An (apparently) vicious circle in logic

Can someone please help me with this following exercise 4.4 (p. 114) from the Mathematical Logic book of Ebbinghaus et al(this is not homework, but rather something that has been bugging me for a long time, which I now have found in form of an exercise, meaning there must be a definite answer to it):

> A reader who has been confused by the discussion in this chapter [I'm that reader!] says: "Now I'm completely mixed up. How can ZFC be used as a basis for first-order logic, while first-order logic was actually needed in order to build up ZFC?" Help such a reader out of his dilemma (*Hint*: Again be careful in distinguishing between the object and the background level.)

(logic)

asked Sep 24 '12 at 14:43
temo
1,581 13 38

1 Is it important to distinguish between meta-theory and theory?, How to avoid perceived circularity when defining a formal language? and some of things linked here in chat seem to be related. – Martin Sleziak Sep 24 '12 at 15:50

4 Answers

In any mathematical study there are two things that have to be clear: the "object" topic that we wish to study and the "background" or "meta" system that we use to study it. We can pick each of these with complete freedom. (Of course, if the background system is too weak, it won't be able to say anything about the object topic, but that's life).

For example, suppose we want to study real numbers as our object topic.

- We could use Euclidean geometry as our background system, so we would end up looking at numbers that can be constructed by ruler and compass. We would declare some segment S to have length 1 "by convention" and then other segments would represent real numbers depending on how the ratio of their length to the length of S. With more work we could even find ways to represent polynomials in terms of finite arrangements of points on the plane, so we could ask which polynomials can be constructed and which constructible polynomials have constructible roots.
- We could use ZFC, and define real numbers as some sort of sets, and study these particular sets.
- We could work more axiomatically by working in ZFC and studying the theory of a complete Archimedean ordered field. One difference between this and the previous bullet is that if we think of a real number just as an element of some field then it makes no sense to ask whether $5 \in \pi$, because there is no "$\in$" in the language of fields; but if we think of a real number as just some particular set then it does make sense to ask whether $5 \in \pi$. Similarly, in geometry we could ask whether two numbers (represented by line segments) are parallel.

The same sort of thing works for any other object theory. For example, if we want to study first-order logic, we could:

- We could work with a very weak theory that is only able to manipulate finite binary strings. By appropriate coding techniques we can view certain strings as representing formulas from first-order logic, in the same way that Euclidean geometry uses line segments to represent real numbers. We can then see what sort of constructions we can do on these strings.
- We could work in ZFC, develop a semantics for first-order logic, and study the relationship between provability and semantics.

The various ways of studying a single topic each act as "filters" that hide certain aspects of the topic and emphasize other aspects. At the same time, as was seen above, many of the meta systems add their own idiosyncracies as well. Most of the commonly used background systems can be arranged in a linear hierarchy of increasing strength (e.g. PRA < PA < ZF < ZFC), which Simpson has called the "Gödel hierarchy". Different portions of this hierarchy provide tools to study different aspects of the object topics.

There is no reason that we cannot use a certain system as a background system to study itself. For example, we can study ZFC within ZFC. This is no more paradoxical than using a wrench to fix the very same casting machine that made that wrench. Similarly, when we want to study ZFC, we may decide to use syntactic first-order logic as our base system, and when we want to study first-order logic, we may decide to take ZFC as our base system.

answered Sep 24 '12 at 19:16
Carl Mummert
56.1k 6 106 207

As always, your answers are the ones that I understand best, especially your description of the object-background language of other parts of mathematics, like real numbers. But this also raised some questions (concerning real numbers as set vs. real numbers as elements of an ordered field): Wouldn't it make sense to say $5 \in \pi$ if we treat numbers as elements of the c.a.o. field as well ? *If* we already are in ZFC, we already have "$\in$" available and the field is also a set (if I got this right) and we don't have a separate "language of fields", since [...] – temo Sep 25 '12 at 9:42

[...] all we have is the language in which we talk about ZFC ? Side question: How should I think about something that is lower than ZFC in the linear hierarchy, like PRA? Are these background theories "build"up in the same way as ZFC, in the sense that we start with logic (meaning that we, as humans, can do logical deductions) and some string manipulation (again meaning that we can discern symbols on paper and think of rule governing these strings) [...] – temo Sep 25 '12 at 9:49

[...] such that specific strings (i.e. the axioms of our background theory) we regard as meaningful and then just "play" with these strings, keeping in mind what they mean. Last question (MOST IMPORTANT): I find it a little bit circular, if you consider real numbers and then look a different background theories for interpreting them. That means that you already have to have the concept of a real number from somewhere, [...] – temo Sep 25 '12 at 10:32

[...] so it seems to me one has to already by in position of at least the idea of a real number. This makes me somehow feel uneasy, since what should we do, if the idea, about which we want to talk in the background, is so complicated, that we would need to formalise it ? (Of course this whole paragraph can be applied to more complicated tings that reals.) Then we would have to be in position of a formal description - but that already means that the background theory is in some sense fixed. What I am trying to express here is, that exchanging background theories seems problematic. [...] – temo Sep 25 '12 at 10:47

[... last one] Can all of what I wrote here be thought in terms of the "spiral" in Hurkyls answer to this (math.stackexchange.com/questions/173735/…) question ? Yes,I know, I shouldn't have asked this many follow-up question in the comments, but I would be extremely happy with even very short answers, since you explain things so damn well, that I would rather have you answer them quickly now that reposting them and get answer which don't tell my anything. – temo Sep 25 '12 at 10:54

|

As cody pointed out to me, this answers the question from the viewpoint of a mathematical platonist.

There are 2 different ZFCs. First is the ZFC that we work in, the one we know and love. In this ZFC, a "proof" means what we mean colloquially by a "proof". In this "casual": ZFC, helped by our real world intuition and experience, we know that "A and B" means what it does in english: both A and B are true. We know what "A(x) is true for all x" means - what it does in english. We know that 2+2=4, not because of number theory or the Peano axioms, but because that reflects our daily experience. We learn how to recognize proofs, not by their ability to be transformed into formal language, but by experience, intuition, and gut feeling.

Within this "casual" ZFC, we create the rigorous discipline known as logic. Within this logic, we formalize the notion of "truth", we formalize the notion of "proof", we formalize, well, everything. Most of our definitions which formalize these notions are based on our feelings/experience in "casual" ZFC. For example, in logic, we define $(A \wedge B)$ to hold if and only if A holds and B holds.

Finally, within logic, we define "formal" ZFC. Here, ZFC is a list of axioms and "proof" means "a chain of sentences, each of which is either a hypothesis, axiom, or can be formally deduced from a previous sentence in the chain by modus ponens." "Truth" becomes a semantic notion, often depending on which particular model of ZFC we are working with.

edited Sep 24 '12 at 15:45 answered Sep 24 '12 at 15:07

Jason DeVito
26.9k 4 64 114

3 So "what we mean colloquially by a 'proof'", which is "casual", is something that is *really* a proof, in the serious sense of the word, whereas the "formal" definition of "proof" as a suitable "chain of sentences", etc., is merely a mathematical model of the phenomenon, as opposed to the real thing. Therefore "casual" = "literal and serious". – Michael Hardy Sep 24 '12 at 15:19

@Michael: Yes, I think about it the same way. Perhaps you (or anyone else) could suggest a more appropriate word than "casual" (which balances well with "formal" but has connotations I'm not necessarily advocating, as you point out.) – Jason DeVito Sep 24 '12 at 15:23

How about "informal"? – Hurkyl Sep 24 '12 at 15:25

8 I would like to elaborate on Michael Hardy's comment: The "logic" that we build up from ZFC is not logic itself, but is a *mathematical model* of logic. As with any model, one can (and some people do) argue about the degree to which it accurately reflects the desired properties of the thing it purports to model. – MJD Sep 24 '12 at 15:32

@Hurkyl: Does "informal" not have the same connotations? I think of an "informal proof" as a hand waivy one which communicates the key ideas, but is not necessarily completely rigorous. – Jason DeVito Sep 24 '12 at 15:42

@Jason DeVito: I have two questions: 1) Could you shortly also describe what the formalist viewpoint would be (or does cody's answer reflect this viewpoint?) 2) I'm confused about the "casual ZFC", since it mixes two very different ingredients together: First the (unintuitive and far from our experience) axioms of ZFC and the (very intuitive, comprehensible by experience) basic notion of math like numbers. Since it seems to me, that the metatheory is for you the aspect of doing math, characterised by routine/intuition etc., so I don't think that mentioning ZFC alongside of it is good. – temo Sep 25 '12 at 5:31

@Jason DeVito In the first [link][1] given by Sleziak, in the last answer someone actually says that the *formalist*, not the platonist viewpoint is that one has to develop logic/ZFC twice. Now I'm also confused about what "formalist" and "platonist" means... [1]: mathoverflow.net/questions/40296/… – temo Sep 25 '12 at 5:33

@temo: 1). Honestly, I haven't really studied the platonist vs formalist debate too much. In particular, I'm just communicating my views, which often line up with the platonist. I don't feel comfortable speaking for the formalist. 2) The axiomatic ZFC I'm using in the "casual" sense is the intuitive idea of what a set is, together with the feelings that, for example, unioning 2 sets together should give a set, and likewise for the other axioms. Of course, some of the axioms are more "obvious" than others. – Jason DeVito Sep 25 '12 at 14:00

Formally, first-order logic is the metatheory in which most mathematics takes place. In this sense, first-order logic is prior to set theory in that when we prove something in set theory, we are using finitistic reasoning which is formalized rigorously in our metatheory. However, it is a phenomenon of sufficiently strong theories (ZF, Robinson Arithmetic) that they are able to formalize what it is to be a proof within themselves. Thus, you have a formula in ZFC ψ that takes Gödel numbers of sentences ϕ ($\ulcorner\phi\urcorner$) s.t $\psi(\ulcorner\phi\urcorner) = 1$ iff there is a proof in ZFC of ϕ. This phenomenon is the source of Gödel's incompleteness theorems. It is important to keep these two things separate, the first order logic formalized in the metatheory and the phenomenon that sufficiently strong theories are able to express fragments of this.

edited Sep 24 '12 at 15:40
MJD
43.3k 26 173 336

answered Sep 24 '12 at 15:11
cody
91 2

I did a spelling correction here: "a phenomena" ---> "a phenomenon". – Michael Hardy Sep 24 '12 at 15:20

It happened three times! Now I've fixed all three. – Michael Hardy Sep 24 '12 at 15:21

@Cody: As to your suggested edit on my post, if you're ok with it, I'll add a paragraph of my own saying that I am taking a platonist view point. Generally, edits which change the meaning or tone of another person's post are discouraged. Alternatively, you could post a comment about this below my answer once you get the rep to post comments. I think you can comment on your own answer, so you should be able to respond to this comment (I think). – Jason DeVito Sep 24 '12 at 15:26

My apologies, I am not yet familiar with the system. I think that, even if one is a platonist, it is important to understand that everything can be formalized in a finitistic metatheory. In particular, we never have to appeal to a 'casual' notion of proof, or an 'intuitive' definition of addition. I have even seen some people argue that this should be a motivation to platonism (which I disagree.) Thank you very much for your kind correction. – cody Sep 24 '12 at 15:34

I've now added a terse paragraph, mostly owing to the fact that I don't really know if I can explain myself clearly enough to add more. If you think I should say something more or different, just let me know. Incidentally, if you type "@Jason" at the beginning of your post, I get pinged on my end letting me know you are trying to talk with me. – Jason DeVito Sep 24 '12 at 15:47

The question what is actually meant by "ZFC is a basis for FOL" and "FOL is a basis for ZFC"?

The idea seems to be that all expressions can be reduced to a formal language ZFC/FOL. That was Hilbert's program. Turns out that this has some problems. And Gödel's incompleteness theory is not the only one in my view.

I think a much better analogy is a compiler which compiles source code into machine code. Machine code is a language which has to be realizable.

I believe this is a very deep question, and that the more or less arbitrary definition of level and meta-level is not the answer. The system itself would have incorporate a definition of levels. This is related to the question how to define definition.

answered Sep 24 '12 at 17:56
RParadox
308 4 9