Solutions: Sufficiently expressive/strong

1. Suppose T is an effectively axiomatized sound theory. Which of the following questions are you currently placed to settle?
   (a) Suppose G is a sentence of T's language which is true iff G is not provable in T: can T decide G?
   (b) Suppose H is a sentence of T's language which is true iff H is provable in T: can T decide H?
   (c) (Looking ahead, but try thinking about it!) Suppose M is a sentence of T's language which is true iff M is not provable in T in less than a million inference steps: can T decide M?

(a) We met this back in Chapter 1! Suppose G is false. Then G is provable in T, i.e. $T \vdash G$. But that contradicts T's soundness. So G is true, i.e. $T \nvdash G$. And since G is true, $\neg G$ is false, so by T's soundness again $T \nvdash \neg G$. So T can't decide G.

(b) (i) No obvious problem arises from the assumption $T \vdash H$. (ii) Suppose alternatively $T \vdash \neg H$. Then, since T is sound and so consistent, $T \nvdash H$. So H would be false, and $\neg H$ true. But then that's consistent with assumption (ii). So far, then, it seems that T could decide H one way or the other, or do neither. [We return though to related questions in §34.5.]

(c) Suppose M is false. Then M *is* provable in T in less than a million steps, so we have $T \vdash M$. But that contradicts T's soundness. So M has to be true: T cannot prove M in less than a million steps. And since M is true, and T is sound, T can't prove $\neg M$.

But the possibility remains that T *can* prove M (albeit in more than a million steps). Can it? Well, *you* can prove M in the following mechanical way: just run through 'in alphabetical order' all the T-proofs with less than a million steps, examining them in turn [this is going to take a lot more than a million steps!]. We know this must verify that none of these short proofs is a proof of M, since M is true, and hence this brute-force proof establishes M. Now suppose that theory T is rich enough to encode claims about this kind of brute-force mechanical process (and you don't know this yet, but T will be able to do this if it contains enough arithmetic) then T will similarly be able to prove M, but in more than a million steps.

2. In this exercise, take 'theory' to mean any set of sentences equipped with deductive rules, whether or not effectively axiomatizable:
   (a) If a theory is effectively decidable, must it be negation complete?
   (b) If a theory is effectively decidable, must it be effectively axiomatizable?
   (c) If a theory is negation complete, must it be effectively decidable?
   (d) Say a first-order theory Q is finitely axiomatizable iff there is a finite set of axioms which together entail every Q-theorem. Much such a theory Q be effectively axiomatizable?

(e) First-order logic is compact: so if $\Gamma \vdash \varphi$ then $\Gamma^* \vdash \varphi$, where Γ^* is a finite subset of Γ. Must an effectively axiomatizable first-order theory therefore be finitely axiomatizable?

(a) No. See *IGT2*. p. 32.

(b) Yes. Suppose T is effectively decidable. That means the set of T-theorems T' is effectively decidable. So consider the theory whose axioms are all of T', and with the trivial deductive system which allows us to infer φ from itself. Then this theory is an effectively axiomatized theory, and its theorems are all the T-theorems.

(c) No. We've only proved that if a theory is effectively axiomatizable and negation complete then it is effectively decidable. We will in due course find negation complete theories which are not effectively axiomatizable and are not decidable. (E.g. the theory whose axioms are all the truths expressible in the language of basic arithmetic.)

(d) Trivially yes. A finite axiomatisation of a theory is a decidable axiomatisation (we can just check whether a wff is an axiom by a finite check against the finite list of axioms).

(e) The reference to compactness is a bit of indirection. As we will later find, we can have effectively axiomatized first order theories which aren't finitely axiomatizable. The proof of any particular theorem uses only a finite number of axioms of course, but there need be no finite set of axioms which proves all theorems. Peano Arithmetic is an example.

3. (a) What is it for a logical theory (a deductive proof system) to be effectively decidable?

(b) Is your favourite proof system for classical propositional logic effectively decidable?

(c) Suppose Q is a finitely axiomatizable theory with a standard first-order logic; then show that there is a single sentence $\hat{Q}$ such that $Q \vdash \varphi$ if and only if $\vdash \hat{Q} \to \varphi$ (where $\vdash$ is deducibility in your favourite first-order logic).

(d) Prove that if there is a consistent, finitely axiomatizable, sufficiently strong theory with a first-order logic, then first-order logic is undecidable.

(a) We said (*IGT2*, §4.4) that a theory T is effectively decidable iff the property of being a theorem of T is effectively decidable. So naturally, we say a logical system is decidable if it is effectively decidable whether a given wff is a theorem or not.

(b) Your favourite deductive proof system of classical propositional logic is (let's hope!) sound and complete with respect to the standard semantics, so that φ is a theorem of the logic (deducible from no assumptions) if and only if it is a tautology. But you can effectively decide whether φ is a tautology by running a truth-table test. So the truth-table test also effectively decides whether φ is a theorem of your favourite proof-system.

(c) This is another trivial exercise, here just to give you a hint about how to prove the theorem in the next part of the question. Let $\hat{Q}$ be the conjunction of the axioms of some finite axiomatisation of Q. Then, trivially, $Q \vdash \varphi$ if and only if $\hat{Q} \vdash \varphi$ if and only if $\vdash \hat{Q} \to \varphi$. (And indeed, this will apply to most sensible deductive relations, not just first-order logic.)

(d) Suppose Q is a consistent finitely axiomatized theory with a first-order logic and which is sufficiently strong. Since it is finitely axiomatized, we can wrap all its axioms together into one long conjunction, $\hat{Q}$. And then, as just noted, $Q \vdash \varphi$ if and only if $\vdash \hat{Q} \to \varphi$; i.e. we can prove φ inside Q if and only if a certain related conditional is logically provable from no assumptions.

So if (1) the logic were decidable, then (2) we could mechanically tell whether the conditional $\hat{Q} \to \varphi$ is a logical theorem, hence (3) we could mechanically decide whether φ is a Q-theorem. But since Q is a consistent sufficiently strong effectively (because finitely) axiomatized theory, (3) is impossible [that's *IGT2*, Theorem 7.2]. So (1) is impossible – the logic must be undecidable.

[Later, we will be able to use this line of proof to show that first-order logic is indeed undecidable, because we will find a finitely axiomatized sufficiently strong first-order arithmetic.]

4. Let *True* be the set of all true sentences of a sufficiently expressive language L with classical negation. We can treat *True* as a theory (with just the trivial rule of inference 'from φ infer φ).
   (a) Show *True* is consistent.
   (b) Show *True* is negation complete.
   (c) Show *True* is sufficiently strong.
   (d) Use Theorem 7.2 to conclude that the set of sentences *True* is not effectively axiomatizable by any theory framed in language L.

(a) Trivial.

(b) *True* is negation complete, for any sentence L-sentence φ, the true one of φ and $\neg\varphi$ is in *True*, so – by the trivial rule – either *True* $\vdash \varphi$ or *True* $\vdash \neg\varphi$.

(c) We show *True* captures any decidable property of numbers, P. By hypothesis, if P is a decidable property, L can express it using some $\varphi(\mathsf{x})$ (since L is sufficiently expressive). So if n has P, $\varphi(\overline{\mathsf{n}})$ is true. But then $\varphi(\overline{\mathsf{n}})$ will be a sentence in *True*, so *True* $\vdash \varphi(\overline{\mathsf{n}})$. While if n does not have P, $\neg\varphi(\overline{\mathsf{n}})$ is true. But then $\neg\varphi(\overline{\mathsf{n}})$ will be a sentence in *True*, so *True* $\vdash \neg\varphi(\overline{\mathsf{n}})$. So depending on whether n has P, either *True* $\vdash \varphi(\overline{\mathsf{n}})$ or *True* $\vdash \neg\varphi(\overline{\mathsf{n}})$. So T captures P.

(d) Suppose *True* were effectively axiomatizable, there would be a consistent effectively axiomatized theory T whose theorems are just the theorems of *True*. So T would be effectively axiomatized, and (like *True*) negation complete and sufficiently strong. But that's impossible by Theorem 7.2.

5. Suppose T is an effectively axiomatized, consistent, sufficiently strong theory. And suppose we augment the language of T and add new axioms to get a new consistent, effectively axiomatized, theory U. Now let U^* be all the theorems of U which are expressed in T's original, unaugmented, language.
   (a) Show U^* is consistent, and sufficiently strong.
   (b) Show that if U^* is negation complete then it is decidable.
   (c) Show that U^* therefore cannot be negation complete.

(a) It is trivial that U^* is consistent if U is, and since U^* must contain every T-theorem it is trivial that it is sufficiently strong (regarded as a theory).

(b) We use the same kind of argument used in proving Theorem 4.2. To spell that out: given any sentence φ in the language of U^*, i.e. the language of T, just mechanically grind out theorems of U until you get either φ or its $\neg\varphi$ (you must get one of them, since by supposition U^* is negation complete).

(c) If U^* is negation complete it is decidable, and so by Exercise 2(b) effectively axiomatizable. But U^* would then be consistent, sufficiently strong, effectively axiomatizable and negation complete, which is impossible by Theorem 7.2. So U^* can't be negation complete.

That last result is interesting. We know that if T is an effectively axiomatized sufficiently strong consistent theory in the language L, then it can't be complete. But we might have wondered if however we could have expanded T to a stronger effectively axiomatized theory (richer langer, more axioms) which – while still negation incomplete overall – at least is negation complete for sentences in the original language L. We now know this can't be done. The richer theory U, if effectively axiomatized, will still be negation complete for sentences in T's language.