Supervaluationism, Laziness and Constructive Logic

Thomas Forster

February 21, 2018

ABSTRACT

We claim that the various sharpenings in a supervaluationist analysis are best understood as possible worlds in a Kripke structure. It's not just that supervaluationism wishes to assert

> $\neg(\forall n)$(if a man with n hairs on his head is bald then so is a man with $n+1$ hairs on his head)

while refusing to assert

> $(\exists n)$(a man with n hairs on his head is bald but is a man with $n+1$ hairs on his head is not)

and that this refusal can be accomplished by a constructive logic (tho' it can)—the point is that the obvious Kripke semantics for this endeavour has as its possible worlds precisely the sharpenings that supervaluationism postulates. Indeed the sharpenings do nothing else. The fit is too exact to be coincidence.

Introduction

Some basic knowledge of possible world semantics for formal logics is assumed. Our accessibility relation will always be a partial order, written '$\geq$' and although there will be no modal operators here for the possible worlds to give us semantics for, we will make use of the following semantics that possible world semantics affords:

$$W \models (A \to B) \ \text{ iff}_{df} \ \ (\forall W' \geq W)(\text{if } W' \models A \text{ then } W' \models B)$$

that defines what it is for a world to satisfy a constructive conditional.

One of the things that bothers people is the thought that our analysis of the concept of baldness might tell us that it has an extension that is both determinate and concealed from us. This is clearly distasteful, but we shouldn't mock ourselves with exaggerated hopes of avoiding this kind of outcome altogether:

something like this is going to happen. The central problem with vague concepts is: how is it that we are ever able to get *definite* answers to *vague* questions? Any theory of vague predicates must explain how this can happen. Because we do get definite answers. It may look tempting to try to work up many-valued logics into a solution for vagueness, but if you try it you then have to explain the process whereby all those extra truth-values get thrown away. Because thrown away they definitely are. If someone were to ask whether or not I am bald the answer isn't "well, 2/3" or "$\pi^2/12$". It's " 'fraid so".

Entrapment in a paradox is like entrapment in a novel or a joke: all three entrapments rely on willing supension of disbelief, and they all promise the victim a reward.

Paradoxes usually involve bringing out some feature of ordinary reasoning that is potentially problematic: that's why they can be interesting and instructive. The set-theoretical paradoxes rely on—and bring out—the assumption that to every intension there corresponds an extension. But the [alleged] paradox before us here relies on *suppressing* a feature of ordinary reasoning about vague concepts, a feature that makes it unproblematic. Let me explain.

The Connection with Lazy Evaluation

Vague concepts are everywhere, as we all know. And, as we all know, real life reasoning with vague concepts is unproblematic: plenty of people who have not written Phil 301 essays on vagueness cope adequately with the routine vague demands of real life. Realistically, when asking whether or not someone is bald, it's with respect to some purpose or other. "Is he too bald to ask out this 25-year-old?"; "Is he so short of hair that (given his age) you think he might be sickening for something?"; "Is he going to need a wig to play the male lead?" Again, the reason why do not not lose sleep over whether or not a particular collection of sand grains is a heap is that in situations where the answer to this question matters to us there is background information that tells us more precisely which question is being asked and thereby steers us towards an answer. If it contains fewer than x grains and is on a building site then the answer is 'no'. If it is on a desk in a room where someone is giving a talk about Sorites then the answer is 'yes'. *Dead* has many physiological sharpenings (some of them with legal significance); *adult* has a variety of sharpenings, and the following table (from [1]) sets out those that have significance in UK law.

Starting compulsory schooling	5
Criminal responsibility	12
Making binding contracts (Scotland), girls	12
Baby sitting	14
Being lent a shotgun, to use without certificate on owner's premisses	14
Making binding contracts (Scotland), boys	14

Entering licensed premisses	14
Being given a shotgun, if holding certificate	15
Stopping compulsory schooling	16
Buying cigarettes	16
Heterosexual intercourse	16
Driving motorcycle	16
Marriage without parental consent (Scotland)	16
Buying/consuming cider/perry on licensed premisses	16
Driving car	17
Buying or hiring shotgun or ammunition	17
Marriage without parental consent (England)	18
Voting	18
Making binding contracts (England)	18
Buying or consuming alcohol on licenced premisses	18
Homosexual intercourse (male)	21
Adopting a child	21

There are other signs of adulthood not legally defined, such as: the age at which a French or German speaker has to be addressed as 'sie' or 'vous', and no longer as 'du' or 'tu'.

The moral of this is that in all cases where we are worried, there are actually lots of different questions that are being asked ("Is this a heap in sense A?", "Is this a heap in sense B?" and so on.

So what is the trick by which you manufacture a puzzle about reasoning with vague concepts? The trick is to remove the cues that tell you which precise question is being asked.

The apprehension that there is background detail being concealed gives rise to *supervaluationism* (tho' that is of course not the way the story is told).

The way to cope with these infinitely many questions is a practice known to logicians as *lazy evaluation*—which we had better sketch briefly. Under any assignment of truth-values to primitive propositions, a complex expression evaluates to a truth-value. Although this truth-value is of course completely determined by the truth-assignment to primitive propositions and the structure of the complex expression, there is more than one way of computing it from the assignment and the structure. Two of those ways are 'eager' evaluation, and 'lazy' evaluation.

If we evaluate the truth-value of $A \wedge B$ "eagerly" we evaluate the two truth-values of A and of B and then take their conjunction. In contrast if we evaluate the truth-value of $A \wedge B$ "lazily" we evaluate the truth-value of A first, and only if it is `true` do we bother to evaluate the truth-value B. (If it is `false` we know the truth-value of $A \wedge B$ to be `false`). More generally, when evaluating lazily the truth-value of $p_1 \wedge p_2 \wedge p_s \ldots p_n$ we calculate the truth values of the p_i starting at 1 and stopping as soon as we see a `false`; only if we see nothing but `true`s do we persist to the end.

In striving to ascertain whether or not a propositus is bald it may be that quite a lot of our investigation is not particularly sensitive to the concept of baldness in play, and that we can procede at least part of the way pending information about which concept is in fact in play.

The lazy strategy is to delay asking which concept of baldness is in play until we reach a stage in our investigation where the answer to this question might make a difference. Indeed, in order to break the logjam it might not even be neccessary to know precisely which notion of baldness is in play, but it might be sufficient merely to narrow down the range of possibilities somewhat. We can conclude our investigation once we have narrowed down the range of possibilities to a set of predicates which all give the same answer. Then we report that truth-value.

Observe that we use the same strategy of lazy evaluation when choosing to reason informally rather than formally. One of the (many) situations in which we do not bother to be formal is when we are in a context where the various formalised versions of what we are reasoning about all give the same answer.

The Connection with Constructive Logic

One of the standard problems with Sorites is the *slippery slope*. Consider the conditionals $C(n)$ (with $n \in \mathbb{N}$): If a man with n hairs on his head is bald, then a man with $n+1$ hairs on his head is likewise bald. We clearly cannot accept *all* these conditionals, but if we are to reject one of them and break the chain, which one should it be? One is reminded of the old puzzle of the chain all of whose links are equally strong: it is unbreakable because there is nowhere for it to break. There is no *prima facie* reason for choosing any particular n, and every choice is arbitrary. Nevertheless we seem to be forced to plump for some n or other, because $\neg(\forall n)C(n)$ implies $(\exists n)\neg C(n)$. In realistic situations—*vide, e.g.* the table above—there are cues to tell us which of the C_n to reject. The conjuror's art in setting up a slippery slope argument is to withhold any cue that might have helped the victim decide which C_n to reject.

If we take the predicates between-which-we-are-equivocating to be possible worlds, we obtain a Kripke model for a constructive logic in which this inference is blocked. Indeed we can do this in more than one way.

Consider the following Kripke structure.

There are worlds W_i for each $i \in \mathbb{N}$, and W_0 is the designated world. Each world knows enough set theory and arithmetic to count the hairs on my head. The language also contains a one-place predicate 'B', and the reader can surely guess what it means. The worlds will all have the same inhabitants. Bear in mind that the refusal of a world to believe p is **not** the same as that world believing $\neg p$. A world believes $\neg p$ iff no world accessible from it believes p. (This is a special case of the definition for conditionals just displayed, since $\neg p$ is just $p \to \bot$.)

We define the extension of the predicate letter B at each world by stipulation, as is usual, and we stipulate that $W_i \models B(x)$ as long as $|x| \leq i$. $|x|$ is of course the number of hairs on x's head so this is saying that worlds with larger and larger subscripts believe more and more people to be bald. In world W_i you are deemed bald if you have no more than i hairs on your head; worlds with larger subscripts are more trigger-happy. Observe that an object bald in one world is bald in all worlds with larger subscript.

Do we want W_i to be able to see W_j for $j > i$? Perhaps we want none of the worlds with nonzero subscripts to be able to see each other. We will explore both these options. Both options result in an analysis where W_0 believes

$$\neg(\forall n)((\forall x)(B(x) \rightarrow |x| \leq n) \rightarrow (\forall x)(B(x) \rightarrow |x| \leq n-1))$$

and W_0 does not believe

$$(\exists n)((\forall x)(B(x) \rightarrow |x| \leq n) \wedge \neg(\forall x)(B(x) \rightarrow |x| \leq n-1)) \qquad \text{(D)}$$

It will transpire that if W_i can see W_j whenever $i \leq j$ then W_0 not only disbelieves (D) but does not even believe the double negation

$$\neg\neg(\exists n)((\forall x)(B(x) \rightarrow |x| \leq n) \wedge \neg(\forall x)(B(x) \rightarrow |x| \leq n-1)) \qquad \text{(E)}$$

Case I: W_i can see W_j as long as $i \leq j$

We will be interested in whether or not W_0—or for that matter any of the other worlds—believes

$$(\forall n)((\forall x)(B(x) \rightarrow |x| \leq n) \rightarrow (\forall x)(B(x) \rightarrow |x| \leq n-1)) \qquad \text{(A)}$$

("if a bald man can have no more than n hairs on his head then a bald man can have no more than $n-1$ hairs on his head.")

For W_i to believe (A) it has to be the case that, for every $j \geq i$ and every n,

$$W_j \models (\forall x)(B(x) \rightarrow |x| \leq n) \rightarrow (\forall x)(B(x) \rightarrow |x| \leq n-1) \qquad \text{(B)}$$

and for (B) to hold it has to be the case that, for all $k \geq j$, and for each n,

$$\texttt{If } W_k \models (\forall x)(B(x) \rightarrow |x| \leq n) \texttt{ then } W_k \models (\forall x)(B(x) \rightarrow |x| \leq n-1) \qquad \text{(C)}$$

And (C) is clearly not true: it fails when $k = n$. This means that no W_i believes (A), and so $W_0 \models \neg$(A).

However, it is also pretty clear that W_0 does not believe

$$(\exists n)((\forall x)(B(x) \rightarrow |x| \leq n) \wedge \neg(\forall x)(B(x) \rightarrow |x| \leq n-1)) \qquad \text{(D)}$$

For W_0 to believe (D) there would have to be an actual n for which $W_0 \models (\forall x)(B(x) \rightarrow |x| \leq n) \wedge \neg(\forall x)(B(x) \rightarrow |x| \leq n-1)$, and there is clearly no such n.

With a bit of work we can check further that W_0 does not even believe

$$\neg\neg(\exists n)((\forall x)(B(x) \to |x| \leq n) \wedge \neg(\forall x)(B(x) \to |x| \leq n-1)) \tag{E}$$

All that is really going on here is that we have a model illustrating the standard fact that constructively $\neg(\forall x)(A(x) \to B(x))$ does not imply $\neg\neg(\exists x)(A(x) \wedge \neg B(x))$.

Case II: W_i cannot see W_j for $0 \neq i \neq j \neq 0$ but W_0 can see itself and all the others

All the W_i with $i > 0$ believe classical logic (since none of these worlds can see any world other than itself), and this simplifies matters considerably. (It also respects the intuition that none of these sharpenings should be privileged with respect to any of the others in any way.)

Let us review the reasoning in the previous section in the light of this modification.

For W_0 to believe (A) it has to be the case that, for every j and every n,

$$W_j \models (\forall x)(B(x) \to |x| \leq n) \to (\forall x)(B(x) \to |x| \leq n-1) \tag{B}$$

as before. This time W_j believes classical logic, which makes it easier than before to see that (B) fails, and it fails when $k = n$ as before.

It is also pretty clear—as before—that W_0 does not believe

$$(\exists n)((\forall x)(B(x) \to |x| \leq n) \wedge \neg(\forall x)(B(x) \to |x| \leq n-1)) \tag{D}$$

However this time W_0 *does* believe (E):

$$\neg\neg(\exists n)((\forall x)(B(x) \to |x| \leq n) \wedge \neg(\forall x)(B(x) \to |x| \leq n-1)) \tag{E}$$

$W_0 \models (E)$ because it cannot see a world that believes

$$\neg(\exists n)((\forall x)(B(x) \to |x| \leq n) \wedge \neg(\forall x)(B(x) \to |x| \leq n-1)) \tag{F}$$

This is because each W_i that W_0 can see, with $i > 0$, believes

$$(\forall x)(B(x) \to |x| \leq i) \wedge \neg(\forall x)(B(x) \to |x| \leq i-1)$$

and therefore believes

$$(\exists n)((\forall x)(B(x) \to |x| \leq n) \wedge \neg(\forall x)(B(x) \to |x| \leq n-1)) \tag{D}$$

Conclusion

I am not claiming that the neatness of this fit between supervaluationism and constructive logic means that if we are to understand vague concepts we need a metaphysics arising from a nonstandard logic. It might rather be that in order to represent how humans do in fact reason about vague concepts (and this is something one might wish to do when trying to simulate human reasoning in an IT system, for example) one might reach for a constructive logic, but the fit reported here need not be read as having any metaphysical significance.

References

[1] Anne McLaren: "Where to draw the line?" Proceeding of the Royal Institution **56** (1984) pp 101–121.