Title
Truth and Approximations

Journal
Proceedings of the Annual Meeting of the Berkeley Linguistics Society, 3(3)

ISSN
2377-1666

Author
Sadock, Jerrold M

Publication Date
1977

Peer reviewed

Truth and Approximations

Jerrold M. Sadock
University of Chicago

A system of language description whose principal goal is the construction of a truth definition for natural-language sentences would seem to be in trouble with regard to approximations such as (1) if it provides only two possible truth values, absolute truth and absolute falsehood.

(1) Sam is approximately six feet tall.

As George Lakoff (1972) argued, the goodness of such sentences with respect to a certain state of affairs is a smoothly varying commodity. Example (1) is a better description of the fact that Sam is 5'11" than it is of the fact that Sam is 5'10". there does not seem to be any sharp dividing line between those states of affairs to which (1) is applicable and those to which it is not.

Lakoff proposed to remedy this apparent deficiency of two-valued logic as a basis for natural-language description by substituting for it a system of fuzzy logic along the lines of Zadeh (1971), in which the truth value of sentences **is** allowed to assume any value between and including absolute truth and absolute falsehood. In treating fuzzy presuppositions, Lakoff indicates that sentence (2) would have variable truth values in his fuzzy logic.

(2) Sam has approximately $10,000 in his savings account.

He says on page 222 of the article referred to above that if it were the case that Sam had $9,992 in his account, "...(it) would be true no matter what." And if Sam had $9,950, "... most people in most situations would still want to say...(it)... was true..." If he had only $9,500, "...in many situations... (it)...would have a high degree of truth..." But when Sam's wealth shrinks to $9,200, "...the degree of truth...gets lower." The truth value of (2) on Lakoff's theory would thus appear to be some function of the difference between the actual amount of money that Sam has in the bank and the amount that is mentioned in the approximation.

Now what sort of function is this to be? Let us consider first a simple function of the ratio between the error and the guess. In (3), A is the approximation, V is the correct value, and the vertical lines indicate absolute value.

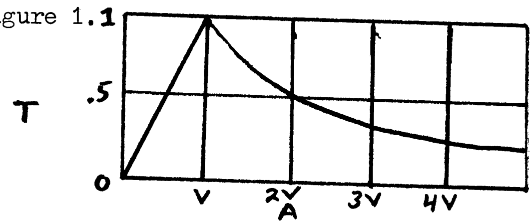
(3) $$T = 1 - \left| \frac{A - V}{A} \right|$$

Unfortunately, this function assumes negative values when the approximation is less than half the actual value and plunges to minus infinity when the approximation itself is zero. Yet expressions like <u>approximately zero</u> are fine. Furthermore it seems that bad guesses that are slightly on the low side are not quite as bad as bad guesses an equal amount off on the high side. Both of these failings are corrected in (4), where the square bracket and comma notation indicates the larger of A and V.

(4) $$T = 1 - \left| \frac{A - V}{[A, V]} \right|$$

Figure 1. gives a plot of this function with truth value on the vertical axis and the value of the approximation, expressed in multiples of the actual value, on the horizontal axis.

Figure 1.

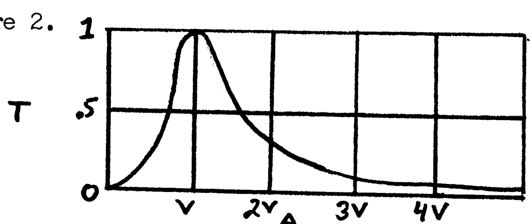


A remaining problem with (4) is that it gives much **too high** values of truth to spectacularly bad guesses. We can rectify this problem by raising the ratio in (4) to some power, as in (5).

(5) $$T = 1 - \left(\left| \frac{A - V}{[A, V]} \right| \right)^n$$

Figure 2. plots this function for n=3.

Figure 2.



The higher we make n, the closer we require the approximation to be for similar truth values under the same circumstances. A feature of (5) that I consider a benefit is that there is only one state of affairs under which an approximation is completely false, namely in case the actual value in question is zero and the estimate is not. Thus sentence (6) is false, while (7) merely has a very, very low truth value.

(6) A geometric point is approximately one tenth of an inch wide.

(7) A drop of water contains approximately one hundred molecules.

It seems to me that any analysis that brands approximations false, except where the property involved is not present at all, is wrong.

While (5) does have this desired property, it still needs some work. Let us compare (1), which I repeat here as (8), with (9).

(8) Sam is approximately six feet tall.

(9) That cockroach is approximately six feet tall.

If Sam is , say, 5'8" tall, I think (8) would be taken as a pretty poor estimate. But if the cockroach mentioned in (9) is 5'8" at the withers (or wherever one measures the height of cockroaches to), (9) would not seem such a bad estimate at all. It would seem, then, that the nature of the item that we make an estimation about has an effect on how the estimate is evaluated. Perhaps in the cases above it is the fact that adult human beings (which, let us assume, Sam is) ordinarily have heights that fall within a rather narrow range, say five to seven feet, that makes (8) so much worse an approximation than (9), even when both sentences miss the mark by the same amount. While cockroaches are usually pretty small, the fact that a six-footer is mentioned in (9) gives us a scale of roughly zero to at least six feet, a much larger scale than we find in (8). This effect of the approximatee can be brought into the equation by making the exponent, n, depend on the size of the scale involved. The smaller the scale, the more accuracy is required and hence, the larger n should be. Now one end of the scale will be determined by the guess itself, if the guess falls outside of the usual range of range of values for the kind of entity involved. Let M_a^+ be a value of some property such that,say, 95% of all members of the group with which the approximatee is being compared have the property to no greater extent than M_a^+. Similarly, M_a^- will be the lower reasonable limit of the scale. The scale, then, will be the largest of the three quantities, $A - M_a^+$, $A - M_a^-$, and $M_a^+ - M_a^-$. Equation (10) takes all of this into account.

(10)

$$T = 1 - \left(\left| \frac{A - V}{[A,V]} \right| \right)^{\frac{C}{[A-M_a^-,\ A-M_a^+,\ M_a^+-M_a^-]}}$$

Here C is some constant which could, presumably, be determined empirically.

But we still are not done trying to bring the truth function for approximations into line with the way they are actually judged. Compare (11) and (12).

(11) Odessa has a population of approximately one million.

(12) Odessa has a population of approximately 990,000.

Let us suppose that the actual population of Odessa is 980,000. Now, surprisingly, (11) is a better approximation than (12) even though the error in (11) is 2% while that of (12) is only a little more than 1%.[1] What seems to be going on here is that (12) has more significant figures than (11) and consequently involves a smaller scale. This effect only comes in to play if the scale established by the last significant figure is smaller than the scale established by the nature of the approximatee. But in any case, the exponent in expression (10) will have to be made more complicated so as to take into account the effect of the form of the approximation. For the most part, all consecutive zeros with no nonzero to one side are not considered significant. While it is only an approximation, let us suppose that the scale can be obtained from the figure that is mentioned in an approximation by subtracting fifty percent from, and adding fifty percent to the last significant figure. Thus the scale set by the numeral 1000 is 500-1500, but that set by 990 is much smaller, only 945-1035.

That it is the form of the approximation, rather than its content, that is operative in determining the scale by which the defensibility of an approximation is to be judged can be seen from the following two considerations. First, a change in the units involved produces a change in the perceived accuracy of the approximation, even though it does not necessarily involve a change in the actual magnitude of the guess. Suppose I tell you that Sam has $10,000 in his Canadian bank account. Now $10,000 in Canadian dollars is worth roughly $9,700 in U.S. dollars. But if I make the same guess in terms of U.S. dollars, that is, if I tell you that Sam has the equivalent of approximately $9,700 U.S. in his Canadian account, you will read me as knowing more about his finances than in the first case. Similarly, an estimate of a mile is taken as cruder than an estimate of 5280 feet, which itself is cruder than an estimate of 63,360 inches. The second thing is that the very same figure can be taken as more or less accurate, depending on exactly how it is put. <u>About a dozen</u> is somehow rougher than <u>about twelve</u>, <u>approximately two and a half tons</u> is not as accurate sounding as <u>approximately two-point-five tons</u>, and so on.

Neglecting these niceties, let me just say that Σ is the scale determined by the form of the estimate. Now the exponent in our truth function involves the larger of the two scales, the one determined by the nature of the approximatee and the one that depends on the form of the approximation. The exponent will look something like (13).

$$(13)\qquad \left[\frac{C}{A-M_a^+,\ A-M_a^-,\ M_a^+-M_a^-},\ \frac{K}{\Sigma}\right]$$

Here K is another constant, also presumably to be determined empirically. The whole truth expression then becomes:

$$(14)\qquad T = 1 - \left(\left|\frac{A - V}{[A,V]}\right|\right)^{\left[\frac{C}{[A-M_a^+,\ A-M_a^-,\ M_a^+-M_a^-]},\ \frac{K}{\Sigma}\right]}$$

If the reader is growing suspicious of this increasingly inelegant equation, I am not surprised. The more it is made to fit our impression of what determines the defensibility of an approximation, the more it diverges from an honest representation of a purposely, and unabashedly inaccurate statement, which is what an approximation is. Furthermore, and more importantly, nothing that I have observed about the various contingencies that seem to play a role in the evaluation of the validity of approximations is really true.

Let me return to the giant cockroaches, with apologies to the squeamish. Let us imagine that at a large state university, roaches roughly the size of human beings have been bred. Now sentence (9), uttered by one of the laboratory technicians involved in this Kafkaesque experiment, would be considered pretty inaccurate if the insect mentioned in (9) were only (!) 5'8". The average cockroach is still a thankfully small thing, but because of the special circumstances, the apparent degree of confidence increases greatly.

When we compare (11) and (12) again, we can see that the discrepancy in implied accuracy disappears under special circumstances. If, for example, (11) is uttered in the context of an argument over whether Odessa is larger than Kiev, which, let us say, is known to have a poulation of 995,000, then (11) would appear to have more significant figures than it did out of context and, indeed, more than (12) has out of context. We see, then, that it is only in the absence of special circumstances that the form of an estimate appears to correlate with its implied degree of accuracy.

What it seems to me is really going on in all the examples discussed so far is this: it is the purpose of the estimate that essentially determines how close to the truth it must be to be warranted. Various facts about the form and content of the approximation can suggest part of the purpose of the approximation, or only seem consonant with certain intentions of the speaker, but these are merely suggestions, and not part of the conventional content of the approximations. These suggestions, as we have seen, can be readily cancelled by specific situational factors and can be denied without contradiction. An accountant could easily say something like (15), and (16) could be felicitously uttered by someone who had only seen Sam drive by in a car.

(15) Sam has approximately $9,983 in his savings account, give or take $100.
(16) The best I can say is that Sam is about six feet tall, give or take four inches.

But all of this means that an explicit truth definition that takes the factors I have discussed into account cannot, indeed should not, be written. The purpose of an approximation is infinitely variable and cannot be encoded in any direct fashion in the linguistic description of sentences.

If a truth definition of some kind is desired, that is, if approximations are to be claimed to be subject to judgments of truth and falsehood at all, then I suggest that the definition will have to be a fairly trivial one. It is always possible to think of situations that will make any approximation, no matter how far off base, at least somewhat defensible. The only case where it seems at all reasonable to label an approximation plain false is when the property in question is one that the approximatee does not have to any degree whatsoever. Even in cases where it is logically impossible for the approximation to be completely accurate, it does not seem right to call an approximation false. I can imagine situations under which (17) would be a fully reasonable thing to say.

(17) Six has approximately five divisors.

The requisite truth definition for an approximation is therefore one that makes it true in all circumstances, or one that makes it true unless the approximatee does not have the property at all. According to this definition, all approximations would have the same truth value under all, or almost all, circumstances. But then how can the fact be explained that their effects clearly differ? Why, in other words, should (18), (19), or indeed (20) not be interchangeable in all contexts?

(18) Alligators have approximately fifty teeth.
(19) Alligators have approximately a thousand teeth.
(20) Linguistics is taught at approximately a thousand universities.

The obvious and, I believe, correct answer is that while these might not differ in possible truth values, they do differ in sense. Just as in the case of Frege's (1975) famous examples, approximations may not be interchangeable because the sense of the components of the statements is different even if their reference is the same. Even more to the point is Grice's analysis of the import of obvious tautologies. His examples are Women are women, and War is war. Though these not only share the same truth value – they have the same truth conditions – they are nevertheless not useful with the same degree of appropriateness in all the same contexts. Grice says (1975,70):

> "They are, of course, informative at the level of what is implicated, and the hearer's identification of their informative content at this level is dependent on his ability to explain the speaker's selection of this particular patent tautology."

I propose that approximations are to be analyzed in the same way. They are so devoid of real semantic content that they simply call attention to the particulars of the form that is chosen. Although the dentition of alligators cannot have an influence on the truth of statements like (18) and (19), it would be misleadingly irrelevant to choose to use the words alligator and teeth if the speaker did not want the addressee to think that something about them was being hinted at. Grice's maxims of quantity and manner are instrumental in the interpretation of approximations, just as they are in the interpretation of tautologies.

I wish to conclude this exercise in linguistic pragmatics by pointing out a few positive advantages that attach to the account of approximations that I have argued for. First of all, if approximating expressions were truth functional, as they are in the fuzzy semantic approach, I can see no reason why they could not take already inexact expressions as arguments. I can see nothing in the fuzzy semantic treatment that would rule out approximations of approximations, yet these are bad.

(21) *Sam is about approximately six feet tall.

According to fuzzy semantics, (21) ought to be grammatical and ought to have a meaning something like the exaggerated approximation, (22).

(22) Sam is very roughly six feet tall.

I can likewise see little reason for the fuzzy semantic account to rule out examples (23) and (24).

(23) *Sam has approximately some money in his savings account.

(24) *Sam has written approximately a few/ several/ many books.

In a fuzzy semantic theory expressions like some, a few, several, and many are presumably truth functional. They differ from the numerical quantifiers and the quantifiers all, every, none, etc., only in that their semantics are fuzzy rather than discrete. One should therefore expect that an approximator used with a vague quantifier would produce a meaningful expression that is just somewhat less precise. But such expressions are, as (23) and (24) show, ungrammatical.

In the pragmatic theory that I am bucking for, on the other hand, an explanation of the ungrammaticality of these examples is forthcoming. The role of an approximator in the pragmatic theory is to trivialize the semantics of a sentence, to make it almost unfalsifiable, to hedge in a genuine sense. Double approximations would therefore be ruled out since a single approximator does as much semantic trivializing as is possible. A second one would be an egregious redundancy. Intensifications will still be possible as indications of diminished confidence, just as expressions like possibly and just possibly differ in the degree of confidence that they indicate without differing in semantic content per se.

The mid-scalar quantifiers some, a few, several, and many will require some comment if my theory is to explain the badness of (23) and (24). These would seem to be approximations in and of themselves, and I propose to treat them as such. Nearly everything I have said about the difficulty of finding an explicit truth definition for approximations applies directly to them. It is difficult, if not impossible, to find a situation in which (25) is clearly false, except in the case where there are no entities that meet this description.

(25) A few professors drive Volvos.

And, as was the case with approximations, the actual number that it takes to justify the use of one of these expresssions varies with the apparent purpose of the utterance. A few Supreme Court Justices are probably a lot fewer than a few stars. But, of course, there are differences among these various quantifiers. They stand in an implicational hierarchy, as described in Horn 1976, such that many implies several, several implies a few, and a few implies some. It seems to me that this is pretty much all that has to be said about these to give a fairly good account of their use and effect. Whereas Horn treated the inexact quantifiers as semantially lower bounded but only conversationally upper bounded, I would like to treat them as both conversationally upper and lower bounded. It is the second maxim of quantity that provides the upper bound,("Do not say less than is required...") but the first maxim of quantity that provides the lower bound ("Do not say more than is re quired..."). Saying many is conventionally indicating more than saying several without making any different semantic commitment. Thus on my theory, it is principally for pragmatic reasons that these inexact quantifiers seem to spread out at arms length.

This long, but still sketchy description of inexact quantifiers is supposed to do no more than make it plausible that they, too, have rather trivial semantics. If this is so, then it is quite natural that further trivialization by means of approximators is impossible.

The second, and only additional positive argument that I have for the pragmatic treatment of approximations comes from the observation, backed up by consulting dictionaries and by the results of an informal (and probably inept) survey that I conducted, that simple approximators do not differ from one another in the communicated degree of accuracy of the approximation. There seems to be no consistent ability on the part of speakers of English to tell which of the words <u>circa</u>, <u>about</u>, <u>around</u>, <u>roughly</u>, and <u>approximately</u> convey greater precision. I am not saying that there are no differences among them at all; indeed they must differ in some way, or the contrasts in (26) would be inexplicable. But it is the case that these approximators do not differ in any striking way as to how close the approximation must be to the truth to be defensible.

(26) John ate approximately/ ?roughly/ ?about/ ?*around/ *circa all of the beans.

Once again, the semantic theory of such forms offers no non-ad hoc explanation for this phenomenon. Certainly it is possible in principle to describe spikier and more gentle fuzzy truth functions. Why then don't these words display such differences? On the pragmatic account, though, a hedge is a hedge. All of these would have to be alike in their ability to turn sentences with interesting, falsifiable semantics into sentences with uninteresting, almost unfalsifiable semantics and import that is almost completely a matter of our knowledge of the rules of cooperative conversational behavior.

Footnote

[1] A similar observation is made in Heinämäki 1975.

Bibliography

Davidson, Donald and Gilbert Harman, eds. (1975) <u>The Logic of Grammar</u>, Dickenson, Encino, California.

Frege, Gottlob (1975) "On Sense and Reference," in Davidson and Harman 1975, pp. 116-28.

Grice, H. P. (1975) "Logic and Conversation," in Davidson and Harman 1975, pp. 64-74.

Hienämäki, Orvokki (1975) "Almost," in Even Hovdhaugen, ed., <u>Papers from the Second Scandinavian Conference of Linguistics</u>, Lysebu.

Horn, Laurence Robert (1976) On the Semantic Properties of Logical Operators in English, reproduced by the Indiana University Linguistics Club, Bloomington, Indiana.
Lakoff, George (1972) "Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts," in P. Peranteau, J. Levi, and G. Phares, eds., Papers from the Eighth Regional Meeting of the Chicago Linguistic Society, Chicago Linguistic Society, Chicago, Illinois.
Zadeh, Lofti (1971) "Quantitative Fuzzy Semantics," Information Sciences, 3.