Solutions: Formal theories

1. Consider a first-order language L for the theory of addition, whose logical apparatus comprises a suitable set of classical connectives, quantifiers, and identity. Its non-logical apparatus is to comprise the constants '0' and '1', the two-place function '+', and the two-place relation '<'.
   (a) Define the terms of L.
   (b) Show that it is algorithmically decidable which expressions are L-terms.
   (c) Define the atomic wffs of L.
   (d) Show that it is algorithmically decidable which expressions are atomic L-wffs.
   (e) Define the wffs of L [allowing wffs with free variables].
   (f) Show that it is algorithmically decidable which expressions are L-wffs.
   (g) Show that it is algorithmically decidable what is an L-wff which contains the variable 'x' free.
   (h) Show that it is algorithmically decidable which expressions are L-sentences (i.e. closed wffs without free variables).

We've intentionally been unspecific about the fine details of L. For example, which connectives does it have built in, and which are introduced (if at all) via definitions in terms of the basic ones? Are both the usual quantifiers built in? Does it typographically distinguish between variables occurring free (as 'parameters') and bound variables? Is the two place function expression '+' to be written infix (so we write e.g. '$(x + y)$') or prefix (so we write '$+(x, y)$' or even '$+xy$')?

You should, at this stage in your logical career, be familiar with the possibility of these variations in detail and also with the fact that, at bottom, they aren't going to make important differences. And you should be able to vary the sketched answers below to fit your own choices easily enough.

(a) The standard kind of definition, with the usual sloppiness about quotation, would be:

1. 0 and 1 are terms.
2. Any variable is a term.
3. If τ_1 and τ_2 are terms, so is $(\tau_1 + \tau_2)$.
4. Nothing else is a term.

If you typographically distinguish free variables/parameters from bound variables, then you will instead put ($2'$) any parameter is a term.

(b) Given an expression e, we can effectively test it to see whether it is 0, 1, or a variable (because you will have set things up so that it is effectively testable whether an expression is a variable in particular). A positive verdict settles the matter: e is a term.

Otherwise proceed to test whether e is of the form $(e_1 + e_2)$ where e_1 contains the

same (possibly zero) number of left and right brackets: that can evidently be done effectively by bracket counting. A negative verdict settles the matter: e is not a term.

Otherwise e is a term iff e_1 and e_2 both are. So start the test again, to determine in turn whether e_1 and e_2 are terms.

Eventually this process of breaking an expression into parts and testing parts will have to terminate, and we will get a final verdict on whether our initial expression is indeed a term.

(c) With the usual sloppiness about quotation,

1. If τ_1 and τ_2 are terms, then $\tau_1 = \tau_2$ is an atomic wff.
2. If τ_1 and τ_2 are terms, then $\tau_1 < \tau_2$ is an atomic wff.
3. Nothing else is an atomic wff.

(d) Since terms cannot contain either '=' or '<', we can simply ask of an expression e: is it of the form $e_1 = e_2$ or $e_1 < e_2$ for some e_1 and e_2? It is a mechanical business to test if so. If 'no', then e isn't an atomic wff. If 'yes', then e is an atomic wff iff both e_1 and e_2 are terms, and that's effectively testable given answer (b).

(e) You will want something like this:

1. If φ is an atomic wff it is a wff.
2. If φ is a wff, so is $\neg\varphi$.
3. If φ, ψ are wffs, so is $(\varphi \to \psi)$.
4. If φ is a wff and ξ is a variable, then $\forall\xi\varphi$ is a wff.
5. Nothing else is a wff.

You might want more binary connectives. If you set up terms and atomic wffs using a distinctive class of variables/parameters, you would say instead that if φ is a wff containing the parameter α and ξ is a variable (new to φ), then $\forall\xi\varphi'$ is a wff (where φ' results from φ by substituting ξ for some or all occurrences of α).

(f) Again you need to consider cases.

1. Given an expression e, test to see if it is an atomic wff, which we can do effectively by part (d). If it is, we are done: e is a wff.
2. If e it isn't atomic, test it to see if it is of the form $\neg e_1$ or $\forall\xi e_1$, which can again be done effectively. If it is, e is a wff iff e_1 is, and we need to start again from the beginning, running this test on e_1.
3. Otherwise, test to see whether e is of the form $(e_1 \to e_2)$ where e_1 has the same number (perhaps zero) of left and right parentheses (that can be done effectively). If it isn't e isn't a wff. Otherwise, e is a wff iff e_1 and e_2 are, and we need to start again, testing e_1 and e_2 in turn.
4. Eventually this process must terminate ...

(g) One reason for setting things up in the first place with free variables (i.e. parameters) typographically distinguished from bound variables is that doing so makes these two last two questions trivially easy. In particular, we just look to see if a wff contains any of those typographically marked variables to determine whether it is a closed wff or not.

If you use only one style of variable things are inevitably messier. Essentially, you need to determine, of a given occurrence of a variable ξ in a wff φ, whether is buried in a wff the form $\forall\xi\psi$ where this wff is part of φ. If it isn't, then that occurrence of ξ is free in φ.

2. Consider the following (uninterpreted) theory H. The alphabet of H's language consists of the symbols $\mathsf{M}, \mathsf{U}, \mathsf{I}$, and any finite string of symbols is a wff. The theory has one axiom: MI. H also has five rules of inference (σ indicates a string of symbols, possibly empty).

   1. Given a wff of the form $\sigma\mathsf{I}$, you can infer the wff $\sigma\mathsf{IU}$. (For example, from MUI infer MUIU.)
   2. Given a wff of the form $\mathsf{M}\sigma$, infer $\mathsf{M}\sigma\sigma$. (For example, from MIIU infer $\mathsf{MIIUIIU}$.)
   3. Given a wff which includes the string UI, infer the wff that results from replacing that string with IU. (For example, from MIUIU infer MIIUU.)
   4. Given a wff which includes the string UU, infer the wff that results from deleting that string. (For example, from MIIUU infer MII.)
   5. Given a wff which includes the string III, infer the wff that results from replacing that string with a U. (For example, from $\mathsf{MUIIIIU}$ infer MUIUU.)

   (a) Does H count as an effectively axiomatized (uninterpreted) formal theory?

   (b) Prove every H-theorem starts with the symbol M, and contains no other occurrence of M.

This question is based on a famous puzzle due to Douglas Hofstadter in his book *Gödel, Escher, Bach*.

(a) We've said how to effectively determine what's an H-wff, and it is effectively decidable what is an instance on an H-rule. Strictly speaking, though, we haven't yet set down what counts as an H-proof. It would be consistent with what we have so far to add, e.g., that a proof is a string of wffs with no more than one appeal to any particular rule of inference. But it is obvious what is intended: a proof is a sequence of wffs, starting from the axiom, where each step applies a rule of inference to the previously derived wff. With that understood, H is an effectively formalized theory.

(b) By inspection, the sole axiom starts with an M. We then just note that no rule of inference deletes an initial M or introduces another one. So, however long a proof goes on, we can only derive wffs of the form $\mathsf{M}\sigma$ where there are no occurrences of M in σ.

(c) Can you derive MIIU as a theorem?

(d) Can you derive MUUIU?

(e) Can you derive MU?

(c) Here's a proof: $\mathsf{MI}, \mathsf{MIU}, \mathsf{MIUIU}, \mathsf{MIIUU}, \mathsf{MII}, \mathsf{MIIU}$.

(d) Here's a proof: $\mathsf{MI}, \mathsf{MIU}, \mathsf{MIUIU}, \mathsf{MIUIUIUIU}, \mathsf{MIIUUIUIU}, \mathsf{MIIIUIU}, \mathsf{MUUIU}$.

(e) This can't be done: see (h).

(f) Show that, for each rule, if it is applied to a wff whose number of contained 'I's is not a multiple of 3, the result is a wff whose number of 'I's is also not a multiple of 3.

(g) Can you derive MIUIIUIII?

(h) Now revisit question (h) again.

(f) Proof is by inspection. Rules (1), (3), (4) don't change the number of 'I's between premiss and conclusion. Rule (2) will double the number of 'I's (so if the number of 'I's in the premiss is not divisible by three, then likewise for the number of 'I's in the conclusion). Rule (5) reduces the number of 'I's by three (so again, if the number of 'I's in the premiss is not divisible by three, then likewise for the number of 'I's in the conclusion).

(g) Since number of 'I's in the sole axiom is not divisible by three, and all the rules of inference take us from wffs where the number of 'I's is not divisible by three to another such wff, no theorem can contain $3n$ 'I's for any n. In particular, it can't contain 6 'I's, which rules out MIUIIUIII as a theorem.

(h) Nor can it contain zero 'I's, which rules out MU as a theorem.

What's interesting about this example is that a simple, purely syntactic argument, involving symbol-counting quickly establishes the non-derivability of certain wffs. That contrasts with the non-derivability proof we give in answering the next question.

3. Mathematicians will recognise this question as being about the theory of *groups*. As part of philosophers' general education, they ought to know about the concept of a group too, for a reason that will emerge.

   Consider the formal first-order theory G whose non-logical vocabulary comprises just a two-place function expression '$\cdot$' and a constant 'e'. We'll in fact write the function 'infix' like ordinary multiplication, so we put e.g. '$(x \cdot y)$' rather than $\cdot(x, y)$. We will also allow the dropping of outer brackets. The axioms of G are:

   1. $\forall x \forall y\, x \cdot (y \cdot z) = (x \cdot y) \cdot z$
   2. $\forall x\, x \cdot e = x$
   3. $\forall x \exists y\, x \cdot y = e$

   (a) Is G an effectively formalized theory?

   (b) Prove $\forall x \exists y\, y \cdot x = e$.

   (c) Prove $\forall x\, e \cdot x = x$.

   (d) Prove that the 'unit' e is unique: in other words, if e and e' both satisfy (2) and (3), then $e = e'$.

   (e) Prove that 'inverses' are unique: i.e., $\forall x \forall y \forall y'((x \cdot y = e \wedge x \cdot y' = e) \rightarrow y = y')$.

(a) Assuming the background first-order logic is correctly formalized, and we set down rules for using brackets with the sole function expression (and for dropping them), then yes, G is an effectively formalized theory.

(b) We argue informally, but you can dress this up inside whatever formal deductive system you have given G. Take an arbitrary x. Then for some y, we have $x \cdot y = e$

(by 3). And for some z, we have $y \cdot z = e$ (by 3 again). Then we have

$$\begin{aligned} y \cdot x &= (y \cdot x) \cdot e \\ &= (y \cdot x) \cdot (y \cdot z) \\ &= y \cdot (x \cdot (y \cdot z)) \\ &= y \cdot ((x \cdot y) \cdot z) \\ &= y \cdot (e \cdot z) \\ &= (y \cdot e) \cdot z \\ &= y \cdot z \\ &= e \end{aligned}$$

Which shows, as we wanted, that for any x, there is a y such that $y \cdot x = e$. And in fact, any y that serves as a 'right inverse' to x is also a 'left inverse'.

(c) Take an arbitrary x. We know that there is a y such that $x \cdot y = y \cdot x = e$. So then we have

$$\begin{aligned} e \cdot x &= (x \cdot y) \cdot x \\ &= x \cdot (y \cdot x) \\ &= x \cdot e \\ &= x \end{aligned}$$

(d) Suppose e and e' are units satisfying (2) and (3).

$$\begin{aligned} e &= e \cdot e' && \text{since } e' \text{ obeys (2)} \\ &= e' && \text{applying result (c) above} \end{aligned}$$

(e) Given an arbitrary x, suppose for some y and y', $x \cdot y = e$ and $x \cdot y' = e$. We can then argue as follows:

$$\begin{aligned} y &= y \cdot e && \text{from (2)} \\ &= y \cdot (x \cdot y') && \text{by assumption} \\ &= (y \cdot x) \cdot y' \\ &= e \cdot y' && \text{by result (b)} \\ &= y' && \text{by result (c)} \end{aligned}$$

(f) Prove G is consistent by giving three interestingly different interpretations for the language of G on which G's axioms are all true.

(g) Is $\forall x \forall y\ x \cdot y = y \cdot x$ a G-theorem?

(h) Is G negation complete?

(f) A structure which interprets the axioms G is said to be a *group*. Here's just three examples.

1. Take the domain of interpretation to be rational numbers greater than zero. Take '$\cdot$' to mean multiplication, and the 'unit' e to be 1. Then the axioms (1) to (3) are all true. [A similar infinite example: take the domain to be all integers, positive and negative. Take '$\cdot$' to mean addition, and the 'unit' e to be 0.]
2. A small finite example. Take the domain to be ways of moving a square so the same four points continue to be coincide with a corner each. (You can leave the square where it is, or rotate by 90°, 180°, or 270°, or flip over on a horizontal, vertical, or diagonal axis). Take $x \cdot y$ to mean that operation x is done and then

operation y, and take e to signify the operation of leaving the square where it is. Axioms (1) and (2) are trivially satisfied. (3) is also satisfied since any rotation can undone by continuing through 360°, and any flip when repeated undoes its effect.

3. Another finite example. Suppose we have a pack of cards. Take the domain to be 'shuffles' or 'permutations' of the cards (where we move the card in position 1 to the position p_1, the card in position 2 to the position p_2, etc. etc.). Again take $x \cdot y$ to mean the result of doing shuffle x followed by shuffle y, and e denotes leaving the pack undisturbed.

We could continue: see http://en.wikipedia.org/wiki/Examples_of_groups.

The point to note here is that there is no single 'intended interpretation' of theory G (in the way that there is, as we shall see, an intended interpretation of various theories of arithmetic we shall be looking at). There's an abstract generality here which is characteristic of much modern mathematics (we are interested here not in this or that particular structure, but about a *kind* of structures, i.e. structures with respect to which the group axioms of G hold.)

(g) Take the example of shuffles of a pack of five cards, and consider these two shuffles: a reverses the order of the first two cards; b reverses the order of the whole pack. Then applied to the cards $[1, 2, 3, 4, 5]$, applying a then b gives $[5, 4, 3, 1, 2]$, applying b then a gives $[4, 5, 3, 2, 1]$. So $a \cdot b \neq b \cdot a$. Hence $\forall x \forall y\ x \cdot y = y \cdot x$ can't be a G-theorem.

Note then that here we are proving the non-derivability of a wff by going via semantic considerations. There is an interpretation which makes the axioms true and the target wff false, so the axioms don't semantically entail the target wff, hence – by the soundness of first-order logic – the axioms don't deductively entail the target wff either.

(h) Since $\forall x \forall y\ x \cdot y = y \cdot x$ is true on some interpretations of G (e.g. the numerical ones we gave first) and false on others, that shows that neither it nor its negation is derivable from the axioms. So G is not negation-complete.

4. A reality check. Suppose T is an effectively axiomatized formal theory. Can T be

(a) Inconsistent and negation-complete?

(b) Consistent, negation-incomplete and decidable?

(c) Inconsistent and undecidable?

(d) Consistent and undecidable?

(e) Consistent, negation-complete and undecidable?

(a) Yes. If T has a classical logic, if it is inconsistent, it will prove both φ and $\neg\varphi$ for any sentence φ, so is automatically negation-complete.

(b) Yes. The null theory with no axioms is consistent, negation-incomplete and decidable. Or take the theory whose language is a propositional language with atomic wffs p and q, with a standard classical logic, and where p is the only axiom. This is consistent, negation-incomplete and decidable (why?).

(c) No, at least if T has a classical logic. For then T it proves everything, so it is trivially decidable whether a wff is a theorem.

(d) Yes. We've already announced that this is going to happen with theories of arithmetic!

(e) No. See Theorem 4.2.