

# What does it take to refer? – A reply to Bojadžiev

Don Perlis\*  
University of Maryland  
College Park MD 20742 USA  
perlis@cs.umd.edu  
www.umd.edu/perlis  
301-405-2685  
fax: 301-405-6707

April 4, 2001

## Abstract

Bojadžiev has taken issue with my distinction between *strong* and *weak* self-reference, in saying that it is reference in general and not simply self-reference, that either is strong or weak. I agree completely. Here I clarify how I intend those notions and why I think that the strong case of self-reference is worthy of special attention. In short, I argue that *all* forms of referring involve a kind of self-referring.

I appreciate the opportunity provided to me by the Editors, to respond to Damjan Bojadžiev's interesting paper. Bojadžiev has taken issue with my notion of *strong self-reference* (SSR), in saying that it is reference in general and not simply self-reference, that either is strong or weak. With this I have no quarrel. I can however offer the following expanded explanations.

When we say “this refers to that”, as in

‘Water’ refers to  $H_2O$

---

\*This research was supported in part by a grant from ONR/AFOSR.

we suppress mention of the agent(s) who use ‘water’ to so refer. Thus the word ‘water’ in itself does no referring. We may call this *weak* reference: it is reference delegated to an unspecified referring agent or community of agents who share a common language.

Moreover, the agents who so use ‘water’ to refer to that liquid do so intentionally, that is, they do not simply mark the letters that appear in ‘water’ on cards and mindlessly float them in pools of water. Rather there is a particular kind of mental tagging activity. Referring is a deliberate act of using one thing, a symbol, for another, the referent. We do this in part because, as Swift makes vivid in *Gulliver’s Travels*, we simply cannot display all the things we wish to talk about; we must use substitutes. There are at least two reasons for our deficient ability here: (i) we have neither the strength nor the temporal position to carry the multitude of things (past, present and future) that we refer to, and (ii) we often do not even *know* in detail what we refer to.<sup>1</sup>

It is this intentional or deliberate act of using a symbol as stand-in that characterizes strong (or perhaps *genuine* would a better term) reference. Bojad/vziev is quite right that any kind of referring can be either strong (ie, explicitly involving a referring agent who makes the deliberate link between symbol and referent, albeit a tenuous link as in the case of “elm” in the earlier footnote) or weak (suppressing the implicit referrer). But when we suppress the referrer, we no longer have an actual case of reference before us. Thus the famous Liar Sentence, ‘This sentence is false,’ does not in itself refer to anything at all, let alone to itself. Or better put: there are two senses of referring: the primary one is what I am calling strong, in which a referrer mentally tags the referent with the symbol; and the secondary or weak one in which, having already accomplished strong reference, we say elliptically of the symbol that it now does the referring.

Thus there is no such thing as weak reference in the absence of strong reference. That is, strong reference genuinely does something. It is an action or process that achieves a referential goal, whereas weak reference is a static form that performs no action at all but that is often *taken by* a referring

---

<sup>1</sup>The latter has been made very clear by Putnam [2] and Kripke [1]. We may not know identifying properties of elm trees, for instance; we rely on the fact that experts can tell them from other trees. (But even the experts cannot carry them about, and instead must often rely on words to convey ideas about elms!) Nevertheless, as stated above, when we use a symbol to refer, even without our knowing much about the referent, we do take ourselves to be referring rather than merely making noises or scribbles.

agent to *represent* a referential relation. The agent actually performs the referential act.

But I want to claim more than this. I claim that not only is strong reference requisite for weak reference, but strong *self*-reference is requisite for (all forms of) reference. This is because, as noted above, there must be a referrer in order for there to be reference, and *to be a referrer requires taking oneself to refer*: it is an intentional or deliberate act. One uses a symbol to stand for a referent in a mental act in which the referrer has both the symbol and the referent in mind and joins them *so that* the symbol becomes a stand-in for the referent *for the referrer*. This requires the referrer to acknowledge his or her own role. To refer is to place oneself in the role of authorship; it is an internally represented three-way relation between oneself, the symbol, and the referent. And so referrers must represent themselves in their referring.

Thus I make the following hypothesis: strong self-reference is the most basic kind of reference, without which reference is impossible.

What kind of processes can perform this seemingly circular act of representing themselves as themselves to themselves, remains unclear. But it seems a bit clearer that we ourselves do in fact accomplish something very like this. We note ourselves as self-noting processes.

Bojad/vziev remarks (end of section 3) that my characterization of strong self-reference in terms of an action seems to apply to any conscious action and thus that consciously reaching for a pencil is as self-referential as is reaching for one's forehead. This I think misses the point I was (admittedly too briefly) trying to make. In reaching consciously for a pencil, on the definition I offered, one not only does the reaching movement but also performs another action, namely one of representing to oneself that one is not only reaching but that one is doing that very representing. Thus one is after all pointing to oneself in consciously reaching for a pencil, though perhaps the pointing in the case is not to one's forehead but rather to a more personal inner self. It is as if one were to mumble "here I am reaching for a pencil", and in the "here I am" one is strongly self-referring. The expanded version would include perhaps "here I am saying this to myself."

Of course I do not mean to say that we usually find ourselves being so verbally explicit; but I do mean to suggest that we nevertheless perform such strong reference each and every time we refer to anything at all, ie, at every moment of being conscious at all.

To comment briefly on Bojad/vziev's use of mirrors as a model of self-reference: to the extent that one takes into account the actual physical pro-

cess involved (that is, the actual changing state of the self-observer as it sees its own image) then we have gone outside traditional modes of formal description, and into action. That is already is a large step, but not enough in my view to achieve (strong) self-reference. Somehow the image must be taken as being of oneself, that is, the very taking of it must be part of what is being taken.

This is the circular-sounding curiousness of strong self-reference that I have advocated. I am embarrassed that I cannot clarify it better, but it seems to me that this is the sort of thing that we must be prepared to struggle with until it yields to our investigation. Perhaps we can take heart in the thought that, not too long ago, it was inconceivable that a molecule could self-replicate. Note that this is more than mere forming of copies, it is *self*-replication; for a particular instance of a DNA molecule itself accomplishes the creation of two copies of itself. We now know how this happens, and it took some very clever chemistry to work it out. Self-referring is not self-replication, to be sure, defined by clearcut start and end states. So our job may be harder. One task is to try to build self-referring agents, aiming particularly at behaviors that may make the clearest use of referring, as I sketched in my paper.

## References

- [1] S. Kripke. *Naming and Necessity*. Harvard University Press, 1980.
- [2] H. Putnam. The meaning of meaning. In H. Putnam, editor, *Mind, Language, and Reality*. Cambridge University Press, London, 1975.