Löb's theorem

Canonical name	LobsTheorem
Date of creation	2013-03-22 17:05:19
Last modified on	2013-03-22 17:05:19
Owner	dankomed (17058)
Last modified by	dankomed (17058)
Numerical id	25
Author	dankomed (17058)
Entry type	Theorem
Classification	msc 03F03
Classification	msc 03F45
Classification	msc 03F40

In this short article we will provide an example where one's intuition can be faulty, and that strict formalization is always needed when one does mathematics in order to avoid possible hidden intuitive inconsistencies. A striking example is the erroneous but intuitively obvious thesis: "if a formal system F proves that there exists proof of formula A in F, then F has proved A" or slightly differently version is this one: "it must be always provable in all formal systems that from formula saying that there exists proof of A in F, indeed follows A in F". Both formulations of this naive intuition concerning basics of math logic severely fail. We will show that a consistent F cannot prove for arbitrary formula ϕ_n that from formula that asserts "exists proof of ϕ_n" follows ϕ_n. This is a major ingredient of a theorem proved by Löb in 1955.

> Löb's theorem: If $F \vdash \exists x P(x, \check{n}) \Rightarrow \phi_n$, where x is the Gödel number of the proof of the formula with Gödel number n, and $\check{n}$ is the numeral of the Gödel number of the formula ϕ_n, then $F \vdash \phi_n$.

The proof presented here follows Karlis Podnieks (2006):

Suppose $F \vdash \exists x P(x, \check{n}) \Rightarrow \phi_n$. Let F^* be the new proof system that is equal to F plus the new axiom $\neg\phi_n$. By modus tollens $F^* \vdash \neg\phi_n \Rightarrow \neg\exists x P(x, \check{n})$. Since F^* has as axiom $\neg\phi_n$ by modus ponens $F^* \vdash \neg\exists x P(x, \check{n})$. In this case F^* proves that ϕ_n is not provable in F (and also that F is consistent). This however leads to proof that F^* is also consistent because it contains only F and $\neg\phi_n$, and we already know both that F is consistent and that F does not prove ϕ_n. Thus F^* proves its own consistency. According to Gödel's second theorem however F^* cannot prove its own consistency unless being inconsistent. Therefore $F + \neg\phi_n$ must be always inconsistent theory, and we conclude that $F \vdash \phi_n$.

Now see that the intuitive notion that $\exists x P(x, \check{n}) \Rightarrow \phi_n$ is not provable within arbitrary formal system. If the formal system is inconsistent obviously it proves every formula so the above intuitive notion will be provable. If however F is consistent it cannot prove $\exists x P(x, \check{n}) \Rightarrow \phi_n$ for arbitrary formula, because *if it could*, for formally refutable formula $\phi_k \equiv [0 = 1]$, F *could* prove by modus tollens $\neg\phi_k \Rightarrow \neg\exists x P(x, \check{k})$, and since $\neg\phi_k$ is formally provable formula, then the formal system *could* infer its own consistency, or said in ordinary language: from $\neg\phi_k$, F *could* have proved "there exists at least one unprovable formula in F", which is impossible according to Gödel's second theorem.

Corollary: If F is consistent formal system then F cannot prove for arbitrary formula ϕ_n that from formula that asserts "exists proof of ϕ_n" follows ϕ_n.

The proof of the corollary is straigthforward either directly from Löb's theorem, or by independent reasoning using the fact that consistent F cannot prove $\exists x P(x, \check{n}) \Rightarrow \phi_k$ for refutable ϕ_k.

The relationship between the direct text of Löb's theorem in the proof provided by Podnieks, and the corrolary is this one: if the system F is consistent, it might be able to prove $\exists x P(x, \check{n}) \Rightarrow \phi_n$ for provable formulas ϕ_n, because the provability of ϕ_n taken together with the consistency of F ensure unprovability of $\neg\phi_n$ and hence modus tollens reversed formula of $\exists x P(x, \check{n}) \Rightarrow \phi_n$ cannot be used for F to prove its own consistency. Also taking into account the second Gödel theorem it is easy to be seen that F should not be able to prove $\exists x P(x, \check{n}) \Rightarrow \phi_k$ for disprovable (refutable) formulas ϕ_k, a result that alone is sufficient to prove the corollary. Still the corollary is a weaker result than Löb's theorem, because the corrolary does not give us clue whether $\exists x P(x, \check{n}) \Rightarrow \phi_p$ is provable for undecidable formulas ϕ_p. This was indeed the "open problem" proposed by Leon . The Löb's theorem answers this question, and shows that it is impossible for consistent F to prove $\exists x P(x, \check{n}) \Rightarrow \phi_p$ for undecidable formulas ϕ_p. Therefore summarized, Löb's theorem says that for refutable or undecidable formulas ϕ, the intuition "if "exists proof of ϕ" then ϕ" is erroneous.

Modus tollens inversed Löb's theorem: If F is consistent formal system then F cannot prove for any unprovable formula ϕ_k that from formula that asserts "exists proof of ϕ_k" follows ϕ_k.

Comparing the above theorem with the corrolary shows small but significant difference, which makes it stronger proposition (note: in consistent formal system F unprovable formulas are all refutable (disprovable) formulas, as well as all undecidable formulas).

References

1. Crossley JN, Brickhill C, Ash C, Stillwell J, Williams N (1972) *What is mathematical logic?* Oxford University Press.

2. Detlovs V, Podnieks K (2006) `http://www.ltn.lv/ podnieks/mlog/ml.htm`*Introduction to Mathematical Logic.* Hypertextbook for students in mathematical logic

3. Gödel K (1931) Über formal unentscheidbare Sätze der Principia und verwandter Systeme I. *Monatshefte für Mathematik und Physik* **38**: 173-198.

4. Löb MH (1955) Solution of a Problem of Leon Henkin. *The Journal of Symbolic Logic* **20**(2): 115-118.

5. Podnieks K (2006) `http://www.ltn.lv/ podnieks/gt.html` *What is Mathematics? Gödel's Theorem and Around*. Hypertextbook for students in mathematical logic