$. In order to apply this assumption to the many-valued case, we present it in a more general form: If some sphere in $\$$ exists with some property P, then there exists a sphere $S \in \$$ having P such that for each other $S' \in \$$ having P, $S \subseteq S'$.

Indeed, in Definition 3 below, Lewis avoids relying on the Limit Assumption, that he considers somehow unnatural. But the cost of this move is to renounce to define a counterfactual as a conditional proposition that is true in the *closest* φ-sphere. Lewis presents another semantics to recover this notion of closest sphere (based on *selection functions*), though these produce unintuitive truth values for counterfactuals when the Limit Assumption fails (see p. 58 of [8]). In contrast, the semantics based on the notion of closest sphere (with the Limit Assumption) seems to be the most intuitive way to solve this issues, since it captures the idea of evaluating a counterfactual in the most similar worlds.

Finally consider counterfactuals nested in other counterfactual: *If it were the case that φ, then if it were the case that ψ then χ would be the case*. In the end, this composite counterfactual is evaluated w.r.t. the actual world, while the inner counterfactual, *If it were the case that ψ then χ would be the case*, is evaluated w.r.t. some φ-world(s). Thus, composite counterfactuals require different systems of spheres centered on different worlds, defined on the same universe W (and w.r.t. a fixed similarity relation among worlds). See Figure 1 for an illustration.

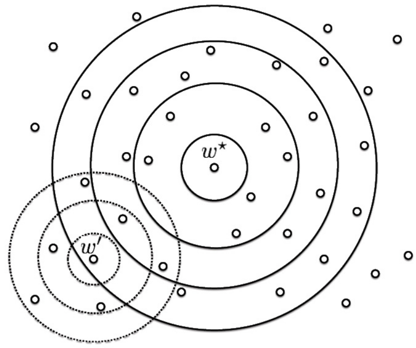

Figure 1. The semantics of counterfactual $\psi \Box\!\!\rightarrow \chi$, nested in $\varphi \Box\!\!\rightarrow (\psi \Box\!\!\rightarrow \chi)$, takes place in each closest φ-world, e.g. in w'.

Despite Lewis directly takes $\Box\!\!\rightarrow$ as primitive, defines $\Diamond\!\!\rightarrow$ from the former and negation, interdefinability can be proved from Definition 3.

LEMMA 4 ([8]) *Let φ, ψ be formulas, then the following formulas are tautologies:*

$$\begin{aligned} \varphi \mathrel{\Box\!\!\rightarrow} \psi &= \neg(\varphi \mathrel{\Diamond\!\!\rightarrow} \neg\psi) \\ \varphi \mathrel{\Diamond\!\!\rightarrow} \psi &= \neg(\varphi \mathrel{\Box\!\!\rightarrow} \neg\psi) \end{aligned}$$

Next we prove an equivalent characterization of the truth-conditions in Definition 3, suggested by Lewis (without proof). This is used later on to show that classical counterfactuals are a particular case of many-valued ones.

LEMMA 5 *Let W be a set of possible worlds with $w \in W$, and \$ a system of spheres satisfying the Limit Assumption. Then:*

(a) A counterfactual $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is true in world *w, iff either $[\varphi] = \emptyset$ or $w' \models \psi$, for every $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$.*

(b) A counterfactual $\varphi \mathrel{\Diamond\!\!\rightarrow} \psi$ is true in world *w, iff both $[\varphi] \neq \emptyset$ and $w' \models \psi$, for some $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$.*

Proof To prove the right-to-left implication of (a) suppose that $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is true in w. Then, by Definition 3, it holds either that (1) no φ-world belongs to any sphere $S \in \$$, or (2) some sphere $S \in \$$ does contain at least one φ-world, and $\varphi \to \psi$ holds at every world in S. The former (1) implies clause $[\varphi] = \emptyset$ and we are done. In case (2), let S' be a sphere (i) containing a φ-world and (ii) making $\varphi \to \psi$ true everywhere in S'. Since spheres are ordered by inclusion, by the Limit Assumption there exists a sphere S^* such that (i) $S^* \cap [\varphi] \neq \emptyset$, (ii) $S^* \subseteq [\varphi \to \psi]$ and (iii) S^* is $\subseteq$-minimal w.r.t. (i) and (ii). Now observe that $S^* = \bigcap\{S \in \$ \mid [\varphi] \cap S \neq \emptyset\}$:

($\subseteq$) Let S be such that (i) $[\varphi] \cap S \neq \emptyset$. Note that it cannot be the case that $S \subsetneq S^*$, because otherwise $S \subseteq S^* \subseteq [\varphi \to \psi]$, since S contains a φ-world and this is jointly incompatible with $\subseteq$-minimality of S^* w.r.t. (i) and (ii).

($\supseteq$) Let $w \in \bigcap\{S \in \$ \mid [\varphi] \cap S \neq \emptyset\}$. For any S'' with (i), (ii) and (iii) above, we have $S'' \in \{S \in \$ \mid [\varphi] \cap S \neq \emptyset\}$. In particular $S^* \in \{S \in \$ \mid [\varphi] \cap S \neq \emptyset\}$, so that $w \in S^*$.

Now, let $w' \in [\varphi] \cap S^*$. Since, by definition, $S^* \subseteq [\varphi \to \psi]$, we have $w' \in [\varphi] \cap [\varphi \to \psi]$. By modus ponens, $w' \in [\psi]$. Thus, it holds that, for every $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$, we have $w' \models \psi$.

To prove the converse implication of (a) assume that, for every w', $w' \models \psi$ is implied by $w' \in [\varphi] \cap \bigcap\{S \in \$ \mid S \cap [\varphi] \neq \emptyset\}$ (since the assumption that $[\varphi] = \emptyset$ trivially implies the claim). Let $S^* = \bigcap\{S \in \$ \mid S \cap [\varphi] \neq \emptyset\}$. By the Limit Assumption, we have S^* is a sphere. Let $w \in S^*$. If, on the one hand, $w \in [\varphi]$, then by assumption $w \in [\psi]$ so that $w \in [\varphi \to \psi]$. If, on the other hand, $w \notin [\varphi]$. Then trivially $w \in [\varphi \to \psi]$. In either case, we have both $S^* \cap [\varphi] \neq \emptyset$ and $S^* \subseteq [\varphi \to \psi]$, so we are done.

For (b), we have $\varphi \mathrel{\Diamond\!\!\rightarrow} \psi$ is true at w iff $\neg(\varphi \to\!\mathrel{\Box\!\!\rightarrow} \neg\psi)$ is true at w (by Lemma 4). This is equivalent to: it is not the case that either $[\varphi] = \emptyset$ or $w' \models \neg\psi$, for every $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$. Thus, in other words, both $[\varphi] \neq \emptyset$ and for some $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$ we have $w' \notin [\neg\psi]$. Therefore we know that $w' \in [\psi]$ for some $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$, and so the proof is done. $\square$

Informally, this characterization reads as follows. A counterfactual $\varphi \mathrel{\Box\!\!\rightarrow} \psi$ is true in w iff either φ is false in every world in W, or ψ is true in every φ-world lying in the closest sphere containing a φ-world. A $\varphi \mathrel{\Diamond\!\!\rightarrow} \psi$ counterfactual is true in w iff both φ is true in some world in W and ψ is true in some φ-world lying in the closest sphere containing a φ-world.

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 notion of possible world when he defines the syntactical calculus of counterfactuals. Since our approach is more a formal than a philosophical one, we will consider a possible world to be just a propositional evaluation of sentences. As we will see, in our case sentences need not be evaluated simply as true or false, but they can take intermediate values.

2.2 Vague sentences and Fuzzy Logic

During the last century there has been a great deal of work in the field of *Propositional Fuzzy Logics*, which are suitable to model reasoning with vague propositions (see Fermüller [4] for an overview of other formalisms). We give below a brief account of such logics (see [7] for details).

We consider a *vague sentence* as a sentence that, by the nature of its meaning, cannot be understood as merely true or false. As an example, if we fix that a tall man is a man whose height is greater or equal to 1.80m̃, we cannot consider a man who is 1.79 m tall as a short man, even if he is not tall. This is related to the well-known Sorites paradox. A way to overcome this paradox is to consider *fuzzy sets*.

As defined in [11] by L.A. Zadeh, a fuzzy set C is a set whose characteristic function,[3] χ_C is a function which returns a real value between 0 and 1, i.e. $\chi_C(a) \in [0,1]$. Intuitively, if C is the set of tall men and a is a man who is 1.79 m tall, then $\chi_C(a) = 0.98$.

The last example shows that, in contrast to classical frameworks where *false* and *not true* are equivalent, the distinction between these two notions becomes fundamental in many-valued frameworks. In the context of predicate sentences, i.e. sentences of the form *a is C*, we can directly adapt the tool of fuzzy sets to give a value (between 0 and 1) to *a being an element of C*. We can also consider the sentence "*a* is *C*" as a whole propositional sentence p, and say that its value lies between 0 and 1. The former approach is known as predicate or first order calculus and the latter approach, propositional calculus. Following the literature on counterfactuals, we will follow the approach which considers atomic propositions as the minimal language entities.

Giving a sentence a propositional value, however, is not a trivial issue: it often depends on the meanings of the predicate and the individuals occurring in it. In the context of natural language, the same predicate, say *high*, may ask for the use of different systems of reference, e.g. depending on whether we want to talk about a high mountain or a high skyscraper. In the former case, we can fix that the membership of a mountain to the set of high mountains has value 1 if the mountain height is greater than or equal to 6000 m, it has value 0 if its height is less or equal to 2000 m, and it has a value between 0 and 1 if its height falls between the given values. In the latter case, we can fix that the

[3] In the classical framework, a *characteristic function* of set C is a function χ_C such that for an individual a, $\chi_C(a) = 1$ if $a \in C$ and $\chi_C(a) = 0$ otherwise.

membership of a skyscraper to the set of high skyscrapers has value 1, if its height is greater or equal to 150 m, has value 0, if its height is less or equal to 60 m, and a value between 0 and 1, if its height falls between the given values.

DEFINITION 6 *A* t-norm *is a binary operation* $*\colon [0,1]^2 \to [0,1]$ *such that:*

1. $*$ *is commutative and associative,*
2. $*$ *is non-decreasing in both arguments,*
3. *for every* $x \in [0,1]$*, it holds that* $1 * x = x$.

If, $*$ *is a continuous mapping from* $[0,1]^2$ *to* $[0,1]$*, we talk about a* continuous t*-norm.*

A t-norm is normally understood as the function giving the truth value of conjunction in a propositional calculus. The basic examples of continuous t-norms are:

1. *Łukasiewicz* t-norm (denoted by Ł), defined by: $x * y = max(0, x + y - 1)$,
2. *Gödel* t-norm (denoted by G), defined by: $x * y = min(x, y)$,
3. *Product* t-norm (denoted by Π), defined by: $x * y = x \cdot y$.

Given a continuous t-norm $*$, we can define its residuum (i.e. the function which gives the semantics for the implication):

DEFINITION 7 *Let* $*$ *be a continuous t-norm, then its residuum is a binary operation* $\Rightarrow_*\colon [0,1]^2 \to [0,1]$ *such that, for every* $x, y \in [0,1]$*:*

$$x \Rightarrow_* y = \sup \{z \in [0,1] \mid x * z \leq y\}$$

Now we define the logic of some continuous t-norm over a propositional language.

DEFINITION 8 *Let* $*$ *be continuous t-norm. The language of a propositional t-norm based logic,* $\mathrm{Fm}(L_*)$ *is defined as follows. Given a countable set* Var *of propositional variables and the connectives in* $\{\perp, \wedge_*, \to_*\}$*, the set* Fm *of formulas is defined as*

$$\mathrm{Fm}(L_*) \quad = \quad \mathrm{Var} \cup \{\perp\} \mid \varphi \wedge_* \psi \mid \varphi \to_* \psi$$

where φ *and* ψ *are formulas.*

We will consider the binary connective $\equiv_*$ *as a connective defined by:*

$$\varphi \equiv_* \psi \quad := \quad (\varphi \to_* \psi) \wedge_* (\psi \to_* \varphi)$$

REMARK 9 *Negation is defined within the language as*

$$\neg\varphi = \varphi \to_* \perp$$

A negation $\neg$ *is* involutive *iff* $\neg\neg\varphi \to \varphi$ *is valid. Among the three basic t-norm based logics, the only logic having an involutive negation is Łukasiewicz logic.*

Under a syntactical point of view, there is a set of axioms[4] for each of the basic continuous t-norms. A t-norm based propositional fuzzy logic is defined as usual:

[4] For reasons of space, we do not report here the set of axioms proper of each t-norm based propositional fuzzy logic. The axiomatization of the logics based on the three basic continuous t-norms can be found in [7].

DEFINITION 10 *Let* $* \in \{\text{Ł}, \text{G}, \Pi\}$, *a* $*$-based propositional fuzzy logic *(denoted by* L_**) is the least set of formulas which includes the axioms of such logic and is closed under Modus Ponens.*

The *evaluation* of a formula is also defined as usual:

DEFINITION 11 *Let* $*$ *be a t-norm with* $* \in \{\text{Ł}, \text{G}, \Pi\}$ *and let* φ, ψ *formulas in* $\mathrm{Fm}(L_*)$. *The* propositional evaluation *is a mapping* $e\colon \mathrm{Var} \to [0,1]$ *defined inductively as follows:*

- $e(\bot) = 0$,
- $e(\top) = 1$,
- $e(\varphi \wedge_* \psi) = e(\varphi) * e(\psi)$,
- $e(\varphi \to_* \psi) = e(\varphi) \Rightarrow_* e(\psi)$.

REMARK 12 *According to the definition of negation (Remark 9) the semantics of negation in each basic t-norm based logic is computed as follows:*

For Łukasiewicz logic,[5] *we have* $e(\neg_{\text{Ł}}\varphi) = 1 - e(\varphi)$

For Gödel and Product logics, we have $e(\neg_{\text{G}}\varphi) = e(\neg_{\Pi}\varphi) = \begin{cases} 1 & \text{if } e(\varphi) = 0, \\ 0 & \text{otherwise} \end{cases}$

In [7], completeness results are provided for each of these basic propositional calculus w.r.t. the corresponding semantics.

3 Vague counterfactuals

By a *vague counterfactual* we understand a counterfactual involving vague sentences, i.e. sentences that are not merely true or false, but are evaluated in $[0,1]$. Actually, the fact that the involved sentences are evaluated in $[0,1]$ implies that the counterfactual, as a formula, is evaluated in $[0,1]$ as well.

The most widely accepted definition of a (classical) counterfactual is that of a conditional with an actually false antecedent. While in the classical framework there is no difference between *false* and *non-true* antecedents, in a many-valued framework non-true is not necessarily false but it can take another value < 1. For example, by sentence: "If *a* were tall, *a* would reach the roof", we mean that, within the actual world, individual *a* is not tall in degree 1. Though this does not mean that *a* is tall in degree 0, because we can think that *a* is not tall without thinking that *a* is short. In this case, we have different choices (now excluding the trivial case where the antecedent nowhere holds):

- the simplest case is to require that the antecedent takes value one (what we call 1-semantics, see Section 4 below),
- more generally, we may consider the value of the antecedent to be at least r, for some r higher than the actual value of the antecedent (called $(\geq r)$-semantics, see Section 5).

[5] Note that this negation turns out to be involutive, because $e(\neg\neg\varphi) = 1 - (1 - e(\varphi)) = e(\varphi)$.

A different approach is needed for sentences like: "If Prague was nearer, I could have arrived by feet". Now the antecedent must indeed be false (i.e. its evaluation, within the actual world, will be 0), since Prague cannot (in the actual world) be nearer than it is. In this case, we can look at possible worlds where the antecedent has an arbitrary value > 0. In this case we can

- look at possible worlds where the antecedent takes a value higher than in the actual world (see Section 6 below).

Clearly the worlds where it assumes the maximum value will be a subset of this set of worlds.

Before going on with a formalization of such truth conditions, it must be pointed out that such a semantics do not perfectly fit in with the framework of each basic t-norm based connective. Indeed we will devote more efforts to explain the case of the Łukasiewicz t-norm, because negation, in the other basic t-norms, presents some shortcomings w.r.t. interdefinability for counterfactual connectives and negated antecedents. We extend the propositional language Fm of a given t-norm based logic L_* with symbols for counterfactual connectives $\Box\!\!\rightarrow_*$, $\Diamond\!\!\rightarrow_*$.

DEFINITION 13 *Let $*$ be a basic t-norm. We define the language for vague counterfactuals $L_{\Box\!\rightarrow_*}$ as the least set of sentences containing* $\mathrm{Fm}(L_*)$ *and all the expressions of the form $\varphi \Box\!\!\rightarrow_* \psi$ and $\varphi \Diamond\!\!\rightarrow_* \psi$, where $\varphi, \psi \in \mathrm{Fm}(L_{\Box\!\rightarrow_*})$:*

$$\mathrm{Fm}(L_{\Box\!\rightarrow_*}) \quad = \quad \mathrm{Fm}(L_*) \mid \varphi \Box\!\!\rightarrow_* \psi \mid \varphi \Diamond\!\!\rightarrow_* \psi$$

We will consider vague counterfactuals as counterfactuals involving vague sentences in the language of some t-norm based logic L_*. So we proceed to define a semantics where the truth of the counterfactual $\varphi \Box\!\!\rightarrow_* \psi$ depends on the truth of the non-modal implication $\varphi \rightarrow_* \psi$ in the appropriate possible worlds, where now the implication admits values between 0 and 1.

We present first a direct generalization of Lewis' proposal to the fuzzy case, where we only consider worlds making the antecedent 1-true. We prove that the classical semantics is a particular case of the many-valued one. The scope of this semantics, though, is restrictive with respect to the use of vague counterfactuals in the natural language. The reason is that it ignores the case where the truth-value of the antecedent could be greater than in the actual world, but not being 1-true. For all these semantics, interdefinability results for *would* and *might* counterfactuals are proved for the case of Łukasiewicz logic.

4 1-semantics for vague counterfactuals

In the classical case considered by Lewis, any formula can only be either true or false. This is true as well for two-valued counterfactuals $\varphi \Box\!\!\rightarrow \psi$ in a given world w. Moreover, by definition of counterfactual, the antecedent φ must be false in the actual world. According to the theory of counterfactuals, we must then evaluate ψ in every closest world where the value of φ is different than its value in the actual world. In the classical case, such a difference necessarily yields the value *true*. Thus, in the classical case, the maximum change in the value of φ is already attained in the closest sphere where this

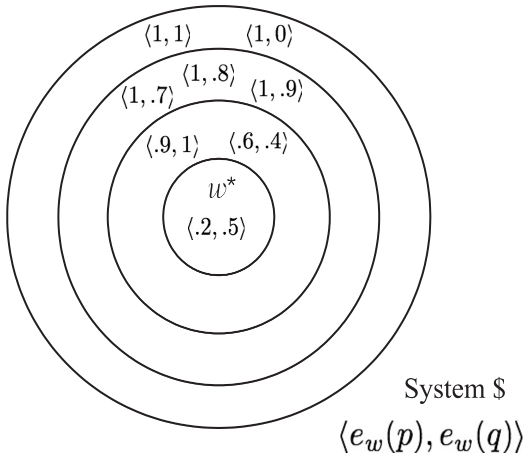

Figure 2. System of spheres from Example 16.

value changes. Outer spheres may contain worlds where many other things change, but the value of φ will be either the actual value (i.e. false) or will remain the same than in the relevant worlds on the closest sphere.

In contrast, as a system of spheres we now consider the set $\mathsf{Mod}(L_{\Box\!\!\rightarrow_*})$ of valuations[6] $e_w\colon \mathrm{Var} \to [0,1]$. As we said, the system of spheres is centered on the actual world $w^\star$ i.e. $\{e_{w^\star}\} \subseteq S$, for any $S \in \$$. So, in outer spheres, the antecedent of a counterfactual may have a value different both than that in the actual world and that in the closest sphere where its value changes.

The semantics for non-counterfactual connectives is as usual. Before we define the semantics for counterfactual connectives we recall a well-known result relating infima and suprema (inf and sup) when the internal negation is involutive (that is, in Łukasiewicz logic). This auxiliary lemma is extensively used later on for several interdefinability results between *would* and *might* counterfactuals.

LEMMA 14 *Let $X \subseteq [0,1]$ be a set of values. Then*

$$1-\inf(X) = \sup(1-X) \quad \textit{and} \quad 1-\sup(X) = \inf(1-X)$$

where $1-X = \{1-x \mid x \in X\}$.

We define first a simple generalization of the semantics for the classical case, now defining φ-worlds as worlds w where φ is 1-true: $e_w(\varphi) = 1$. The other semantics will consider instead φ-worlds to be defined by some weaker condition, e.g. $e_w(\varphi) \geq r$ (for some chosen value $r \in [0,1]$).

[6] We will use indistinctively the possible-world notation w and the algebraic one e_w.

DEFINITION 15 *Let $*$ be a basic t-norm and φ a formula. Let*

$$\mathbb{K}^1_\varphi = \{w' \in W \mid e_{w'}(\varphi) = 1 \text{ and } w' \in \bigcap\{S \in \$ \mid \exists w'' \in S \text{ such that } e_{w''}(\varphi) = 1\}\}$$

Then the 1-semantics of would and might counterfactuals $\varphi \Box\!\!\rightarrow_* \psi$ and $\varphi \Diamond\!\!\rightarrow_* \psi$ is defined by:

$$e^1_w(\varphi \Box\!\!\rightarrow_* \psi) = \begin{cases} 1, & \text{if } \{w' \mid e_{w'}(\varphi) = 1\} = \emptyset \\ \inf\{e_{w'}(\varphi \rightarrow_* \psi) \mid w' \in \mathbb{K}^1_\varphi\} & \text{otherwise} \end{cases}$$

$$e^1_w(\varphi \Diamond\!\!\rightarrow_* \psi) = \begin{cases} 0, & \text{if } \{w' \mid e_{w'}(\varphi) = 1\} = \emptyset \\ \sup\{e_{w'}(\varphi \wedge_* \psi) \mid w' \in \mathbb{K}^1_\varphi\} & \text{otherwise} \end{cases}$$

We assume, as a convention that, whenever $\{w' \in W \mid e_{w'}(\varphi) = 1\} \neq \emptyset$ and $\bigcap\{S \in \$ \mid \exists w'' \in S$ such that $e_{w''}(\varphi) = 1\} = \emptyset$, the value of the counterfactual is undefined.

In cases where there is no φ-world, we take the counterfactual to be *vacuously true*. If we assume Lewis' Limit Assumption, the value of any counterfactual is always defined. Recall that we consider only systems of spheres that satisfy the Limit Assumption.

EXAMPLE 16 *Let $*$ be any basic t-norm. Let $p =$* I am rich *and $q =$* I am happy. *Consider the system of spheres \$ from Figure 2. We identify each world w with a pair of truth-values $\langle e_w(p), e_w(q)\rangle$. For example, the actual world $w^\star$ is such that I am rather not rich but moderately happy. The 1-semantics for the would counterfactual $p \Box\!\!\rightarrow_* q$, i.e.* If I were rich, I would be happy*, considers worlds in the third sphere $S = \{\langle 1, 0.7\rangle, \langle 1, 0.8\rangle, \langle 1, 0.9\rangle\}$, giving*

$$e^1_{w^\star}(p \Box\!\!\rightarrow_* q) = \inf\{1 \Rightarrow_* 0.7,\ 1 \Rightarrow_* 0.8,\ 1 \Rightarrow_* 0.9\} = \inf\{0.7, 0.8, 0.9\} = 0.7,$$

so that this counterfactual is quite true. The same sphere S is where the truth-value for the might counterfactual $p \Diamond\!\!\rightarrow_* q$ is computed, now giving

$$e^1_{w^\star}(p \Diamond\!\!\rightarrow_* q) = \sup\{1 \wedge_* 0.7,\ 1 \wedge_* 0.8,\ 1 \wedge_* 0.9\} = \sup\{0.7, 0.8, 0.9\} = 0.9,$$

so that If I were rich, I might be happy *is highly true.*

Restricting the set of truth-values to $\{0,1\}$ makes our 1-semantics from Definition 15 and Lewis' semantics from Lemma 5 equivalent. In other words, Lewis' semantics for counterfactuals is a particular case of the preceding semantics under the condition that the evaluations are restricted to $\{0,1\}$ as in classical logic.

PROPOSITION 17 *Let $*$ be a basic t-norm. Let $e^c_w(\cdot)$ denote Lewis' semantics and W a set of classical possible worlds (i.e. $W \subseteq \{e^c_w \colon \mathrm{Var} \rightarrow \{0,1\}\}$). For any classical system of spheres \$ and any world $w \in W$, the 1-semantics $e^1_w(\cdot)$ definition gives:*

$$e^c_w(\varphi \Box\!\!\rightarrow \psi) = e^1_w(\varphi \Box\!\!\rightarrow_* \psi)$$

$$e^c_w(\varphi \Diamond\!\!\rightarrow \psi) = e^1_w(\varphi \Diamond\!\!\rightarrow_* \psi)$$

Proof ($\Box\!\!\rightarrow$) We have $e^c_w(\varphi \Box\!\!\rightarrow \psi) = 1$ iff either $[\varphi] = \emptyset$ or or $w' \models \psi$, for every $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$. This is equivalent to the fact that either $\mathbb{K}^1_\varphi = \emptyset$ or the infimum of $e^1_{w'}(\varphi \rightarrow_* \psi)$, for each such world w', is 1. But this is just $e^1_w(\varphi \Box\!\!\rightarrow_* \psi) = 1$.

($\Diamond\!\!\rightarrow$) We have $e^c_w(\varphi \Diamond\!\!\rightarrow \psi) = 1$ iff both $[\varphi] \neq \emptyset$ and $w' \models \psi$, for some world $w' \in [\varphi] \cap \bigcap\{S \in \$^w \mid S \cap [\varphi] \neq \emptyset\}$. This is equivalent to the fact that both $\mathbb{K}^1_\varphi \neq \emptyset$ and the supremum of the values of $\varphi \wedge_* \psi$ is 1 in such a sphere, which is the definition of 1-semantics for might counterfactuals; hence, we have that $e^c_w(\varphi \Diamond\!\!\rightarrow \psi) = 1$ iff $e^1_w(\varphi \Diamond\!\!\rightarrow_* \psi) = 1$. □

In the particular case of Łukasiewicz, we also have that classical interdefinability of *would* and *might* counterfactuals is preserved.

PROPOSITION 18 *Let W be a set of possible worlds, $w \in W$ and let $\neg_{Ł}$ be Łukasiewicz negation, then:*

$$\begin{aligned} e^1_w(\varphi \Diamond\!\!\rightarrow_{Ł} \psi) &= e^1_w(\neg_{Ł}(\varphi \Box\!\!\rightarrow_{Ł} \neg_{Ł}\psi)) \\ e^1_w(\varphi \Box\!\!\rightarrow_{Ł} \psi) &= e^1_w(\neg_{Ł}(\varphi \Diamond\!\!\rightarrow_{Ł} \neg_{Ł}\psi)) \end{aligned}$$

Proof Let \$ and $w \in W$ be given, then $e^1_w(\neg_{Ł}(\varphi \Box\!\!\rightarrow_{Ł} \neg_{Ł}\psi)) =$

$=$	$1 - e^1_w(\varphi \Box\!\!\rightarrow_{Ł} \neg_{Ł}\psi) =$	Remark 12
$=$	$1 - \inf\{e^1_{w'}(\varphi \rightarrow_{Ł} \neg_{Ł}\psi) \mid w' \in \mathbb{K}^1_\varphi\} =$	Definition 15
$=$	$1 - \inf\{e^1_{w'}(\varphi \rightarrow_{Ł} (\psi \rightarrow_{Ł} \bot)) \mid w' \in \mathbb{K}^1_\varphi\} =$	Remark 9
$=$	$1 - \inf\{e^1_{w'}((\varphi \wedge_{Ł} \psi) \rightarrow_{Ł} \bot) \mid w' \in \mathbb{K}^1_\varphi\} =$	Definition 7
$=$	$1 - \inf\{e^1_{w'}(\neg_{Ł}(\varphi \wedge_{Ł} \psi)) \mid w' \in \mathbb{K}^1_\varphi\} =$	Remark 9
$=$	$1 - \inf\{1 - e^1_{w'}(\varphi \wedge_{Ł} \psi) \mid w' \in \mathbb{K}^1_\varphi\} =$	Remark 12
$=$	$1 - (1 - \sup\{e^1_{w'}(\varphi \wedge_{Ł} \psi) \mid w' \in \mathbb{K}^1_\varphi\}) =$	Lemma 14
$=$	$\sup\{e^1_{w'}(\varphi \wedge_{Ł} \psi) \mid w' \in \mathbb{K}^1_\varphi\} =$	Double neg. law
$=$	$e^1_w(\varphi \Diamond\!\!\rightarrow_{Ł} \psi)$	Definition 15

Hence, by Definition 15, we have exactly the definition of $e^1_w(\varphi \Diamond\!\!\rightarrow_{Ł} \psi)$:

$$e^1_w(\neg_{Ł}(\varphi \Box\!\!\rightarrow_{Ł} \neg_{Ł}\psi)) = \begin{cases} \neg_{Ł} 1 = 0, & \text{if } \{w' \mid e_{w'}(\varphi) = 1\} = \emptyset \\ \sup\{e_{w'}(\varphi \wedge_{Ł} \psi) \mid w' \in \mathbb{K}^1_\varphi\}, & \text{otherwise} \end{cases}$$

For the second equivalence it is enough to apply twice the Double Negation law to obtain: $e^1_w(\varphi \Box\!\!\rightarrow_{Ł} \psi) = e^1_w(\neg_{Ł}\neg_{Ł}(\varphi \Box\!\!\rightarrow_{Ł} \neg_{Ł}\neg_{Ł}\psi))$, and then by the former equivalence we have $e^1_w(\neg_{Ł}\neg_{Ł}(\varphi \Box\!\!\rightarrow_{Ł} \neg_{Ł}\neg_{Ł}\psi)) = e^1_w(\neg_{Ł}(\varphi \Diamond\!\!\rightarrow_{Ł} \neg_{Ł}\psi))$. □

5 *r*-semantics for vague counterfactuals.

Having, in a many-valued framework, truth-values other than 0 and 1, permits us to look at worlds satisfying the antecedent φ of a counterfactual in a fixed degree r, with r lying between the actual truth-value of φ and 1: $e_{w^\star}(\varphi) \leq r \leq 1$. (See Figure 3.)

Figure 3. Possible truth-values for antecedent φ in r-semantics.

We define next the corresponding semantics, called r-semantics, in terms of the set of worlds w with $e_w(\varphi) \geq r$.

DEFINITION 19 *Let $*$ be a basic t-norm and φ a formula. For a given $r > 0$, let*

$$\mathbb{K}^r_\varphi = \{w' \in W \mid e_{w'}(\varphi) \geq r \text{ and } w' \in \bigcap\{S \in \$ \mid \exists w'' \in S \text{ such that } e_{w''}(\varphi) \geq r\}\}$$

Then we define the r-semantics of $\Box\!\!\rightarrow$ *and* $\Diamond\!\!\rightarrow$ *as follows:*

$$e^r_w(\varphi \Box\!\!\rightarrow_* \psi) = \begin{cases} 1, & \text{if } \{w' \mid e_{w'}(\varphi) \geq r\} = \emptyset \\ \inf\{e_{w'}(\varphi \rightarrow_* \psi) \mid w' \in \mathbb{K}^r_\varphi\}, & \text{otherwise} \end{cases}$$

$$e^r_w(\varphi \Diamond\!\!\rightarrow_* \psi) = \begin{cases} 0, & \text{if } \{w' \mid e_{w'}(\varphi) \geq r\} = \emptyset \\ \sup\{e_{w'}(\varphi \wedge_* \psi) \mid w' \in \mathbb{K}^r_\varphi\}, & \text{otherwise} \end{cases}$$

We assume, as a convention that, whenever $\{w' \in W \mid e_{w'}(\varphi) \geq r\} \neq \emptyset$ and $\bigcap\{S \in \$ \mid \exists w'' \in S(e_{w''}$ such that $\varphi \geq r)\} = \emptyset$, the value of the counterfactual is undefined.

The Limit Assumption on systems of spheres excludes counterintuitive examples, and makes counterfactuals' truth-values always to be defined. The first result for this semantics is that for the particular case $r = 1$, it collapses to the previous 1-semantics.

PROPOSITION 20 *Let $*$ be a basic t-norm. Setting $r = 1$ for a given system of spheres \$ and world w, r-semantics collapses to the 1-semantics.*

$$\begin{aligned} e^1_w(\varphi \Box\!\!\rightarrow_* \psi) &= e^{r=1}_w(\varphi \Box\!\!\rightarrow_* \psi) \\ e^1_w(\varphi \Diamond\!\!\rightarrow_* \psi) &= e^{r=1}_w(\varphi \Diamond\!\!\rightarrow_* \psi) \end{aligned}$$

Proof ($\Box\!\!\rightarrow_*$) Obvious, since the set of worlds in Definition 15 is identical to $\mathbb{K}^1_\varphi$ (because $e_w(\cdot) \geq 1 \iff e_w(\cdot) = 1$). □

Moreover, in the two-valued case, Proposition 17 above makes Lewis' semantics (Definition 3) a particular case of the r-semantics for $r = 1$.

COROLLARY 21 *Let $*$ be a basic t-norm. Let W a set of classical possible worlds (i.e. $W \subseteq \{e^c_w \colon \mathrm{Var} \rightarrow \{0,1\}\}$). For any classical system of spheres \$ and any world $w \in W$, the r-semantics $e^r_w(\cdot)$ definition gives:*

$$e^c_w(\varphi \Box\!\!\rightarrow \psi) = e^r_w(\varphi \Box\!\!\rightarrow_* \psi) \text{ and } e^c_w(\varphi \Diamond\!\!\rightarrow \psi) = e^r_w(\varphi \Diamond\!\!\rightarrow_* \psi)$$

Proof If the evaluation is in $\{0,1\}$, then the r-semantics reduces to $r = 1$ case. We apply Propositions 20 and 17. □

Finally we prove that, in the r-semantics, $\Box\!\!\rightarrow_*$ and $\Diamond\!\!\rightarrow_*$ can also be defined from each other for the Łukasiewicz case: $* = $ Ł.

PROPOSITION 22 *Let W be a set of possible worlds, $w \in W$ and let $\neg_Ł$ be Łukasiewicz negation, then:*

$$\begin{aligned} e^r_w(\varphi \Diamond\!\!\rightarrow_Ł \psi) &= e^r_w(\neg_Ł(\varphi \Box\!\!\rightarrow_Ł \neg_Ł \psi)) \\ e^r_w(\varphi \Box\!\!\rightarrow_Ł \psi) &= e^r_w(\neg_Ł(\varphi \Diamond\!\!\rightarrow_Ł \neg_Ł \psi)) \end{aligned}$$

Proof We have the following $e^r_w(\neg_Ł(\varphi \,\Box\!\!\rightarrow_Ł \neg_Ł \psi)) =$

$$
\begin{array}{lll}
= & 1 - e^r_w(\varphi \,\Box\!\!\rightarrow_Ł \neg_Ł \psi) = & \text{Remark 12} \\
= & 1 - \inf\{e_{w'}(\varphi \rightarrow_Ł \neg_Ł \psi) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Definition 19} \\
= & \sup\{1 - e_{w'}(\varphi \rightarrow_Ł \neg_Ł \psi) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Lemma 14} \\
= & \sup\{1 - e_{w'}(\varphi \rightarrow_Ł (\psi \rightarrow_Ł \bot)) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Remark 9} \\
= & \sup\{1 - e_{w'}((\varphi \wedge_Ł \psi) \rightarrow_Ł \bot) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Definition 7} \\
= & \sup\{1 - e_{w'}(\neg_Ł(\varphi \wedge_Ł \psi)) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Remark 9} \\
= & \sup\{e_{w'}(\neg_Ł\neg_Ł(\varphi \wedge_Ł \psi)) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Remark 12} \\
= & \sup\{e_{w'}(\varphi \wedge_Ł \psi) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Double Neg. Law} \\
= & e^r_w(\varphi \,\Diamond\!\!\rightarrow_Ł \psi) & \text{Def. 19}
\end{array}
$$

Hence, by Definition 19, we have exactly the definition of $e^r_w(\varphi \,\Diamond\!\!\rightarrow_Ł \psi)$:

$$
e^r_w(\neg_Ł(\varphi \,\Box\!\!\rightarrow_Ł \neg_Ł \psi)) \;=\; \begin{cases} \neg_Ł 1 = 0, & \text{if } \{w' \mid e_{w'}(\varphi) \geq r\} = \emptyset \\ \sup\{e_{w'}(\varphi \wedge_Ł \psi) \mid w' \in \mathbb{K}^r_\varphi\}, & \text{otherwise} \end{cases}
$$

The second equivalence is proven applying to the first the double negation law. □

EXAMPLE 23 *Recall the system* \$ *from Figure 2. As before, let* $p =$ I am rich*,* $q =$ I am happy *but now we consider the Łukasiewicz t-norm* $*_Ł$ *only. The r-semantics with* $r \geq 0.5$ *is as follows. For the would counterfactual* $p \,\Box\!\!\rightarrow_Ł q$*,* If I were rich, I would be happy*, we look at worlds in the second sphere* $S = \{\langle 0.9, 1\rangle, \langle 0.6, 0.4\rangle\}$*, giving*

$$e^{0.5}_{w^\star}(p \,\Box\!\!\rightarrow_Ł q) = \inf\{0.9 \Rightarrow_Ł 1,\, 0.6 \Rightarrow_Ł 0.4\} = \inf\{1,\, 0.8\} = 0.8$$

so that this counterfactual is very true. The same sphere S is selected for evaluating the might counterfactual $p \,\Diamond\!\!\rightarrow_Ł q$*, now giving*

$$e^{0.5}_{w^\star}(p \,\Diamond\!\!\rightarrow_Ł q) = \sup\{0.9 \wedge_Ł 1,\, 0.6 \wedge_Ł 0.4\} = \sup\{0.9,\, 0\} = 0.9$$

so that If I were rich, I might be happy *is highly true.*

6 More-than-actual semantics

In the previous sections, we presented different semantics for counterfactuals $\varphi \,\Box\!\!\rightarrow_* \psi$ where relevant worlds assigned φ a fixed value higher than that of the actual world $w^\star$. A weaker, yet perhaps more intuitive, semantics for counterfactuals will consider worlds assigning φ a value that is *minimally* higher than that of $w^\star$. We will call this semantics *more-than-actual* semantics. See Figure 4 for the set of possible truth values for antecedent φ in relevant worlds.

Figure 4. Possible truth-values for antecedent φ in *more-than-actual* semantics.

There is a noticeable difference between the finite-valued case and the infinite-valued one. In the finitely-valued case, say $\{0 = r_0, r_1, \ldots, r_{k-1}, r_k = 1\}$ with $r_m < r_{m+1}$, we could simply set the lower bound r for the value of the antecedent to be *minimally* higher than the actual value: so, if $e_{w^\star}(\varphi) = r_m$, then the designed value to be considered should be r_{m+1}. Each such case is a particular instance of the r-semantics considered above. In contrast, if the set of truth-values is dense, the notion of a next-degree does not make sense; so we cannot apply a r-semantics, because there is not any minimal value higher than r and, by definition of counterfactuals, we cannot set $r = e_{w^\star}(\varphi)$ either.

In order to capture this notion of more-than-actual semantics, we propose a definition based both on the r-semantics and the notion of limit lim. As usual, $\lim_{x\to c^+} f(x)$ denotes the limit of $f(x)$ as x approaches to c from the right, i.e. as x decreases its value, while remaining strictly greater than c.

DEFINITION 24 *Let $*$ be a basic t-norm, φ, ψ formulas and r ranges on all values $> e_w(\varphi)$. The* more than actual *semantics $e^\star(\cdot)$ for would and might counterfactuals is:*

$$\begin{aligned} e^\star_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi) &= \lim_{r\to e_w(\varphi)^+} e^r_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi) \\ e^\star_w(\varphi \mathrel{\Diamond\!\!\rightarrow}_* \psi) &= \lim_{r\to e_w(\varphi)^+} e^r_w(\varphi \mathrel{\Diamond\!\!\rightarrow}_* \psi) \end{aligned}$$

We have that $e^\star_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi)$ is undefined *iff for all $r > e_w(\varphi)$, $e^r_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi)$ is undefined.*

It is obvious that the finite-valued case is a particular instance of this definition:

$$e^\star_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi) \quad = \quad e^{r_{m+1}}_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi), \qquad \text{where } r_m = e_w(\varphi)$$

It is interesting to observe the behavior of function $f(r) := e^r_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi)$ in comparison to the actual value of the (non-modal) residuated implication $e_{w^\star}(\varphi \rightarrow_* \psi)$. We may have examples where the limit of $f(r)$ (for counterfactuals) coincides with that of the implication, and examples where these two values do not coincide:

EXAMPLE 25 *Consider the comparative counterfactual* If I were richer, I would be happy. *We assume a continuous system of spheres, with sphere S_r containing the most similar worlds where* I am rich (= φ) in degree *r, for any $r > e_{w^\star}(\varphi)$. That is, worlds in S_r are minimal w.r.t. changes in propositions other than φ. The distance of worlds in sphere S_r is $r - e_{w^\star}(\varphi)$. In the case that a small increase (say less than 1 cent) did not make a difference, the previous semantics gives a function $f(r)$ whose limit coincides with the actual value of the residuated implication $\varphi \rightarrow_* \psi$. In this case, $e^\star_{w^\star}(p \mathrel{\Box\!\!\rightarrow}_* q)$ $= e_{w^\star}(p \rightarrow_* q)$. See Figure 5 (left). In contrast, consider the following situation: you have* \$*10, and you want to buy a music CD, which is sold for* \$*10 plus some (obligatory but arbitrary) tip. Buying this CD would make you happier. The limit of function $f(r)$ does not coincide now with the implication $e_{w^\star}(\varphi \rightarrow_* \psi)$. See Figure 5 (right).*

As before, we can prove for the Łukasiewicz t-norm that $\mathrel{\Box\!\!\rightarrow}_Ł$ and $\mathrel{\Diamond\!\!\rightarrow}_Ł$ are interdefinable according to the *more-than-actual*-semantics.

PROPOSITION 26 *Let W be a set of possible worlds, $w \in W$. Then:*

$$\begin{aligned} e^\star_w(\varphi \mathrel{\Diamond\!\!\rightarrow}_Ł \psi) &= e^\star_w(\neg_Ł(\varphi \mathrel{\Box\!\!\rightarrow}_Ł \neg_Ł \psi)) \\ e^\star_w(\varphi \mathrel{\Box\!\!\rightarrow}_Ł \psi) &= e^\star_w(\neg_Ł(\varphi \mathrel{\Diamond\!\!\rightarrow}_Ł \neg_Ł \psi)) \end{aligned}$$

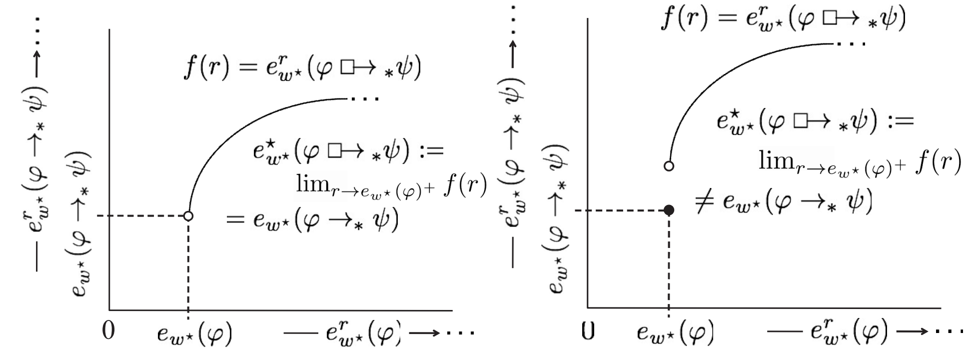

Figure 5. The limit of function $f(r)$ and the actual truth-value of $\varphi \to_* \psi$ coincide (left) and do not coincide (right). In both graphics, w_r denotes an arbitrary world in sphere S_r.

Proof We show the former claim $e^\star_w(\varphi \,\Diamond\!\!\rightarrow_{\text{Ł}}\, \psi) = e^\star_w(\neg_{\text{Ł}}(\varphi \,\Box\!\!\rightarrow_{\text{Ł}}\, \neg_{\text{Ł}}\psi))$, since the proof of the other is analogous. We have $e^\star_w(\neg_{\text{Ł}}(\varphi \,\Box\!\!\rightarrow_{\text{Ł}}\, \psi)) =$

$$
\begin{array}{lll}
= & \lim_{r \to e_w(\varphi)^+} e^r_w(\neg_{\text{Ł}}(\varphi \,\Box\!\!\rightarrow_{\text{Ł}}\, \neg_{\text{Ł}}\psi)) = & \text{Definition 24} \\
= & \lim_{r \to e_w(\varphi)^+} e^r_w(\varphi \,\Diamond\!\!\rightarrow_{\text{Ł}}\, \neg_{\text{Ł}}\psi) = & \text{Proposition 22} \\
= & e^\star_w(\varphi \,\Diamond\!\!\rightarrow_{\text{Ł}}\, \psi) & \text{Definition 24} \quad \square
\end{array}
$$

We give some examples of (non-)tautologies for the semantics presented above:

EXAMPLE 27 *For any of the above semantics, and any basic t-norm $*$, counterfactuals of the following form are tautologies (in systems \$ that are universal, centered and satisfying the Limit Assumption):*

$$\varphi \,\Box\!\!\rightarrow_*\, (\psi \to_* \varphi), \quad (\varphi \wedge_* \psi) \,\Box\!\!\rightarrow_*\, \varphi$$

In contrast, these expressions are not tautologies for all φ, ψ:

$$\varphi \,\Box\!\!\rightarrow_*\, (\psi \,\Box\!\!\rightarrow_*\, \varphi), \quad \varphi \to_* (\psi \,\Box\!\!\rightarrow_*\, \varphi)$$

Proof: About the first tautology: this follows from the fact that for any t-norm, any fuzzy model e_w satisfies: $e_w(\varphi \to_* (\psi \to_* \varphi)) = 1$. Hence, whatever be $r \in [0,1]$, and $w \in \mathbb{K}^r$, the r-semantics assigns this counterfactual the value 1. Finally, for the more-than-actual semantics, say $> r = e_{w^\star}(\varphi)$, we just have the limit of r-semantics is again 1, for any $r' > r$.

The reasoning about the second tautology in similar, using that $(\varphi \wedge_* \psi) \to_* \varphi$ is a tautology in any model e_w, for an arbitrary t-norm.

For the first non-tautology we show it for Łukasiewicz t-norm and r-semantics with $r = 0.7$. Consider a set of possible worlds W and $w^\star, w', w'' \in W$ such that $e_{w^\star}(\varphi) = 0.3$, $e_{w'}(\varphi) = 0.7$, $e_{w''}(\varphi) = 0.2$, and $e_{w''}(\psi) = 0.7$. Let it be a similarity relation that induces a system of spheres $\$^{w^\star}$ in which $e_{w^\star}(\varphi \,\Box\!\!\rightarrow_*\, (\psi \,\Box\!\!\rightarrow_*\, \varphi)) = \inf\{e_w(\varphi \to_* (\psi \,\Box\!\!\rightarrow_*\, \varphi)) \mid w \in \mathbb{K}^{0.7}\} = e_{w'}(\varphi \to_* (\psi \,\Box\!\!\rightarrow_*\, \varphi))$ and a system of spheres $\$^{w'}$ in which $e_{w'}(\psi \,\Box\!\!\rightarrow_*\, \varphi)$

$= \inf\{e_w(\psi \rightarrow_* \varphi)) \mid w \in \mathbb{K}^{0.7}\} = e_{w''}(\psi \rightarrow_* \varphi) = 0.5$. So, $e_{w^\star}(\varphi \,\Box\!\!\rightarrow_* (\psi \,\Box\!\!\rightarrow_* \varphi)) = \inf\{e_w(\varphi \rightarrow_* (\psi \,\Box\!\!\rightarrow_* \varphi)) \mid w \in \mathbb{K}^{0.7}\} = e_{w'}(\varphi \rightarrow_* (\psi \,\Box\!\!\rightarrow_* \varphi)) = 0.7 \Rightarrow_* 0.5 = 0.8 < 1$.

For the second non-tautology we show it for Łukasiewicz t-norm and r-semantics with $r = 0.7$. Consider a set of possible worlds W and $w^\star, w' \in W$ such that $e_{w^\star}(\varphi) = 0.6$, $e_{w'}(\psi) = 0.8$ and $e_{w'}(\varphi) = 0.1$. Let it be a similarity relation that induces a system of spheres $\$^{w^\star}$ in which $e_{w^\star}(\psi \,\Box\!\!\rightarrow_* \varphi) = \inf\{e_w(\psi \rightarrow_* \varphi) \mid w \in \mathbb{K}^{0.7}\} = e_{w'}(\psi \rightarrow_* \varphi) = 0.8 \Rightarrow_* 0.1 = 0.3$. So, $e_{w^\star}(\varphi \rightarrow_* (\psi \,\Box\!\!\rightarrow_* \varphi)) = 0.6 \Rightarrow_* e_{w^\star}(\psi \,\Box\!\!\rightarrow_* \varphi) = 0.6 \Rightarrow_* 0.3 = 0.7 < 1$.

7 Expanding the language

It is usual in the literature on fuzzy propositional logic, to study expansions of the language by some set of truth-constants or Baaz' operator Δ. In this section we study the expressivity of the language of counterfactuals when expanded in some of these ways. Informally, expanding the language by truth-constants permits to talk about truth-values within formulas, thus making possible to express that some formula is true *at least* (*at most*, *exactly*) in a certain degree. On the other hand, the Delta operator Δ allows us to distinguish whether a formula is 1-true or not.

7.1 Expanding the language with Δ

The Δ operator was introduced by Baaz in [1]. The semantics for this operator is:

$$e(\Delta\varphi) = \begin{cases} 1 & \text{if } e(\varphi) = 1 \\ 0 & \text{otherwise} \end{cases}$$

For instance, for $e(\varphi) = 0.9$, we have $e(\Delta\varphi) = 0$, while if $e(\varphi) = 1$, then $e(\Delta\varphi) = 1$.

Thus, this operator requires that the formula under its scope is absolutely true. In our context, this operator has the following consequences:

- If Δ is applied to the antecedent of a counterfactual $\varphi \,\Box\!\!\rightarrow_* \psi$, i.e. for $\Delta\varphi \,\Box\!\!\rightarrow_* \psi$, then the r-semantics collapses to the 1-semantics (the latter in a language without Δ).
- If, in addition, Δ is also applied to the consequent, i.e. $\Delta\varphi \,\Box\!\!\rightarrow_* \Delta\psi$, then each of our semantics collapse to the classical case of Lewis (applied to a language with Δ).

These claims are shown by the following propositions:

PROPOSITION 28 *Let $W, w \in W$ and $\$$ be given, and let $*$ be a basic t-norm. For any formulas φ, ψ, and a fixed $r > e_w(\varphi)$,*

$$e^r_w(\Delta\varphi \,\Box\!\!\rightarrow_* \psi) \quad = \quad e^1_w(\varphi \,\Box\!\!\rightarrow_* \psi)$$

Proof We have $e^r_w(\Delta\varphi \,\Box\!\!\rightarrow_* \psi) =$

$$\begin{array}{rlr}
= & \inf\{e_{w'}(\Delta\varphi \rightarrow_* \psi) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{Definition 19} \\
= & \inf\{e_{w'}(\Delta\varphi \rightarrow_* \psi) \mid w' \in \mathbb{K}^1_\varphi\} = & \text{since } e_{w'}(\Delta\varphi) \geq r \Leftrightarrow e_{w'}(\Delta\varphi) = 1 \\
= & \inf\{e_{w'}(\varphi \rightarrow_* \psi) \mid w' \in \mathbb{K}^1_\varphi\} = & \text{since } e_{w'}(\Delta\varphi) = 1 \Leftrightarrow e_{w'}(\varphi) = 1 \\
= & e^1_w(\varphi \,\Box\!\!\rightarrow_* \psi) & \text{Definition 15}
\end{array}$$

□

PROPOSITION 29 *Let $e^c_w(\cdot)$ denote Lewis' semantics and W a set of multi-valued possible worlds. For any system of spheres \$, any world $w \in W$ and any $r > 0$, the r-semantics $e^r_w(\cdot)$ definition gives:*

$$\begin{aligned} e^c_w(\varphi \mathrel{\Box\!\!\rightarrow} \psi) &= e^r_w(\Delta\varphi \mathrel{\Box\!\!\rightarrow}_* \Delta\psi) \\ e^c_w(\varphi \mathrel{\Diamond\!\!\rightarrow} \psi) &= e^r_w(\Delta\varphi \mathrel{\Diamond\!\!\rightarrow}_* \Delta\psi) \end{aligned}$$

Proof We prove the result just for the would counterfactual $\mathrel{\Box\!\!\rightarrow}_*$, since the prove for the other operator is analogous. We have that $e^c_w(\varphi \mathrel{\Box\!\!\rightarrow} \psi) = 1$ iff either $[\Delta\psi] = \emptyset$ or each φ-world w' in the \$-closest sphere S containing a φ-world, is a ψ-world. This is equivalent to the infimum of $e^r_{w'}(\Delta\varphi \rightarrow_* \Delta\psi)$ being 1. But this is just $e^r_{w'}(\Delta\varphi \mathrel{\Box\!\!\rightarrow}_* \Delta\psi) = 1$ for each such world w'. □

Hence, in counterfactuals with Δ in the antecedent and the consequent, there is no need for many-valued truth conditions since these counterfactuals cannot take values besides 0 and 1.

This language expansion permits to define truth-conditions for the following examples: (1) If I were absolutely rich, I would be happy, $\Delta\varphi \mathrel{\Box\!\!\rightarrow}_* \psi$, and (2) If I were absolutely rich, I would be absolutely happy, $\Delta\varphi \mathrel{\Diamond\!\!\rightarrow} \Delta\psi$.

7.2 Expanding the language with truth-constants

Truth-constants were introduced by Pavelka (in [9]; see also Hájek [7]) for the case of Łukasiewicz t-norm (for other t-norms, see [2] and [3]). This formalism consists in introducing a suitable set of truth-constants into the language propositional fuzzy logic. By *suitable* we mean a subset C of the set of truth-values, closed under the operations of the truth-value algebra. Each new constant $\overline{r}$ can only take the truth-value r in any evaluation: $e(\overline{r}) = r$, for any e. For any t-norm $*$, it is always true that $\overline{r} \wedge_* \overline{s} \Leftrightarrow \overline{r * s}$ and $\overline{r} \rightarrow_* \overline{s} \Leftrightarrow \overline{r \Rightarrow_* s}$. As an example, $e(\overline{0.4} \rightarrow_* p) = 1$ iff $e(p) \geq 0.4$. An *evaluated formula*, as defined in [2], is a formula of the form $\overline{r} \rightarrow_* \varphi$, or $\varphi \rightarrow_* \overline{r}$, or $\varphi \equiv_* \overline{r}$, where φ contains no truth-constant but $\perp$ or $\overline{1}$. In the following, we assume that the set of formulas expanded by truth-constants is restricted to evaluated formulas of the form $\overline{r} \rightarrow_* \varphi$.

In our context, this operator has the following consequences:

- If evaluated formulas (of the mentioned kind) occur only in the the antecedent of a counterfactual, i.e. $(\overline{s} \rightarrow_* \varphi) \mathrel{\Box\!\!\rightarrow}_* \psi$, then the 1-semantics captures the r-semantics, for any $r > 0$.
- The r-semantics (without truth-constants) collapses to the classical semantics of Lewis applied to counterfactuals whose antecedent and consequent are evaluated formulas of the mentioned kind.

These claims are proved next:

PROPOSITION 30 *Let W, w and \$ be given, and let $*$ be a basic t-norm. For any formulas φ, ψ, and a fixed $r > e_w(\varphi)$,*

$$e^r_w(\varphi \mathrel{\Box\!\!\rightarrow}_* \psi) = e^1_w((\overline{r} \rightarrow_* \varphi) \mathrel{\Box\!\!\rightarrow}_* \psi)$$

Proof We have $e^1_w((\overline{r} \rightarrow_* \varphi) \mathrel{\Box\!\!\rightarrow_*} \psi) =$

$$\begin{array}{lll} = & \inf\{e_{w'}((\overline{r} \rightarrow_* \varphi) \rightarrow_* \psi) \mid w' \in \mathbb{K}^1_\varphi\} = & \text{Definition 15} \\ = & \inf\{e_{w'}(\varphi \rightarrow_* \psi) \mid w' \in \mathbb{K}^r_\varphi\} = & \text{since } e_{w'}(\overline{r} \rightarrow_* \varphi) = 1 \Leftrightarrow e_{w'}(\varphi) \geq r \\ = & e^r_w(\varphi \mathrel{\Box\!\!\rightarrow_*} \psi) & \text{Definition 19} \quad \Box \end{array}$$

PROPOSITION 31 *Let $e^c_w(\cdot)$ denote Lewis' semantics and W a set of many-valued possible worlds. For any system of spheres \$ and any world $w \in W$,*

$$e^r_w(\varphi \mathrel{\Box\!\!\rightarrow_*} \psi) \geq s \iff e^c_w((\overline{r} \rightarrow_* \varphi) \mathrel{\Box\!\!\rightarrow_*} (\overline{s} \rightarrow_* \psi)) = 1$$

$$e^r_w(\varphi \mathrel{\Diamond\!\!\rightarrow_*} \psi) \geq s \iff e^c_w((\overline{r} \rightarrow_* \varphi) \mathrel{\Diamond\!\!\rightarrow_*} (\overline{s} \rightarrow_* \psi)) = 1$$

Proof We will prove the result just for the would counterfactual $\Box\!\!\rightarrow$, since the proof for the other operator is analogous.

We have that $e^r_w(\varphi \mathrel{\Box\!\!\rightarrow_*} \psi) \geq s$ if and only if each world w' which gives φ a value greater or equal than r in the \$-closest sphere S containing a world w'' which gives φ a value greater or equal than r, gives $\varphi \rightarrow_* \psi$ a value greater or equal than s. So, w' is a world such that (i) gives the formula $\overline{r} \rightarrow_* \varphi$ value 1, (ii) is an element of the \$-closest sphere S containing a world w'' giving the formula $\overline{r} \rightarrow_* \varphi$ value 1, and (iii) gives the formula $\overline{s} \rightarrow_* \psi$ value 1. $\Box$

The next examples illustrate instances of natural language expressions that can be expressed in a language expanded by truth-constants.

EXAMPLE 32 *Assume we interpret truth at degree (at least)* 0.8 *as* very *true. Then the following sentences:*

(1)	*If I were very rich, I would be happy.*	$(\overline{0.8} \rightarrow_* \varphi) \mathrel{\Box\!\!\rightarrow} \psi$
(2)	*If I were very rich I would be quite happy.*	$(\overline{0.8} \rightarrow_* \varphi) \mathrel{\Diamond\!\!\rightarrow} (\overline{0.7} \rightarrow_* \psi)$

8 Conclusions

We have addressed the problem of generalizing Lewis' system of spheres semantics to the many-valued case, making use of t-norm based semantics. This permits to provide truth-conditions for vague counterfactuals in a faithful way. Due to the plurality of truth-values, we do not obtain a unique semantics as a result. We define several semantics, each one meaningful in a different context a vague counterfactuals can be uttered.

For the case of Łukasiewicz t-norm, we prove that would and might counterfactuals are interdefinable just like in the classical case. Then, we also show for which pairs of semantics, one can be reduced to the other (including here semantics defined on expanded languages). See the scheme in the next figure for a summary of this reducibility.

BIBLIOGRAPHY

[1] Mathias Baaz: Infinite-valued Gödel logics with 0-1-projections and relativizations. In Gödel '96 (Brno, 1996), Springer, 1996, pp. 23–33.

[2] Francesc Esteva, Joan Gispert, Lluís Godo and Carles Noguera: Adding truth-constants to logics of continuous t-norms: Axiomatization and completeness results. Fuzzy Sets and Systems, 2007 (158), pp. 597–618.

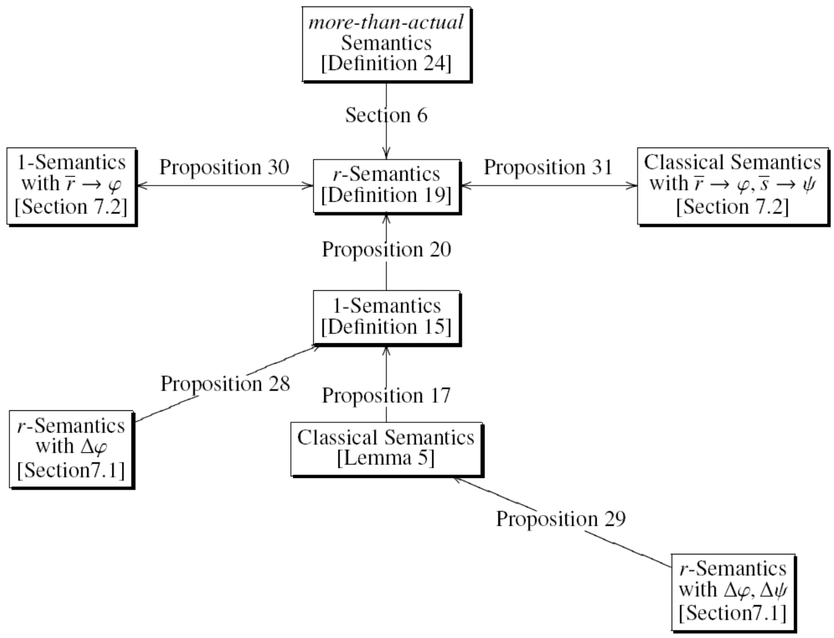

Figure 6. Reducibility relations among semantics for vague counterfactuals.

[3] Francesc Esteva, Lluís Godo and Carles Noguera: Expanding the propositional logic of a t-norm with truth-constants: completeness results for rational semantics. Soft Computing, 2010 (14), pp. 273–284.
[4] Christian G. Fermüller: Theories of Vagueness versus Fuzzy Logic: can logicians learn from philosophers? Neural Network World, 2003 (5), pp. 455–466.
[5] Christian G. Fermüller: Fuzzy logic and vagueness: can philosophers learn from Petr Hájek? In P. Cintula, Z. Haniková and V. Švejdar (eds.) *Witnessed years: Essays in Honour of Petr Hájek*, College Publications, 2009, pp. 373–386.
[6] Lluís Godo and Ricardo O. Rodríguez: Graded Similarity-based semantics for nonmonotonic inferences. Annals of Mathematics and Artificial Intelligence, 2002 (34), pp. 89–105.
[7] Petr Hájek: *Metamathematics of Fuzzy Logic*. Kluwer Academic Publisher, 1998.
[8] David Lewis: *Counterfactuals*. Oxford Basil Blackwell 1973.
[9] Jan Pavelka: On fuzzy logic I, II, III. Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 1979 (25), pp. 45–52, 119–134, 447–464.
[10] Judea Pearl: Causality: models, reasoning and inference. Cambridge University Press, 2000.
[11] Lofti A. Zadeh: Fuzzy Sets. *Information and Control*, 1965 (8), pp. 338–353.

Marco Cerami and Pere Pardo
Institut d'Investigació en Intel·ligència Artificial (IIIA – CSIC)
Campus UAB s/n 80193, Bellaterra, Catalunya
Email: {cerami, pardo}@iiia.csic.es