T. Parent
Phil 5506: Metalogic

GÖDEL'S PROOF OF INCOMPLETENESS

We assume that System H is a consistent formal theory of arithmetic, and we already showed that it strongly represents all recursive functions. (The Expressibility Lemma for H.) Below, we will see it also has a decidable set of wff and a decidable set of proofs. Thus, GG secures that H is not negation-complete. (Gödel's First Incompleteness Theorem for H.)

However, Gödel's own proof of incompleteness is a bit different. It also relies crucially on EL, the consistency of the formal system, and the decidability of wff/proofs. But instead of going via GG, it instead unearths a particular sentence such that neither it nor its negation is a theorem.

Preliminary: Gödel Numbering

We need to make decidable the set of wff of H, and the set of its proofs. To this end, we assign numbers to symbols and sequences of symbols, but we need our coding to be more nuanced than before. It is one thing for '*p*' to be a wff, and another thing for '*p*' to be a one-line proof within the system (e.g., in a system where '*p*' is axiomatic). So '*p*' as a wff requires a different code than '*p*' as a proof.

Gödel invented a numbering system which enables such distinctions. We begin by assigning the first odd numbers >1 to the basic symbols of the language:

Symbol	*Number*	*Symbol*	*Number*
p	3	*	15
′	5	~	17
x	7	⊃	19
a	9	∧	21
f	11	(	23
F	13	)	25

To determine the Gödel number for a wff, we take the Gödel numbers of its *n* basic symbols, and assign them (in order) as exponents to the first *n* members of the sequence of primes ⟨2, 3, 5, 7, …⟩. The Gödel number for the wff is then the product of the exponentiated primes. Thus, the sentence '*F′a′*' is assigned $2^{13} \cdot 3^5 \cdot 5^9 \cdot 7^5$, whereas the sentence '*p*' is assigned $2^3 = 8$. And in general, a wff will always have an even Gödel number, unlike a basic symbol.

Consider now a proof consisting of the wff $A_1, A_2 \ldots A_n$. To determine its Gödel number, take the Gödel numbers for those wff, and assign them (in order) as the exponents to the first *n* members of the sequence of primes. The number for the proof is then the product of *those* exponentiated primes. E.g., the proof consisting just of '*p*' will have a Gödel number of $2^8 = 16$. And the proof consisting of '*p*' followed by '$p \supset p$' will have the Gödel number $2^8 \cdot 3^k$, where *k* is the Gödel number for '$p \supset p$'. Note that the exponent on 2 will be even iff the Gödel number being calculated is a number for a *proof*.

It should be clear, then, that each basic symbol, wff, and proof is assigned a distinct Gödel number. And we can tell from the Gödel number whether it is the number for a basic symbol, wff, or a proof. More than that, we can recover *which* symbol/wff/proof it is a number for. This is because each Gödel number is a product of primes, and because of the following:

> Fundamental Theorem of Arithmetic: Any integer >1 is the product of a *unique* prime factorization.

Thus, if we recover which prime factorization yields a Gödel number as its general product, we can thereby determine which symbol/wff/proof it is the Gödel number of. The upshot is that the wff/proofs of the system are *decidable*.

The Arithmetization of the Proof Relation

Given the decidability of wff and proofs, the (metamathematical) relation "P is a proof in H of A" is also decidable. So assuming Church's Thesis, the following function *f* is recursive. Where P is a proof with Gödel number $^{\#}$P, and A is a wff with Gödel number $^{\#}$A,

$$f(^{\#}\mathrm{P}) = \begin{cases} ^{\#}\mathrm{A} & \text{if P is a proof in H of A.} \\ 0 & \text{otherwise} \end{cases}$$

Thus by EL, we know that *f* is strongly represented in the system. That is to say, there is a formula '$\mathrm{Pf}_{\mathrm{H}}(n, m)$' [intuitively, the proof predicate for H] such that:

(i) $\vdash_{\mathrm{H}} \ulcorner \mathrm{Pf}_{\mathrm{H}}(\underline{^{\#}\mathrm{P}}, \underline{^{\#}\mathrm{A}}) \urcorner$ if $f(^{\#}\mathrm{P}) = {}^{\#}\mathrm{A}$, and

(ii) $\vdash_{\mathrm{H}} \ulcorner \sim\mathrm{Pf}_{\mathrm{H}}(\underline{^{\#}\mathrm{P}}, \underline{^{\#}\mathrm{A}}) \urcorner$ if $f(^{\#}\mathrm{P}) \neq {}^{\#}\mathrm{A}$.

This means there is an *arithmetic* relation denoted by '$\mathrm{Pf}_{\mathrm{H}}(n, m)$' which holds just in case the *metamathematical* relation holds "*n* codes a proof in H of the wff coded by *m*." In this way, there is an indication *within* H for when something is true *about* H (in particular, about a wff having a proof in H).

Diagonal or Fixed Point Lemma

FPL: For any formula B(v) in the language of H, with exactly v free, there is a sentence Φ such that $\vdash_{\mathrm{H}} \ulcorner \Phi \equiv \mathrm{B}(\underline{^{\#}\Phi}) \urcorner$.

-FPL guarantees there is a Gödel sentence G such that $\vdash_{\mathrm{H}} \ulcorner \mathrm{G} \equiv \sim \bigvee x\, \mathrm{Pf}_{\mathrm{H}}(x, \underline{^{\#}\mathrm{G}}) \urcorner$.

Proof of FPL

Define the **diagonaliation** of a formula B(v) with exactly v free as $\ulcorner \mathrm{B}(\underline{^{\#}\mathrm{B(v)}}) \urcorner$. A diagonalization is a sentence which claims that its one-place formula is satisfied by the Gödel number for that very formula. Let diag(*n*) be a function that outputs the Gödel number of the diagonalization of B(v) if *n* is the Gödel number of B(v); undefined otherwise. Intuitively, this function is computable; hence CT implies it is recursive. And thus, EL secures that it is strongly represented in H, meaning there is a formula D(*x*, *y*) such that:

(i) $\vdash_{\mathrm{H}} \ulcorner \mathrm{D}(\underline{m}, \underline{n}) \urcorner$ if $\mathrm{diag}(m) = n$, and

(ii) $\vdash_{\mathrm{H}} \ulcorner \sim\mathrm{D}(\underline{m}, \underline{n} \urcorner$ if $\mathrm{diag}(m) \neq n$.

Observe that if diag(m) = n, then $\ulcorner D(m, n) \urcorner$ is equivalent to $\ulcorner \bigwedge y\, (D(m, y) \equiv y = n) \urcorner$. So given condition (i), it is equally true thanks to the logical axioms and axioms for '=' that:

1. $\vdash_H \ulcorner \bigwedge y\, (D(\underline{m}, y) \equiv y = \underline{n}) \urcorner$ if diag(m) = n.

Next, consider the one-place formula $\ulcorner \bigvee y\, (D(x, y) \wedge B(y)) \urcorner$. Call it Abby and assume it has Gödel number a. Then, $\ulcorner \bigvee y\, (D(\underline{a}, y) \wedge B(y)) \urcorner$ is the diagonalization of Abby. For convenience, let Φ be identical to that diagonalization, so that trivially:

2. $\vdash_H \ulcorner \Phi \equiv \bigvee y\, (D(\underline{a}, y) \wedge B(y)) \urcorner$

Now since Φ is the diagonalization of Abby, we know that:

3. $\mathrm{diag}(a) = {}^{\#}\Phi$

Therefore:

4.	$\vdash_H \ulcorner \bigwedge y\, (D(\underline{a}, y) \equiv y = \underline{{}^{\#}\Phi}) \urcorner$	[From 1 and 3]
5.	$\vdash_H \ulcorner \Phi \equiv \bigvee y\, (y = \underline{{}^{\#}\Phi} \wedge B(y)) \urcorner$	[From 2 and 4]
6.	$\vdash_H \ulcorner \Phi \equiv B(\underline{{}^{\#}\Phi}) \urcorner$	[From 5 and the axioms for '=']

The Proof of Incompleteness

Argument that G is not provable.

(1) $\vdash_H \ulcorner G \equiv \sim\bigvee x\, Pf_H(x, \underline{{}^{\#}G}) \urcorner$	[From FPL]
(2) $\vdash_H G$	[Suppose for *reductio*]
(3) $\vdash_H \ulcorner \sim\bigvee x\, Pf_H(x, \underline{{}^{\#}G}) \urcorner$	[From (1) and (2)]
(4) $\vdash_H \ulcorner \bigvee x\, Pf_H(x, \underline{{}^{\#}G}) \urcorner$	[From (2) and the arithmetization of the proof relation]
(5) $\nvdash_H G$	[By *reductio*; (3) and (4) contradict that H is consistent]

Argument that ~G is not provable.

(6) $\vdash_H \ulcorner \sim Pf_H(\underline{n}, \underline{{}^{\#}G}) \urcorner$, for any $\underline{n}$	[From (5) and the arithmetization of the proof relation]
(7) $\vdash_H \sim G$	[Suppose for *reductio*]
(8) $\vdash_H \ulcorner \bigvee x\, Pf_H(x, \underline{{}^{\#}G}) \urcorner$	[From (1) and (7)]
(9) $\nvdash_H \sim G$	[By *reductio*; (6) and (8) contradict that H is ω-consistent][1]

Bonus: Argument that G is true.

(10) $\mathcal{N} \vDash \ulcorner G \equiv \sim\bigvee x\, Pf_H(x, \underline{{}^{\#}G}) \urcorner$	[From (1) assuming $\mathcal{N} \vDash$ H]
(11) $\mathcal{N} \vDash \ulcorner \sim\bigvee x\, Pf_H(x, \underline{{}^{\#}G}) \urcorner$	[From (6) assuming $\mathcal{N} \vDash$ H]
(12) $\mathcal{N} \vDash G$	[From (10) and (11)]

---

[1] Consistency is weaker than ω-consistency. If ~G is added as an axiom to H, the resulting system H* is consistent (since G is not derivable from H). But H* is ω-inconsistent, for the derivability claims at (6)-(8) are all true of H*.

After Gödel, Rosser (1936) showed how to prove incompleteness by assuming consistency only. Briefly: FPL yields a sentence R that says "if there is a proof of R, then there is an earlier proof of ~R" (where 'earlier' is defined with respect to the Gödel numbers of these proofs.)

Appendix: Gödel's Second Incompleteness Theorem

G2: There is no derivation in H of the *consistency sentence* (assuming H is a respectable formal arithmetic system). Specifically, $\nvdash_{\text{H}}$ $\ulcorner{\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\bot})\urcorner$, where $\bot$ is the wff '0 = 1'.

Assuming $\mathcal{N} \vDash$ H, ${\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\bot})$ is true in $\mathcal{N}$ iff H is consistent, i.e., the truth of ${\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\bot})$ indicates the consistency of H. And so, G2 means you could derive this consistency-indicator in H only if H is inconsistent. Put a different way, if H is respectable, there is no means within H to derive the consistency sentence.

From another angle, neither the consistency sentence nor its negation is provable in H, assuming H is respectable…so the consistency sentence is another point at which H is incomplete.

The Basic Argument

(13) $\vdash_{\text{H}}$ $\ulcorner{\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\bot})\urcorner$	[Suppose for *reductio*]
(14) $\vdash_{\text{H}}$ $\ulcorner{\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\bot}) \supset {\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\text{G}})\urcorner$	[Lemma]
(15) $\vdash_{\text{H}}$ $\ulcorner{\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\text{G}})\urcorner$	[From (13) and (14)]
(16) $\vdash_{\text{H}}$ G	[From (1) and (15)]
(17) $\nvdash_{\text{H}}$ $\ulcorner{\sim}\bigvee x\ \text{Pf}_{\text{H}}(x, \underline{{}^{\#}\bot})\urcorner$	[By *reductio*; contradiction at (5) and (16)]

The proof of the Lemma at (14) is omitted here. But basically, you can prove *inside* H that G is not provable in H…assuming that H is consistent (which is what the antecedent in (14) effectively says).