The sorites paradox

PHIL 20229
Jeff Speaks

April 22, 2008

1 Some examples of sorites-style arguments

The paradox we're discussing today and next time is not a single argument, but a family of arguments. Here are some examples of this sort of argument:

1. Someone who is 7 feet in height is tall.
2. If someone who is 7 feet in height is tall, then someone 6'11.9" in height is tall.
3. If someone who is 6'11.9" in height is tall, then someone 6'11.8" in height is tall.

   . . . . . .

C. Someone who is 3' in height is tall.

The '. . . ' stands for a long list of premises that we are not writing down; but the pattern makes it pretty clear what they would be. This is a paradox, since it looks like each of the premises is true, but the conclusion is clearly false. Nonetheless, the reasoning certainly appears to be valid.

We could also, rather than giving a long list of premises 'sum them up' with the following *sorites premise*:

> For any height h, if someone's height is h and he is tall, then someone whose height is $h - 0.1''$ is also tall.

(Though we'll see below that on some views, these two ways of running the argument are not equivalent.)

Once we see this, it is easy to come up with other instances of the paradox:

1. 10,000 grains of sand is a heap of sand.
2. 10,000 grains of sand is a heap of sand, then 9999 grains of sand is a heap of sand.
3. 9999 grains of sand is a heap of sand, then 9998 grains of sand is a heap of sand.

   ......

C. 1 grain of sand is a heap of sand.

Sorites premise: For any number n, if n grains of sand is a heap, then $n-1$ grains of sand is a heap.

1. A man with 1 hair on his head is bald.
2. If a man with 1 hair on his head is bald, a man with 2 hairs on his head is bald.
3. If a man with 2 hairs on his head is bald, a man with 3 hairs on his head is bald.

   ......

C. A man with 100,000 hairs on his head is bald.

Sorites premise: For any number n, if someone with n hairs on his head is bald, then someone with $n + 1$ hairs on their head is bald.

There's also a special case of this sort of argument that has to do with our powers of perceptual discrimination. Suppose that we line up 10,000 color swatches, which range from bright red (swatch 1) and the beginning to bright orange at the end (swatch 10,000). It seems as though, with 10,000 swatches, there will be no discernible difference between any two adjacent swatches, so that every swatch will look the same as the one next to it in the series. We can then construct the following argument:

1. Swatch 1 is red.
2. If swatch 1 is red, then swatch 2 is red.
3. If swatch 2 is red, then swatch 3 is red.

   ......

C. Swatch 10,000 is red.

If we assume that each swatch looks the same as the one next to it, then this version of the sorites argument can be thought of as having a rather special sorites premise: For any objects x, y, if x looks the same as y and x is red, then y is red. This certainly seems hard to deny. This is sometimes called the 'phenomenal sorites', and it raises interesting issues of its own.

Each of the above arguments uses a certain predicate — ‘tall’, ‘bald’, ‘heap’, ‘red.’ Not every predicate can be used in a sorites argument. Consider: ‘above average in height for someone in 105 Pasquerilla Center right now.’ (What would the sorites premise be for this predicate?)

One important feature of ‘tall’ and the other predicates which can be used to generate sorites-type paradoxes is that they all admit of borderline cases: each of these predicates is such that there are things to which we aren’t sure whether or not the predicate applies, no matter how much we know about the thing. Normally, to say that a predicate is vague is just to say that it has borderline cases, in this sense.

This feature — having borderline cases — should, as Sainsbury emphasizes be distinguished from two others: relativity to a reference class and ambiguity.

Can we generate a sorites paradox for any vague predicate?

Assuming that we cannot accept the conclusion of the instances of the sorites argument we discussed, any plausible response to the sorites paradox will fall into one of three categories:

- Rejecting the initial premise.
- Rejecting one of the other premises, and/or the sorites premise.
- Rejecting the validity of the argument.

We’ll discuss these in turn.

2 Rejecting the initial premise: nihilism

The simplest but also the most drastic response to sorites arguments is to reject their first premise. This involves, for example, denying that a 7-foot man is tall (and that he is tall for a man), that a man with one hair is bald, etc. Since we can set the values of the relevant quantities however we like in the first premise of a sorites argument, this fairly clearly involves denying that *anything* is bald, tall, etc. Yet more generally, this involves saying that no vague predicate — i.e., no predicate that admits of borderline cases — really applies to anything.

This looks pretty clearly like a last-ditch solution; before adopting such an extreme view, we should want to see if there are any better options.

3 Rejecting one or more of the other premises

3.1 The epistemic view

The first view is that there must be, and that the fault must lie with one of the premises in the sorites argument other than the first one. Consider, for illustration, the instance of the sorites argument concerned with baldness:

1. A man with 1 hair on his head is bald.
2. If a man with 1 hair on his head is bald, a man with 2 hairs on his head is bald.
3. If a man with 2 hairs on his head is bald, a man with 3 hairs on his head is bald.

   . . . . . .

C. A man with 100,000 hairs on his head is bald.

Suppose that we have to reject one of the premises other than the first. Which one should it be? Suppose that it is number 125:

125. If a man with 124 hairs on his head is bald, a man with 125 hairs on his head is bald.

If this is the faulty premise, then men with 125 hairs on their head who don't want to be bald should be extremely careful: if they lose just one more hair, that will push them over the edge into baldness.

The idea that the right solution to versions of the sorites paradox is to reject one of the premises — like number 125 — is called *the epistemic view*.

Many people find the epistemic view extremely hard to believe, for the following sorts of reasons:

- If one premise of this sort is false, then it is fair to say that no one knows which premise it is. Moreover, it is hard to see what we could do to find out what it is. So facts about whether our word 'bald' applies to someone with, say 130 hairs is forever unknowable. But is it plausible to think that there are unknowable facts of this sort about the application of our own words?
- Presumably, words like 'bald' have the meanings they do because of the way that we use them. But how could we use our words in ways which determined standards which even we not only don't know, but couldn't know?

3.2 Truth-value gaps

3.2.1 The simple truth-value gap theory

The most natural response to these questions, many think, is that the reason why we cannot know whether the sharp cut-off point for words like 'bald' falls is that there is nothing of that sort to know: there just is no sharp cut-off point.

The intuitive idea is this: there is one group of people of whom it is simply true to say that they are bald. There's another group of people of whom it is simply false to say that they are bald — of these people, it is true to say that they are not bald.

But there are also some people in the middle. If you say that one of them is bald, you haven't said anything true; but you haven't said anything false, either. Just so, if you say that one of them is *not* bald, you haven't said anything true, but you haven't said anything false, either. The rules for applying the word 'bald' just don't deliver a verdict for these people — it is 'undefined' when it comes to them.

When you think about the purposes for which we use the word 'bald' and other vague terms, this can seem quite plausible. We want to be able to use the word to be able to distinguish one group of people, the bald ones, and to say of some other people that they don't belong to that group. But it's not as though we have a big interest in providing an exhaustive division of the world's people into two groups, the bald and the non-bald.

What does this have to do with the sorites argument? One thing that the proponent of truth-value gaps can say is that some number of the premises in a typical sorites argument will fail to be true. Consider, for example, the one we considered earlier:

125. If a man with 124 hairs on his head is bald, a man with 125 hairs on his head is bald.

Let's suppose that it is neither true nor false to say that someone with 124 hairs, or 125, is bald. Then this premise is an example of a conditional whose antecedent and consequent are 'undefined.' The proponent of truth-value gaps might say that sentences of this sort are also undefined. Since these sentences are undefined rather than true, not all premises of the argument are true; perhaps this is enough to explain why the conclusion of the sorites argument is false.

But this approach to sorites arguments also has some curious features. Consider, for example, the sentence

> Either it is raining or it is not raining.

Ordinarily, we think of sentences of this sort as logical truths — they are true no matter what the weather. We think the same of

> If it is raining, then it is raining.

But on the present approach, it looks like these sentences can, in some cases, be false. For presumably there are borderline cases of rain; 'is raining' is a vague predicate. Suppose that, as often in South Bend, it is a borderline case of rain. Then both 'It is raining' and 'It is not raining' will be undefined, rather than true or false; but then it looks like neither of the above sentences will be true.

A different sort of problem comes from a seeming asymmetry between the following two sentences (suppose that Bob has 125 hairs on his head — or whatever number you think would make him a 'borderline case' of baldness):

> If Bob is bald, then with one less hair he would still be bald.
>
> If Bob is bald, then with one more hair he would still be bald.

It looks like, intuitively, the first one is definitely true; the second looks quite plausible (it's a premise that might be used in a sorites argument), but surely not as clearly true as the first. In any case, there appears to be a definite asymmetry between them. But what would the proponent of truth-value gaps say about these conditionals?

3.2.2 Supervaluationism

This might lead us to think that we want some approach to the sorites paradox which captures the idea that there are 'middle cases' for which words like 'bald' are undefined, but which avoids the problematic results discussed above. This is the aim of the proponent of *supervaluationist* approaches to vagueness.

The core idea behind supervaluationism is as follows: as above, there are a host of middle cases of thinly haired men which are such that the rules for 'bald' don't dictate that it is true to say of them that they are bald, but also don't dictate that it would be false to say this of them. So, in a certain sense, it is 'up to us' to say what we want about such cases. Let's call the act of 'drawing the line' between the bald and non-bald a *sharpening* of 'bald.' Then we can say that there are many possible sharpenings of 'bald' which are consistent with the rules governing the word.

Then the supervaluationist can give the following definitions:

> A sentence is true if and only if it is true with respect to every sharpening.
>
> A sentence is false is and only if it is false with respect to every sharpening.
>
> A sentence is undefined if and only if it is true with respect to some sharpenings, and false with respect to others.

Since the supervaluationist believes that some sentences are undefined rather than true or false, this is a version of the truth-value gap family of solutions to the sorites paradox.

To see why this view can seem plausible, consider the problems discussed above for the non-supervaluationist believer in truth-value gaps. First, consider the sentences

> Either it is raining or it is not raining.
>
> If it is raining, then it is raining.

According to the supervaluationist, these sentences are, as they seem to be, true in every circumstance. Can you see why?

Similarly, compare the pair of sentences

> If Bob is bald, then with one less hair he would still be bald.
>
> If Bob is bald, then with one more hair he would still be bald.

We said that there appears to be an asymmetry between these sentences. The supervaluationist can capture this asymmetry, since he takes the first to be true, but (given that Bob is a borderline case of baldness) the second to be undefined. This matches our intuitions about these sentences nicely.

How does this view escape the sorites paradox? Many premises in the typical instance of the sorites paradox will be true on some sharpenings, but false on others. So, as above, some of these premises will be undefined; this makes room for the view that the reasoning is valid and the conclusion false.

So far, so good for the supervaluationist. But this view too has some odd consequences:

- Above we counted it as a point in favor of the supervaluationist that

  > Either it is raining or it is not raining.

  always comes out true, even if neither of 'It is raining' and 'It is not raining' comes out true. But isn't this also a bit odd? How could a sentence of the form 'p or q' be true if neither p nor q is?

- Consider also the sentence:

  > There is some number such that if you have that number of hairs you are not bald, but if you have one fewer you are bald.

  The supervaluationist, like the epistemic theorist but unlike the simple truth-value gap theorist, must say that this sentence is true, since it is true on every sharpening. This seems counter-intuitive. A further weirdness is that the supervaluationist must say, for every number n in the borderline region, that

  > If you have n hairs you are not bald, but if you have one fewer you are bald.

  is undefined. But how could the first of these sentences be true, and the second false?

3.2.3 The problem of higher-order vagueness

Futher, all truth-value gap approaches — whether of the simple sort — of supervaluationist face the problem of *higher-order vagueness.*

This problem is that just as there are borderline cases between 'bald' and 'not bald', there are also borderline cases between cases where 'bald' applies and those borderline cases. But both simple truth-value gap theories and supervaluationist theories assume that there is a dividing line between the cases where 'bald' applies and the cases in which it is undefined.

(One way to see this is to define 'definitely bald' as a predicate that applies to everything to which 'bald' applies, and does not apply to everything to which 'not bald' applies and everything with respect to which 'bald' is undefined. The idea that there is higher-order vagueness can then be expressed as the idea that 'definitely bald' is itself vague.)

The epistemic theorist might then reply to the truth-value gap approaches by saying that if they posit a sharp dividing line between the cases to which 'bald' applies and the cases for which it is undefined, why not simplify the theory and posit such a sharp dividing line between 'bald' and 'not bald'? This would then avoid all of the problems discussed above. Of course, then we'd be stuck with the idea that there's some number of hairs such that, if you have that number you are not bald, but if you lost just one, you would be bald.

How should a proponent of truth-value gap approaches reply? What should she say about higher-order vagueness? The idea that the boundaries of vague predicates are context-sensitive.

Is there 3rd order vagueness as well? Does 'definitely definitely bald' have a different extension that 'definitely bald'?

4 Rejecting the validity of the argument: degrees of truth

There remains one other option: we might accept each of the premises of the argument, but deny that the conclusion of the argument follows. Initially, this does not look very plausible, since the argument seems only to employ the relatively uncontroversial inference from

> p
> If p, then q

to

> q

Of course, a typical sorites argument will contain very many instances of this sort of inference, but normally we would think that the combination of a bunch of valid inferences is a valid argument.

This is denied by partisans of the idea that in the standard case, sentences are not simply true or false, but rather true to a certain degree.

How can this help with an instance of the sorites paradox? The idea would be that a sentence like

> A 7' tall man is tall for an adult person.

is true to a very high degree – let's say, degree .95. However, a sentence like

A 6'11" tall man is tall for an adult person.

is true to a slightly lesser degree – say, degree .94. Now consider a typical premise in a sorites argument, like

If a 7' tall man is tall for an adult person, then a 6'11" tall man is tall for an adult person.

This is an 'if-then' statement in which the 'if' part is more true than the 'then' part. So the intuitive idea is that since we are going from a more true statement to a less true one, the whole 'if-then' statement is true to some degree less than 1. It is more true than most statements, surely — but not perfectly true.

On this view, what should we expect when we have a very long string of 'if-then' statements, none of which are perfectly true? How might this help explain how the reasoning in the sorites paradox can lead us from (almost complete) truth to (almost complete) falsity?

A concluding problem for the 'degree' approach: suppose that we have some statement 'p and q', in which both component statements are only true to a certain degree. What should we say about the degree of truth of the whole? Presumably, it should at least be determined by the degrees of truth of the components.

But now consider a statement like

It is raining.

that has degree of truth 0.5; presumably its negation,

It is not raining.

will also have degree of truth 0.5. This seems to lead to the conclusion that the following two conjunctions will have the same degree of truth:

It is raining and it is raining,

It is raining and it is not raining.

Could this be right?