Mathware & Soft Computing 11 (2004) 5-11

Wide Sets, Deep Many-Valuedness and Sorites Arguments

C. Pelta
Sec. Psychology
Univ. Complutense de Madrid. 28040 Madrid, Spain.
carlospelta@inicia.es

Abstract

In this article I show how to obtain a powerful and truthful explanation of the failure of sorites arguments combining an adaptation of the Wide Set Theory formulated by Formato and Gerla and the concept of deep many-valuedness established by Marraud. It is shown that if the premises of a sorites argument are conceived as a succession of indexed consequence operators (where indices express the accuracy of the inferences) prefixing sentences, the argument fails because the transitive property for the inclusion between wide data sets is not always maintained.

Key words: Wide set, consequence operator, sorites argument.

1 Introduction

One of the most recent and interesting set-theoretic creations is Wide Set Theory (cf. [3]). In a general sense, a wide set is a finite set as big as we want it to be. Given a binary predicate EQ interpreted as a multivalued extension of the typical relation of equipotence between sets, the notion of infinite set may be modelized by a fuzzy subset representing the class of (finite) wide sets. Wide Set Theory is the ZF theory without the infinity axiom, and adding axioms which establish the fact that EQ represents an equivalence relation plus the following axioms:

1. $(\mathrm{EQ}(X, X\cup\{X\})\wedge(X\subseteq Y))\rightarrow \mathrm{EQ}(Y, Y\cup\{Y\})$ (monotonicity),
2. $\neg\mathrm{EQ}(\emptyset, \emptyset\cup\{\emptyset\})$ ($\emptyset$ is not infinite)
3. $\exists X\ \mathrm{EQ}(X, X\cup\{X\})$ (infinity axiom)

The authors of this axiomatization demonstrate the existence of a model associated with a function eq: $N \times N \rightarrow[0,1]$ which is the correlate of EQ such that (for an interpretation **I**), $\mathbf{I}(\mathrm{EQ})(X,Y)=\mathrm{eq}(\mathrm{card}(X),\mathrm{card}(Y))$. But more fundamental to our aim is the fact that on page 383 of their article they declare that sorites paradoxes are based on the notion of wide sets. Because this, in fact, is the main objective of our work: to explain the actual genesis and failure of this kind of arguments using the mighty construction of Formato and Gerla. In order to do this, I vividly recommend reading their original writings, as only in this way the reader will understand to what degree our "solution" owes itself to the tools employed by the Italian logicians. Another source of influence is to be found in the idea of deep many-valuedness developed by Marraud and Pelta (see [8],[10]). Until now the construction of superficial many-valued logics, that is, logics with an arbitrary number (bigger than two) of truth-values but always incorporating a binary consequence relation, has prevailed in investigations on logical many-valuedness. According to Marraud many-valuedness has been excluded from the consequence operation and the metalogic of a logical system, although its object language may be many-valued. Many authors are pleased to hold the existence of logics with a number bigger than two of truth-values but in which sentences of the kind "Sn+1 follows from S1,...,Sn" only can take the assignation of two truth-values. This is why Marraud speaks of deep many-valued logics, that is, many-valued logics in which the consequence relation is $n>2$-ary. He propounds graded systems of sequents (GSS) (cf. [8]) which, by mixing different consequence relations (in a similar style to the labelled deductive systems of Dov Gabbay [4], although Gabbay only uses labels for formulas), express the concept of deep many-valuedness. For a given ordered set $< I, \leq >$, a graded system of sequents would have rules of the form

$$\frac{X1 \vdash i\ Y1, \ldots, Xn \vdash j\ Yn}{X \vdash k\ Y}$$

with $i,j,k \in I$ being the subindices joined to the symbols of assertion and interpreted as meaning accuracy of the inferences. This system does not define a unique consequence relation but a family of them, although its definitions can be dependent between them. In a similar vein, Pavelka [9] was the first author introducing the concept of fuzzy consequence operator extending the notion of consequence operator in Tarski's sense; Chakraborty [2] used Fuzzy Set Theory for obtaining graded consecuence in many-valued logic while Castro and Trillas [1] develop an approach, which using fuzzy consequence operators, allows to explain the break of transitivity in argument chains that include vague predicates.

Given an initial set of data $D=\{\vdash 1A1, \ldots, \vdash 1An\}$, for every $i \in I$, it is convenient to introduce:

$Ci(D)=\{B: \vdash j\ B$, for some $j \geq i$, is GSS-provable from $D\}$.

If 1 is the maximally trustworthy item of I and 0 is the minimally trustworthy element, then it is obvious that $C1(D) \subseteq \ldots \subseteq C0(D)$, expressing the relative inclusion of the sets Ci(D) that the quantity of information provable from the set D of premises grows when the control of accuracy of the inferences is relaxed. Thus the point of equilibrium between initial information and accuracy of the reasoning would be the least $i \in I$ where Ci(D) is consistent.

Let us consider the following sorites argument (see [8, p. 64]):

1. Whoever is fifteen years old is young

2. Whoever is one month older than a young person is young

.

.

$\vdash$ Whoever is ninety years old is young

There are strategies for solving sorites arguments that resolve out the difficulty on the basis of the gradual character of the assertion of truth of fuzzy predicates such as young (see [5]), in this example. Nevertheless, from a deeply many-valued perspective the matter is not that the second premise in the argument is less true (with a determined degree of truth) than the first one: it happens that the second sentence displays a different consequence relation, that is, a weaker consequence relation. And so the second premise could be paraphrased as saying that "it is provable (more weakly than the proved by the first one) that one who is fifteen years plus one month old is young unless new contradictory information (consistency criterion) is introduced". Any reader can appreciate in this formulation a certain similarity with the prototype of a defeasible rule.

Let us consider data sets incorporating a determined consequence operator and being small theories. Let a database D={A1, A2, A3,...,An} closed under a consequence relation C1 so that C1(D)={$\vdash$1A1, $\vdash$1A2,...,$\vdash$1An}. In a simplified manner, we will consider a merely quantitative criterion (the number of premises contained) for distinguishing between databases closed under consequence relations: a theory C1(D) is more trustworthy than a theory C2(D) if and only if the quantity of information derivable from D is bigger under C2, that is, C2 is more lax than C1. If, e.g., D={Whoever is fifteen years old is young},

C1(D)={$\vdash$1Whoever is fifteen years plus one month old is young}; if C2 is applied then C2C1(D)={$\vdash$1Whoever is fifteen years old is young, $\vdash$2$\vdash$1Whoever is fifteen years plus one month old is young}. In this case, the premise "Whoever is fifteen years old is young" is the basis for the weaker deduction C2. In a more general sense a continuum of degrees of accuracy covering the threshold [0,1] is taken into account so that C1(D) would be the strongest theory and C0(D) would be the weakest theory (the theory with more defeasible consequences when new information is added).

2 Applying Wide Set Theory to sorites arguments

Let $I =< I, \leq \ >$ be a structure where I$\neq \emptyset$, i$\in$I; there is a function f: I$\rightarrow$[0,1] and $\leq$ is a partial order for indices belonging to I. Let g be a function such that g=h(f), being h: [0,1]$\rightarrow$ C, where C is a set of consequence operators $\vdash$0,...,$\vdash$1. Let D a data set and Ci(D) a data theory. Two consequence operations $\vdash$i and $\vdash$j are extensionally equivalent if and only if they close data sets of the same cardinality, that is, $\vdash$i$\equiv$$\vdash$j, iff card(Ci(Di))=card(Cj(Dj)).

Data theories Ci(Di) are collected in partition classes of an equivalence relation induced because consequence operators are extensionally equivalent.

So, $\vdash$i$\equiv$$\vdash$i (reflexivity),

($\vdash$i$\equiv$$\vdash$j)$\rightarrow$ ($\vdash$j$\equiv$$\vdash$i) (symmetry), and ((⊢i≡⊢j)∧(⊢j≡⊢k))→ (⊢i≡⊢k) (transitivity).

We are going to define the idea of degree of extensional equivalence between consequence operators (deec). In order to do so, first of all we define a function dee of degree of extensional equivalence:

dee: N × N →[0,1], where the domain of the function is the Cartesian product of data sets. Actually, dee is a typical measure of the fuzzy set theory (see e.g. [6] for an excellent comparison between fuzzy equivalence relations and pseudo-metrics) but we now adapt it for relating sets of sentences closed under consequence operators taking into account the cardinality of the sets. The degree of correlation between ⊢i and ⊢j is given by the difference of cardinality between Di and Dj. The more distance of cardinality between sets the more difference in the accuracy of the generated consequences from Di and Dj. Therefore, deec(Ci(Di), Cj(Dj))=dee(card(Di),card(Dj)).

According to the operations traditionally defined on t-norms (see [3], [11]), we can give a quantitative interpretation for dee: for m,n∈ N, and it is supposed that card(Di)=m and card(Dj)=n, it happens that dee(m,n)=min[m,n]/max[m,n]. It is possible to introduce a contextual function c for the measure dee so that, according to the nature of the premises of the argument, fix more or less conventionally the degree of separation between premises with respect to their derivational accuracy: deec(m,n)=dee(m,n)^c; c acts as a parameter for reflecting the influence of the context on the variation of the measure. If c takes a conventional value then dee(m,n)^c=(min[m,n]/max[m,n])^c. But do not forget that we want to apply these measures to lax theories closed by consequence operators Cn+1 from theories closed by operators Cn. That is, what I want to say is that sorites arguments have an argumentative structure similar to the form of a waterfall (if the reader accepts this metaphor): premises "spurt" from a given one to the next one, seemingly without a break. Certainly this is the key perspective of this essay and nothing better for reflecting it that the introduction of a function of measure ws adapted directly of the definition of wide set (see [3]). So, we have ws(n)=dee(n,n+1) and always considering the quantitative definition of dee(m,n)=min[m,n]/max[m,n], then ws(n)=min[n,n+1]/max[n,n+1] and ws(n)=n/n+1. If the parameter indicating context is used, we have ws(n)=(n/n+1)^c. Following Formato and Gerla [3, p. 385], definitional properties of our measures dee and ws would be:

1. ws(0)=0,
2. dee(m,m)=1
3. dee(m,n)=dee(n,m)
4. dee(m,n)≥dee(m,q)•dee(q,n),
5. $\lim_n$ ws(n)=1,
6. ws(n) is a monotonic function on N.

3 Some aspects of Williamson's solution

In his book entitled *Vagueness*, Williamson (cf. [12]) proposes an epistemic "solution" to sorites paradox. I so not wish to evaluate here its conceptual framework based upon the path sketched by Sorensen which subordinates the knowledge of an entity ei (Kei) to the truth of its precedent ei-1. But I shall be interested in the question of the failure of the transitive property for the conditional connective suggested by Williamson because it is related to our ideas about the failure of the transitivity for sets of premises of sorites arguments conceived like wide sets or data sets unfolded like the water of a cascade.

On page 123 of *Vagueness*, Williamson presents the following pattern of a sorites argument (being the symbol "→" by the conditional):

p0 (a totally true premise)
p0→p1
p1→p2...
p99.999→p100.000
----
p100.000 (a totally false conclusion)

Of course, the degree of truth of pn decreases in an almost imperceptible way. Williamson does not see the structure of the argument as a chain of progressively weaker consequences, but he reduces the matter to consider the degree of truth of the conditionals integrating the premises. The degree of truth of each premise exerts influence over the application of the rule of *modus ponens* (MP) which is not already totally valid. There are clues confirming that the author is worried not only by the traditional perspective based on the notion of truth, and he concedes importance to the idea of inferential validity. So, on page 124 of his essay, he suggests a fuzzy measure β-(1-α) of validity for degree β of the premises and for a degree α of the conclusion. Nevertheless, on the same page 124, Williamson wonders about the fulfilment of the transitive property by the conditional. Because it is obvious that transitivity is broken when the conditional premises are chained to consequents acting like almost totally true antecedents but finishing in a totally false consequent. Thus, quasi-truth is not transitively preserved by enchained conditionals. We think that a more precise explanation of this phenomenon is to be found in locating the failure of the transitivity on the inclusion between sets of premises conceived as sets of sentences prefixed by consequence operations and having progressively weaker or more defeasible indices of accuracy to the presence of new information (conceived as enunciative items) added. By this means, we wish to construct a model for data sets D closed by consequence operators. This model will be adapted to the pattern of a sorites argument.

4 Failure of transitivity in sorites arguments

Let us introduce a new relation GEE (graded extensional equivalence) so that it holds the following properties (correlates of the properties for EQ in [3, p. 384]):

1. GEE(Ci(D),Ci(D)) (reflexivity),
2. (GEE(Ci(D),Cj(D')))→(GEE(Cj(D'),Ci(D))) (symmetry),
3. (GEE(Ci(D),Cj(D'))∧GEE(Cj(D'),Ck(D")))→(GEE(Ci(D), Ck(D"))) (transitivity)
4. (D≡D')→((GEE(Ci(D),Ci(D")))↔GEE(Ci(D'),Ci(D"))) (compatibility).

We expect that the model reflects the behaviour "in cascade" of the sorites for data theories and by this reason we establish the existence of a wide set such that it holds GEE(Ci(D),Ci(D)∪{Ci(D)}). Naturally, the empty set of data does not hold the requisite. Besides I add and prove the property of inclusion between data sets closed under consequence:
(GEE(Ci(D),Ci(D)∪{Ci(D)})∧(Ci(D)⊆Cj(D')))→(GEE(Cj(D'),Cj(D')∪ ∪{Cj(D')})).

Proof. Let M = <C(D),**I**> be a model in which C(D) is the class of data sets closed by consequence relations and let **I** be the function of interpretation on the structure. **I**(GEE) is the degree of extensional equivalence and **I**(GEE)(Ci(D),Cj(D')) expresses the grade of extensional equivalence existing between theories Ci(D) and Cj(D'). For every Ci(D)∈C(D), **I**(GEE)(Ci(D),Ci(D)∪{Ci(D)}) is understood as the degree of unfold of Ci(D) and **I**(GEE)(C(D)) is the class of unfolded data sets under the consequence operator of the argument.

I(GEE)(Ci,Cj)=deec(Ci,Cj)=dee(card(Di),card(Dj)) and ws(n)=dee(n,n+1), dee being a measure of extensional equivalence and ws a measure of the unfold of a data set closed under consequence relation. We conclude that M satisfies
(GEE(Ci(D),Ci(D)∪{Ci(D)})∧(Ci(D)⊆Cj(D')))→GEE(Cj(D'),Cj(D')∪ ∪{Cj(D')}) iff dee(card(D),card(D∪{D}))≤ dee(card(D'),card(D'∪ ∪{D'})), for every D⊆D', iff ws(n) is a monotonic function on N.

We can get an interpretation for the inclusion between data sets (⊆D) using the approach of Kosko based on conditional probability (see [7]) by setting: **I**(⊆D)(D,D')=card(D∩D')/card(D), for every D≠ ∅.

(⊆D)(D,D') can be denotated as GEE(Ci(D)∩Cj(D'),Ci(D)) because in classical set theory, D⊆D' iff D∩D'=D. in fact, **I**(⊆D)(D,D')=1 for card(D)≤card(D'); and it is established that if card(D)≥card(D') then **I**(⊆D)(D,D')=card(D')/card(D).

Taking into account that conditional probabilities can be seen as fuzzy measures in the continuum determined by **I**(⊆D), transitivity has been broken because it does not hold ((⊆D)(D,D')∧(⊆D)(D',D"))→ (⊆D)(D,D") whenever for Ci(D), Cj(D'), Ck(D")∈C(D), Ci(D)∩Cj(D')≠ ∅ and Cj(D')∩ ∩Ck(D")≠ ∅ but Ci(D)∩Ck(D")=∅. Actually this circumstance happens with respect to the conclusion of the sorites : it does not coincide at all with respect to the precedent neighbouring consequents. So it is explained that while the grade of accuracy of the first application of ⊢1 is absolute, the grade of accuracy of the last turnstile is null, that is, the expression of the greatest derivational laxness. This is the proper understanding of the derivational failure of sorites arguments conceived as arguments by a succession or chain of consequence operators, each one generating weaker or more defeasible theoıies than the precedent ones. And with this explanation my aim has been to offer a deeply many-valued version of that kind of problematic arguments.

Acknowledgement

I thank the anonymous referee comments which have improved this article.

References

[1] J.L. Castro and E. Trillas, "Tarski's Fuzzy Consequences", *Proceedings of the International Fuzzy Engineering Symposium 91*, vol. 1, 70-81, Yokohama, 1991.

[2] M.K. Chakraborty, "Use of Fuzzy Set Theory in Introducing Graded Consequence in Multiple Valued Logic, in *Fuzzy Logıc in Knowledge- Based Systems, Decision and Control*, M.M. Gupta and T. Yamakawa (eds.), (North-Holland, Amsterdam, 247-257).

[3] F. Formato and G. Gerla, "Grasping Infinity by Finite Sets", Mathematical Logic Quarterly 44 (1998), 383-393.

[4] D. Gabbay, *Labelled Deductive Systems.* (Oxford University Press, Oxford, 1996).

[5] J. Goguen, "The Logic of Inexact Concepts", Synthese 19 (1968-69), 325-373.

[6] F. Klawonn and J.L. Castro, "Similarity in Fuzzy Reasoning", Mathware and Soft Computing 2 (1995), 197-228.

[7] B. Kosko, *Neural Networks and Fuzzy Systems.* (Prentice- Hall, Eaglewood Cliffs, 1992).

[8] H. Marraud, "Razonamiento aproximado y grados de consecuencia", Endoxa 10 (1998), 55-70, (in Spanish).

[9] J. Pavelka, "On Fuzzy Logic I, II, III, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 25 (1979), 45-52, 119-131, 447-464.

[10] C. Pelta, "Deep Many-Valuedness", Logique et Analyse 167-168 (1999), 361-372.

[11] E. Trillas, C. Alsina and J.M. Terricabras, *Introducción a la lógica borrosa.* (Ariel, Barcelona, 1995), (in Spanish).

[12] T. Williamson, *Vagueness.* (Routledge, London, 1994).