Lecture IX

Cantor's Diagonal Principle

A relation is called *arithmetical* if it is definable in L, the language of arithmetic. Since L contains RE, it follows that all r.e. relations are arithmetical. Also, since L contains negation, it follows that all complements of r.e. relations are arithmetical. That L contains negation also implies that the enumeration theorem fails for arithmetical sets, i.e. there is no arithmetical relation that enumerates all the arithmetical relations; similarly, there is no recursive relation that enumerates all the recursive relations.

The best way to see this is by proving a general theorem. As in the enumeration theorem for r.e. sets, if R is a two-place relation, we write R_x for {y: R(x, y)}. We give the following

Definition: Let X be a set, F be a family of subsets of X, and R a two place relation defined on X. R is said to *supernumerate* F iff for any $S \in F$, there is an $x \in X$ such that $S = R_x$. R is said to *enumerate* F iff R supernumerates F and for all $x \in X$, $R_x \in F$.

The content of the enumeration theorem is thus that there is an r.e. relation which enumerates the r.e. sets. Next we have

Cantor's Diagonal Principle: The following two conditions are incompatible:

(i) R supernumerates F

(ii) The complement of the*Diagonal Set* is in F (the Diagonal Set is $\{x \in X: R(x, x)\}$).

Proof: Suppose (i)-(ii) hold. Then by (ii) $X-\{x \in X: R(x, x)\} = \{x \in X: \sim R(x, x)\} \in F$. By (i), $\{x \in X: \sim R(x, x)\} = R_y$ for some y. But then R(y, x) iff ~R(x, x) for all $x \in X$, so in particular R(y, y) iff ~R(y, y), contradiction.

Cantor applied this lemma in the case F = power set of X to show that a set is never in 1-1 correspondence with its own power set. We can apply it to formal languages by letting F be the family of sets definable in a given language and letting R be a relation definable in the language. Unless the language is very strange indeed, (ii) will be satisfied, so (i) will be false. In the case of RE, we know from the enumeration theorem that (i) is satisfied, so it follows that (ii) fails, and therefore that negation is not definable in RE. In the case of L, on the other hand, (ii) holds, so (i) must fail. The same applies to the language Lim. Finally, if we let F be the family of recursive sets and R be an arbitrary recursive relation, (ii) clearly holds, so (i) fails and no recursive relation enumerates the recursive relations. (To see that

(ii) holds in this case, let R be a recursive relation, and let A(x, y) and B(x, y) define R and -R, respectively, in RE. Then the diagonal set is defined in RE by A(x, x), and its complement is defined by B(x, x).)

A First Version of Gödel's Theorem.

We now know enough to prove a version of Gödel's first incompleteness theorem. A sentence is said to be *undecidable* in an axiom system Γ if neither it nor its negation is a theorem of Γ, and Γ is said to be *incomplete* if some sentence is undecidable in it. We normally use the same letter to denote a set of axioms and its set of theorems. If a set of axioms is r.e., so is its set of theorems (by the Generated Sets Theorem). Similarly, if a set of axioms is arithmetical, so is its set of theorems. We have the following

Theorem: Every arithmetical set Γ of true sentences of the language L is incomplete.
Proof: Since Γ consists of true sentences, if Γ were complete, then the true sentences of L would be precisely the theorems of Γ. But as Γ is arithmetical, the set of theorems of Γ is also arithmetical, i.e. definable in L. And as we have seen earlier, the set of true sentences of L is not definable in L.

The theorem implies that every r.e. set of true sentences of L is incomplete.

In Gödel's original result the assumption that Γ is a set of true sentences was weakened, and hence Gödel's original result is stronger. An axiom system Γ in the language L is said to be *ω-consistent* if there is no formula A(x) such that Γ fi (∃x)A(x) but Γ fi ~A($\mathbf{0}^{(n)}$) for all n. Obviously, an axiom system consisting of true sentences of L is ω-consistent. An axiom system can be consistent without being ω-inconsistent, however. Gödel showed (in effect) that if Γ is an r.e. ω-consistent extension of Q, then Γ is incomplete. We shall not prove the full result this time, though we shall prove some related results. Rosser later showed that the assumption of ω-consistency can be weakened still further, and that no *consistent* r.e. extension of Q is complete.

One of Gödel's main intents was to prove the theorem we just gave. The reason he gave a stronger result must be understood in the light of the fact that, in the discovery and presentation of his results, he was oriented by Hilbert's program. In a nutshell, Hilbert's program demanded a proof of the consistency of the formal systems that codified the theories of classical mathematics, a proof in which, roughly, no appeal to notions or principles involving infinities was made: only so called 'finitistic' principles and methods of proof were to be employed in proofs about the properties of formal systems. The notion of truth in the standard interpretation of the language of arithmetic is a typically non-finitistic one, and hence not usable within the context of Hilbert's program. However, the notions of ω-consistency and consistency are finitistic.

More Versions of Gödel's Theorem

If B(x_i) is a formula of L that defines a set S, let us say that a system Γ is *correct* for B if Γ fi B($\mathbf{0}^{(n)}$) implies that n ∈ S, and *complete* for B if n ∈ S implies that Γ fi B($\mathbf{0}^{(n)}$).

Theorem: If Γ is r.e. and B defines a set which is not r.e., then Γ is not both correct and complete for B. That is, the set S' = {n: Γ fi B($\mathbf{0}^{(n)}$)} is different from S.
Proof: This is simply because S' is r.e., since it is defined by the formula $(\exists x)(\exists m)(\mathrm{Num}(m,n) \wedge \mathrm{NSubst}(\mathbf{0}^{(k)},x,[\mathbf{0}^{(1)},\mathbf{0}^{(i)}],m) \wedge \mathrm{Th}(x))$ where Th(x) is an RE formula defining the set of theorems of Γ and k is the Gödel number of B(x_i) (we can use naive substitution because we may assume that B does not contain bound occurrences of x_i).

So when S is not recursively enumerable there's a difference between being an element of S and being provably an element of S. If Γ is a true set of axioms, and thus correct, there will be an instance B($\mathbf{0}^{(n)}$) that is true but unprovable from Γ.

This is a slight generalization of a result due to Kleene. Kleene's result was that no r.e. axiom system can be complete and correct for any formula that defines -K, and thus in particular for the formula ~W(x, x). In fact, this holds for formulae defining -S whenever S is a nonrecursive r.e. set.

Thus the interest of the theorem depends on the previous proof that there *are* r.e. nonrecursive sets (which in turn depends on the Enumeration Theorem). We can, however, state a theorem which does not depend on this fact (or on any important fact of recursion theory), and which says that any formal system must be incomplete for any formula defining the complement of *some* r.e. set:

Theorem: If Γ is an r.e. set of true axioms, then there is an r.e. set S such that if A(x_1) defines -S, some instance A($\mathbf{0}^{(n)}$) is true but unprovable.
Proof: Suppose, for a contradiction, that for every r.e. set S at least one formula A(x_1) defining -S is such that Γ is complete for A. Then the following relation would be a supernumeration of the complements of the r.e. sets: R(m,n)={<m,n>: m is a Gödel number of a formula A(x_1) and m is provable of n}; this relation is clearly r.e., using the same reasoning as in the proof above. But now we can use Cantor's Diagonal Principle, and conclude that the complement of the diagonal set {n:R(n,n)} cannot be the complement of an r.e. set. But this is absurd, since {n:R(n,n)} is an r.e. set (if B(x,y) is an RE formula that defines R, then B(x,x) defines {n:R(n,n)}).

Q is RE-Complete

Call a set Γ *RE-complete* if every true sentence of RE is a theorem of Γ, and *RE-correct* if every theorem of Γ which is a sentence of RE is true. Whenever Γ fi $A(\mathbf{0}^{(n)})$ iff $n \in S$ for all n, A(x) is said to *weakly represent* S in Γ, and S is said to be *weakly representable* in Γ if some formula weakly represents it in Γ. (We also say that S is *numerable* in Γ.) Thus, any r.e. set is weakly representable in any RE-complete and correct axiom system. Moreover, if Γ is an r.e. set which is RE-complete and correct, then the sets weakly representable in Γ are *precisely* the r.e. sets, since any set weakly representable in an r.e. axiom system Γ is r.e. (To see this, recall that if $A(x_i)$ weakly represents S in Γ, k is a Gödel number of $A(x_i)$, and Th(x) is an RE formula that defines the set of theorems of Γ, then the RE formula $(\exists x)(\exists m)(\mathrm{Num}(m,n) \wedge \mathrm{NSubst}(\mathbf{0}^{(k)},x,[\mathbf{0}^{(1)},\mathbf{0}^{(i)}],m) \wedge \mathrm{Th}(x))$ defines S.)

It turns out that Q is RE-complete and correct. Q is obviously RE-correct, because all of its axioms are true; it takes a bit more work to show that Q is RE-complete. The main fact we need to show this is

(Fact 1) Q fi $(x_1)(x_1 < \mathbf{0}^{(n)} \equiv (x_1 = \mathbf{0} \vee \ldots \vee x_1 = \mathbf{0}^{(n-1)}))$ for all n > 0, and
Q fi $(x_1) \sim(x_1 < \mathbf{0})$

Another useful fact is

(Fact 2) For all n, Q fi $(x_1)(x_1 = \mathbf{0}^{(n)} \vee x_1 < \mathbf{0}^{(n)} \vee \mathbf{0}^{(n)} < x_1)$

Fact 2 is not necessary to prove that Q is RE-complete, however. We shall not prove either fact, but we shall give a proof that Q is RE-complete.

It is also worth noting that a natural strengthening of Fact 2, namely that Q fi $(x_1)(x_2)(x_1 = x_2 \vee x_1 < x_2 \vee x_2 < x_1)$, is false. We can show this by constructing an interpretation in which the axioms of Q are true but the statement $(x_1)(x_2)(x_1 = x_2 \vee x_1 < x_2 \vee x_2 < x_1)$ is false. The domain of this interpretation is $\mathbf{N} \cup \{\alpha, \beta\}$, where α and β are two new elements not in **N**. The constant **0** still denotes 0, and successor, addition and multiplication have the same interpretations as before when restricted to the natural numbers. When the arguments include α or β, we make the following stipulations:

Successor: $\alpha' = \alpha$, $\beta' = \beta$
Addition: $n + \alpha = \alpha + n = \alpha$; $n + \beta = \beta + n = \beta$; $\alpha + \alpha = \alpha + \beta = \alpha$; $\beta + \beta = \beta + \alpha = \beta$
Multiplication: $\alpha \cdot 0 = \beta \cdot 0 = 0$; $\alpha \cdot n = \alpha$, $\beta \cdot n = \beta$ $(n > 0)$; $n \cdot \alpha = \alpha$, $n \cdot \beta = \beta$ (all n); $\alpha \cdot \alpha = \beta \cdot \alpha = \alpha$; $\beta \cdot \beta = \alpha \cdot \beta = \beta$

(where n ranges over the natural numbers) We leave it to the reader to verify that the axioms of Q are true in this interpretation, but that neither $\alpha < \beta$ nor $\beta < \alpha$ holds.

We are now ready to prove our theorem about Q.

Theorem: Q is RE-complete.
Proof: We show by induction on the complexity of sentences that every true sentence of RE is a theorem of Q.

(1) Atomic sentences. First, note that every true atomic sentence involving = is provable, since any such sentence is of the form $\mathbf{0}^{(n)} = \mathbf{0}^{(n)}$ and therefore follows from the identity axioms. Next, we show by induction on n that $A(\mathbf{0}^{(m)}, \mathbf{0}^{(n)}, \mathbf{0}^{(p)})$ is provable for all m, where $p = m + n$. $A(\mathbf{0}^{(m)}, \mathbf{0}, \mathbf{0}^{(m)})$ follows from axiom 4 of Q and is therefore provable. If Q fi $A(\mathbf{0}^{(m)}, \mathbf{0}^{(n)}, \mathbf{0}^{(p)})$, then by axiom 5 we see that Q fi $A(\mathbf{0}^{(m)}, \mathbf{0}^{(n+1)}, \mathbf{0}^{(p+1)})$. So Q proves all the true atomic sentences involving A; that Q proves all the true atomic sentences involving M follows similarly from the recursion axioms for multiplication.

(2) Conjunctions. Suppose A and B are theorems of Q if true. If their conjunction is true, both of them are true, so both are provable, and so is their conjunction.

(3) Disjunctions. Similar to the preceding case.

(4) Existential quantification. Suppose any statement less complex than $(\exists x)A(x)$ is a theorem of Q if true. If $(\exists x)A(x)$ is true, so must be one of its instances $A(\mathbf{0}^{(n)})$, which is then provable. But then so is $(\exists x)A(x)$.

(5) Bounded universal quantification. Suppose $(x_i < \mathbf{0}^{(n)})A$ is a true RE sentence. Then all of $A(\mathbf{0})$, ... $A(\mathbf{0}^{(n-1)})$ are true, and hence provable by the inductive hypothesis. Therefore Q fi $(x_i)((x_i = \mathbf{0} \vee \ldots \vee x_i = \mathbf{0}^{(n-1)}) \supset A)$, and so by Fact 1, Q fi $(x_i)(x_i < \mathbf{n} \supset A)$.

It follows, as we have seen, that the sets representable in Q are precisely the r.e. ones. We have a related result about Lim:

Theorem: Q proves, among the sentences of Lim, exactly the true ones.
Proof: This time, we show by induction on the complexity of sentences that for all sentences A of Lim, Q fi A if A is true and Q fi ~A if A is false.

(1) Atomic sentences. We have already proved half our result; we only need to show that all false atomic sentences are refutable in Q. Moreover, if we can show this for sentences involving =, the result will follow for those involving A and M: if $p \neq m + n$, then Q fi $A(\mathbf{0}^{(m)}, \mathbf{0}^{(n)}, \mathbf{0}^{(k)})$ (where $k = m + n$) and Q fi $\mathbf{0}^{(k)} \neq \mathbf{0}^{(q)}$, so by the uniqueness axiom for A, Q fi $\sim A(\mathbf{0}^{(m)}, \mathbf{0}^{(n)}, \mathbf{0}^{(p)})$; and similarly for multiplication.

First, observe that by axiom 1 of Q, Q fi $\mathbf{0} \neq \mathbf{0}^{(n)}$ when $n > 0$ (since then $\mathbf{0}^{(n)}$ is a term of the form t'). Next, note that axiom 2 is equivalent to $(x_1)(x_2)\ (x_1 \neq x_2 \supset x_1' \neq x_2')$, so we can show by induction on k that Q fi $\mathbf{0}^{(n)} \neq \mathbf{0}^{(p)}$ where $p = n + k$. It follows that whenever $n < m$, Q fi $\mathbf{0}^{(n)} \neq \mathbf{0}^{(m)}$. Finally, by the identity axioms we have that Q fi $\mathbf{0}^{(m)} \neq \mathbf{0}^{(n)}$.

(2) Negation. Suppose A is the sentence ~B. If A is true, then B is false and by the inductive hypothesis Q fi ~B, i.e. Q fi A. If A is false, then B is true, so by the inductive hypothesis Q fi B, so Q fi ~~B (= ~A).

(3) Conjunction and disjunction. These are straightforward, and we shall only do the case of conjunction. Suppose A = (B $\wedge$ C). If A is true, then so are B and C, so Q fi B, Q fi C, and so Q fi (B $\wedge$ C). If A is false, then either B or C is false; suppose B is. Then Q fi ~B, so Q fi (~B $\vee$ ~C), and so Q fi ~(B $\wedge$ C).

(4) Bounded universal and existential quantification. Again, we only do universal quantification, as the other case is similar. If ($x_i < \mathbf{0}^{(n)}$)A is true, then A($\mathbf{0}$), ..., A($\mathbf{0}^{(n-1)}$) are true, so Q fi A($\mathbf{0}$), ..., Q fi A($\mathbf{0}^{(n-1)}$), and by Fact 1, Q fi ($x_i < \mathbf{0}^{(n)}$)A. If ($x_i < \mathbf{0}^{(n)}$)A is false, then A($\mathbf{0}^{(k)}$) is false for some k < n, so Q fi ~A($\mathbf{0}^{(k)}$); Less($\mathbf{0}^{(k)}$, $\mathbf{0}^{(n)}$) is a true sentence of RE, so Q fi Less($\mathbf{0}^{(k)}$, $\mathbf{0}^{(n)}$), and so Q fi ($\exists x_i$)(Less(x_i, $\mathbf{0}^{(n)}$) $\wedge$ ~A) and Q fi ~($x_i < \mathbf{0}^{(n)}$)A.

A formula A(x) is said to *binumerate* a set S in a system Γ iff for all n, n $\in$ S iff Γ fi A($\mathbf{0}^{(n)}$) and n $\notin$ S iff Γ fi ~A(n). If some formula binumerates S in Γ, then we say that S is *binumerable* in Γ (or *numeralwise expressible*, or *strongly representable*, or even simply *representable*). Clearly, if a set is binumerable in Γ then both it and its complement are numerable, so in particular if Γ is r.e., then any set binumerable in Γ is recursive. So not all r.e. sets are binumerable in Q. The converse, that all recursive sets are binumerable in Q, is true but not evident at this point: if S is recursive, then we have some formula A which numerates S in Q and some formula B which numerates -S, but we don't yet have a *single* formula which numerates both S and -S. The theorem we just proved shows that all sets definable in Lim are binumerable in Q, since if A(x) is a formula of Lim that defines S, then A(x) binumerates S.

The facts about weak representability in Q just given also hold for arbitrary r.e. extensions of Q that have true axioms. However, they do not hold for *arbitrary* extensions of Q, or even arbitrary r.e. extensions. For example, let Γ be an inconsistent set. Then Γ clearly extends Q, but only one set is weakly representable in Γ, namely **N** itself (since for any A and any n, A($\mathbf{0}^{(n)}$) is a theorem of Γ). Also, no set is strongly representable in Γ (since we will always have Γ fi A($\mathbf{0}^{(n)}$) and Γ fi ~A($\mathbf{0}^{(n)}$)). However, they do hold for arbitrary *consistent* r.e. extensions of Q. That is, if Γ is a consistent r.e. extension of Q, then the sets weakly representable in Γ are precisely the r.e. ones. (Again, it is easy to show that all sets weakly representable in Γ are r.e.; the hard part is showing that all r.e. sets are representable in Γ.) Moreover, as Shepherdson has shown, every r.e. set is weakly represented in Γ by some formula that actually defines it, though it is not necessarily the case that every formula that defines it weakly represents it in Γ. The proof of this result is tricky, however. It is easier to prove if we only require Γ to be ω-consistent; we will prove this later.

Let Γ be any consistent extension of Q whatsoever, and let A(x) be a formula of RE that defines a set S. Then whenever Γ fi ~A($\mathbf{0}^{(n)}$), n $\notin$ S. To see this, suppose Γ fi ~A($\mathbf{0}^{(n)}$) and n $\in$ S. Then A($\mathbf{0}^{(n)}$) is a true sentence of RE, and so Q fi A($\mathbf{0}^{(n)}$); since Γ extends Q, Γ fi A($\mathbf{0}^{(n)}$). But then both A($\mathbf{0}^{(n)}$) and ~A($\mathbf{0}^{(n)}$) are theorems of Γ, contradicting our assumption that Γ is consistent. We thus have the following

Theorem: Any consistent extension of Q is correct for negations of formulae of RE.

Lecture X

True Theories are 1-1 Complete.

Theorem: If Γ is an r.e. set which is RE-complete and correct, then the theorems of Γ form a 1-1 complete set.
Proof: Suppose Γ is such a set. Let S be any r.e. set, and let $A(x_i)$ be a formula of RE that defines it; we can assume A to contain no bound occurrences of x_i. We define $\phi(n)$ to be the least among the Gödel numbers of $A(\mathbf{0}^{(n)})$. ϕ is recursive: its graph $\{<n,y>: \phi(n)=y\}$ is defined in RE by the formula $(\exists m\le y)(\mathrm{Num}(m,n) \wedge \mathrm{NSubst}(\mathbf{0}^{(k)},y,[\mathbf{0}^{(1)},\mathbf{0}^{(i)}],m) \wedge \mathrm{Th}(x) \wedge (w<y)(\sim\mathrm{NSubst}(\mathbf{0}^{(k)},w,[\mathbf{0}^{(1)},\mathbf{0}^{(i)}],m)))$ (where Th(x) is an RE formula defining the set of theorems of Γ and k is the Gödel number of $B(x_i)$); notice that the use of negation in the last conjunct is legitimate, since the formula it affects is equivalent to a formula of Lim^+. Clearly ϕ is 1-1, and for any n, $n \in S$ iff $A(\mathbf{0}^{(n)})$ is true, iff Γ fi $A(\mathbf{0}^{(n)})$, iff $\phi(n)$ belongs to the set of Gödel numbers of theorems of Γ. So ϕ: $S \le_1$ {theorems of Γ}. Finally, the set of theorems of Γ is r.e., and therefore is 1-1 complete.

It follows that the theorems of Q form a 1-1 complete set, and hence a nonrecursive set. In fact, we can prove the stronger result that if Γ is any r.e. set of true axioms, the set of theorems of Γ is 1-1 complete, as we will see shortly.

Let us say that a formula A(x) of the language of arithmetic *nicely* weakly represents a set S in a theory Γ if it weakly represents S in Γ and also defines S. We may similarly define "nicely *strongly* represents". Similarly, a formula $A(x_1, ..., x_n)$ nicely weakly (strongly) represents an n-place relation R in Γ if it both weakly (strongly) represents R in Γ and also defines R.

It follows from our results of the last lecture that any r.e. set is nicely weakly representable in Γ whenever Γ is true and extends Q. We shall now see that the latter requirement, that Γ extend Q, is unnecessary: any r.e. set is nicely weakly representable in any set Γ of true axioms of the language of arithmetic. Before proving this, we shall need the following theorem:

Deduction Theorem: For any set Γ and any sentences A and B (of any first-order language), if Γ, A fi B then Γ fi $A \supset B$. (Here, Γ, A fi B means $\Gamma \cup \{A\}$ fi B.)
Proof: Suppose Γ, A fi B, and let M be a model of Γ (i.e. an interpretation in which every element of Γ is true). If A is true in M, then M is a model of $\Gamma \cup \{A\}$, and so by the soundness of the predicate calculus B is true in M, so $A \supset B$ is true in M. If A is false in M, then again $A \supset B$ is true in M. So $A \supset B$ is true in all models of Γ, and therefore by the completeness theorem Γ fi $A \supset B$.

The proof we just gave is model-theoretic; however, it is possible to establish the deduction theorem proof-theoretically, by showing how to transform any proof of B from $\Gamma \cup \{A\}$ into a proof of $A \supset B$ from Γ. Such a proof-theoretic argument might be more satisfying, since the model-theoretic argument merely shows that whenever a proof of A from $\Gamma \cup \{A\}$ exists, then a proof of $A \supset B$ from Γ exists, and leaves it an open question whether there is any direct way to transform the former into the latter.

Now let $A(x_1)$ be any sentence of RE that defines a set S; we claim that the formula $Q \supset A(x_1)$ nicely weakly represents S in any system Γ with true axioms. (By Q we mean here some conjunction of the axioms of Q; such a conjunction exists because Q's axioms are finite in number.) Clearly, $Q \supset A(x_1)$ defines S; we must show that $Q \supset A(\mathbf{0}^{(n)})$ is a theorem of Γ iff $n \in S$. First, suppose that $n \in S$. Since Q is RE-complete, Q fi $A(\mathbf{0}^{(n)})$. Clearly, Γ, Q fi $A(\mathbf{0}^{(n)})$. By the deduction theorem, Γ fi $Q \supset A(\mathbf{0}^{(n)})$. Conversely, suppose Γ fi $Q \supset A(\mathbf{0}^{(n)})$. Then since Γ is true, $Q \supset A(\mathbf{0}^{(n)})$ is also true. But Q is true, so $A(\mathbf{0}^{(n)})$ is true. But $A(\mathbf{0}^{(n)})$ is a sentence of RE, and is true iff $n \in S$. So $n \in S$, and we are done. Therefore we have established this

Theorem: If Γ is a set of true sentences of L, then every r.e. set is nicely weakly representable in Γ.

Corollary: For such a Γ, the set of all theorems of Γ is a set to which all r.e. sets are 1-1 reducible. If Γ is r.e., then Γ's theorems form a 1-1 complete set.

Note that, while every r.e. set is nicely weakly representable in such a Γ, we have not shown that every r.e. set is nicely representable by every formula that defines it, or even every RE formula that defines it. If we require Γ to extend Q, on the other hand, then every RE formula that defines S represents it in Γ, because any such Γ is RE-complete and correct.

Church's Theorem

Note that the empty set Ø is trivially a set of true axioms; it follows from our theorem that every r.e. set is nicely weakly representable in Ø, and therefore that Ø's theorems, i.e. the valid formulae of L, form a 1-1 complete set (since Ø is r.e.). So the set of valid formulae of L is not recursive (when the set of theorems of a theory is not recursive, the theory is called *undecidable*; this use of the term 'undecidable' must not be confused with the use we are familiar with, in which the term applies to sentences). This is called *Church's Theorem*. Note that whether a formula is valid does not depend on the interpretation of the nonlogical vocabulary, and therefore that Church's theorem does not depend on the interpretation of the predicates and function symbols of L: for any language with two 3-place predicates, one 2-

place predicate, a constant, and a 1-place function symbol, the set of valid formulae of that language is undecidable, and indeed 1-1 complete.

(Actually, there are two versions of Church's theorem, depending on whether the identity predicate is regarded as a logical symbol. We have been regarding it as nonlogical; when it is regarded as logical, so that the identity axioms are taken as logical axioms, Church's Theorem states that the set of valid formulae (in the present sense of "valid") of a language with two 3-place predicates, a constant and a 1-place function symbol is an undecidable set. The proof is exactly the same.)

Clearly, this result also applies to first-order languages extending L. In fact, we can use a few tricks to show that it also applies to some languages smaller than L. We already know that the constant **0** is redundant in L, since the formula x = **0** is equivalent to (y)A(x, y, x). We can also eliminate the successor function sign, since the graph of the successor function is defined by (∃z)[(w)M(w, z, w) ∧ A(x, z, y)]. Reasoning in this way, we can show that Church's theorem applies to any language with two 3-place predicates. Using a trick that we saw in an exercise, we can eliminate these predicates in favor of a single 3-place predicate defining the graph of the exponentiation function plus a constant for 0 and a 1-place function letter for successor. Using still more devious tricks, we can show that Church's theorem applies to a language which contains only a single 2-place predicate. However, we cannot go any further: the set of valid formulae of a language with only 1-place predicates (with or without identity) is recursive.

(The reasoning we have given is not wholly rigorous. For one thing, while we can find a language K which is properly included in L and which has the same expressive power, we must also show that the above remarks about Q hold for some translation of Q into K. We shall not enter into these considerations here; they will be addressed when we prove the Tarski-Mostowski-Robinson theorem.)

The name "Church's Theorem", though traditional, does not make full justice to its discoverers, since Turing proved the same theorem in his famous original paper on computability; "the Church-Turing theorem" would be a more appropriate name. Also Gödel, in his paper on the incompleteness theorems, stated a very closely related result which, from our vantage point, establishes Church's theorem; but Gödel may not have realized that this was a consequence of his result. Gödel's result is that for any formal system with a primitive recursive set of axioms we can always find a sentence which is not quantificationally valid, but such that the statement that it is not quantificationally valid is not provable in the system.

Complete Theories are Decidable

Theorem: Consider a language in which the set of all sentences is recursive, and let Γ be a set of axioms in this language. If Γ is r.e. and the set of closed theorems of Γ is not

recursive (i.e., undecidable), Γ is incomplete.
Proof: We first give an informal proof using Church's Thesis. We can assume that Γ is consistent, for otherwise the set of theorems of Γ is simply the set of all sentences, which by hypothesis is recursive. Suppose Γ is complete, and let A be any expression. We shall show that either A or ~A is a theorem of Γ. Since the set of sentences of the language of Γ is recursive, we can tell effectively whether A is a sentence. If A is a sentence, then since Γ is complete, either A or ~A is a theorem of Γ. So to see whether A is a theorem of Γ or not, simply run through all the proofs from Γ. If you encounter a proof of A, then A is a theorem; if you encounter a proof of ~A, then A is not a theorem; and since Γ is complete, you will eventually encounter a proof of A or of ~A, so this procedure will eventually terminate. So we can effectively tell whether an arbitrary expression is a theorem of Γ, and so the set of theorems of Γ is recursive.

We now reason more formally. Suppose Γ is complete; again, we may suppose that Γ is consistent. Since Γ is r.e., the set of theorems of Γ is defined in RE by some formula Th(x). Since Γ is complete, an expression A is a nontheorem of Γ just in case either A is a nonsentence or ~A is a theorem. Thus the set of nontheorems of Γ is defined by an RE formula $N(x) \lor (\exists y)(Neg(x, y) \land Th(y))$, where N(x) defines the set of nonsentences of the language of Γ and Neg(x, y) defines the relation *y is the negation of x*. We know that such an RE formula N(x) exists because by hypothesis the set of sentences of the language of Γ is recursive, and Neg(x, y) is easily defined using concatenation. It follows that the set of theorems of Γ is recursive.

It was a while before logicians realized this fact, despite the simplicity of its proof. This may be because the decision procedure given is not intuitively a "direct" one, i.e. a procedure which determines whether A is a theorem of Γ or not by examining A itself.

Note that the requirement that Γ be r.e. is essential here, as is seen by letting Γ equal the set of true sentences of the language of arithmetic: the set of theorems of Γ, i.e. Γ itself, is certainly not recursive, but it is complete, and the set of sentences of L is recursive.

Replacing Truth by ω-Consistency

Let us now state some incompleteness results about ω-consistent, but not necessarily true formal systems. *School* is the set of (consequences of) true atomic sentences of the language of arithmetic and of true negations of atomic sentences.

Theorem: If Γ is arithmetical, ω-consistent and contains School, Γ is incomplete.
Proof: Suppose Γ was complete. Since Γ is arithmetical, it does not coincide with the set of truths, so there must be a sentence A which is either true but unprovable or false but provable from Γ. If A is false but provable, since the system is complete, ~A must be true

but unprovable (ω-consistency implies consistency), and so is any prenex version of it. Take the shortest example B of a true but unprovable sentence: among the prenex true unprovable sentences, take one with the shortest number of quantifiers. This sentence must have some quantifiers, since if Γ contains School, all sentences made up from atomic sentences by use of connectives are decidable in Γ. The first quantifier in B will not be existential, because if it were, some instance of it would be true, and thus a shorter true but unprovable statement (unprovable, because if it were provable so would be its existential generalization). So the first quantifier in B must be universal, which means that all its instances must be true and provable, since they are shorter. Since Γ is complete, ~B must also be provable. But this contradicts the hypothesis that Γ is ω-consistent.

The theorem does not hold if we replace ω-consistency by consistency. There are consistent arithmetical, even recursive, sets of sentences containing School (and extensions of School) which are complete. An example is the set of truths in the structure of the real numbers with a constant for 0 and function letters for successor, addition and multiplication. This is naturally a complete set, which, by a celebrated result of Tarski, is recursive.

On the other hand, if Γ is r.e. and contains Q (not just School), the hypothesis of ω-consistency in the theorem can be weakened to consistency. We will prove this later. Now we can establish the following

Theorem: If Γ is r.e., ω-consistent and contains Q, then Γ is incomplete.
Proof: We know that if Γ contains Q, it is RE-complete and hence that it is correct for negations of RE sentences. Let $A(x_1)$ be an RE-formula that defines K. Then $\sim A(x_1)$ defines -K. Since Γ is r.e., the set {n: $\sim A(\mathbf{0}^{(n)})$ is a theorem of Γ} is r.e. and thus it does not coincide with -K. Since Γ is correct for negations of RE sentences, there is no false provable statement of the form $\sim A(\mathbf{0}^{(n)})$, so there must be a statement of that form which is true but unprovable. This does not yet tell us that $A(\mathbf{0}^{(n)})$ is also unprovable, since Γ need not be a set of true sentences. Let's take again, from among the prenex true unprovable sentences of the form $\sim A(\mathbf{0}^{(n)})$, one with the shortest number of quantifiers. This sentence must have some quantifiers, for the same reason as before. And it cannot begin with an existential, again for the same reason. So the sentence must be of the form (x)C(x), and such that all of its instances are provable. Now, $A(\mathbf{0}^{(n)})$ cannot be provable, since it is equivalent to ~(x)C(x) and that would contradict ω-consistency.

The Normal Form Theorem for RE.

Although our incompleteness results are quite powerful, it is a bit unsatisfying that we have not been able to construct effectively examples of undecidable sentences. One way to do this uses a result that we will prove now.

Normal Form Theorem for RE: Every r.e. relation is definable by a formula of the form $(\exists y)B$, where B is a formula of Lim in which there are no occurrences of negation.

This is a version of a theorem of Kleene, though what he showed was something weaker, namely that every r.e. formula is defined by $(\exists y)B$ for some formula B that defines a *primitive recursive* relation.

To prove it, we prove by induction on the complexity of formulae that every RE formula A is equivalent to a formula $(\exists y)B$, with B a formula of Lim without negation.

(1) A is atomic. Then A is already in Lim without negation, and is equivalent to $(\exists y)A$, where y is a variable that does not occur in A. (That is, we get a formula of the required form by adding a vacuous existential quantifier. Note that A is also equivalent to $(\exists y)(y = y \land A)$, so the use of vacuous quantifiers is not really necessary.)

(2) A is $A_1 \lor A_2$. By the inductive hypothesis, A_1 and A_2 are equivalent to formulae $(\exists y)B_1$ and $(\exists y)B_2$, where B_1 and B_2 are formulae of Lim without negation; so A is equivalent to $(\exists y)B_1 \lor (\exists y)B_2$, which is equivalent to $(\exists y)(B_1 \lor B_2)$, which is of the required form.

(3) A is $A_1 \land A_2$. Again, A_1 and A_2 are equivalent to $(\exists y)B_1$ and $(\exists y)B_2$, with B_1 and B_2 formulae of Lim without negation, so A is equivalent to $(\exists y)B_1 \land (\exists y)B_2$. By rewriting bound variables, we see that A is equivalent to $(\exists z)B_1' \land (\exists w)B_2'$, where B_1' and B_2' come from B_1 and B_2 by changing bound occurrences of y to z (or w) throughout. This is in turn equivalent to $(\exists z)(\exists w)(B_1' \land B_2')$. This is not yet in the required form, since we have two unbounded quantifiers. However, this is equivalent to $(\exists y)(\exists z < y)(\exists w < y)(B_1' \land B_2')$, which is of the required form.

[The usual way to reduce this pair of unbounded quantifiers to a single quantifier uses the pairing function; however, the present approach is simpler.]

(4) A is $(\exists z)A_1$. Then A_1 is equivalent to $(\exists w)B_1$ for some formula B_1 of Lim without negation, and so A is itself equivalent to $(\exists z)(\exists w)B_1$. As in (3), this is equivalent to $(\exists y)(\exists z < y)(\exists w < y)B_1$.

(5) A is $(z < t)A_1$. This is the trickiest case. Let A_1 be equivalent to $(\exists w)B_1$, with B_1 in Lim without negation. A is equivalent to $(z < t)(\exists w)B_1$. We claim that this is equivalent to $(\exists y)(z < t)(\exists w < y)B_1$. To see this, first fix an assignment of values to the free variables. Suppose $(\exists y)(z < t)(\exists w < y)B_1$ holds; then $(z < t)(\exists w < n)B_1$ holds for some n, so *a fortiori* $(z < t)(\exists w)B_1$ holds. Conversely, suppose $(z < t)(\exists w)B_1$ holds, and let n be the denotation of t. Then for each $m < n$, $(\exists w)B_1(w, m)$ holds, so $B_1(k, m)$ holds for some particular k. For each $m < n$, pick a k_m such that $B(k_m, m)$ holds. Since there are only finitely many k_m's, there is a number p such that $p > k_m$ for all $m < n$. So for all $m < n$, there is a $k < p$ (namely k_m) such that $B_1(k, m)$ holds. Therefore, $(z < t)(\exists w < p)B_1$ holds, and so $(\exists y)(z < t)(\exists w < y)B_1$ holds. This completes the proof.

The normal form theorem yields very strong results when combined with the enumeration theorem. From the normal form theorem, we see that the relation W(e, x) is defined by some formula $(\exists y)T(e, x, y)$, where T is a formula of Lim without negation. T is a particular formula, and therefore has a *fixed* number of bounded quantifiers and *no* unbounded quantifiers. It follows that there is a single fixed n such that every r.e. set is defined by some formula with at most n bounded quantifiers and only one unbounded quantifier. This leaves open the possibility that n must be very large; in fact, however, it is known that n can be made rather small.

Whenever we have a normal form theorem, we can combine it with the enumeration theorem to get an analogous result. The most spectacular enumeration theorem we have mentioned is that proved by Matijasevic (building on earlier work by Davis, Putnam and Julia Robinson), that every r.e. set or relation is definable by a formula of the form $(\exists x_1)...(\exists x_n)\ t_1 = t_2$, where t_1 and t_2 are terms involving only the function symbols ', + and · (and the constant **0** and the variables). Note that the formula $t_1 = t_2$ is simply a polynomial equation. Thus the decision problem for any r.e. set is equivalent to the decision problem for some polynomial equation with integer coefficients. The Matijasevic theorem alone does not give us an indication of how large the degree of such equations can be, or of how many variables they may contain. If we apply the enumeration theorem, however, we see that the relation W(e, x) is defined by some *particular* formula $(\exists x_1)...(\exists x_n)\ t_1 = t_2$, whose free variables are e and x. Let us indicate the free variable e by writing this formula as $(\exists x_1)...(\exists x_n)\ t_1(e) = t_2(e)$. Every r.e. set is therefore defined by the formula $(\exists x_1)...(\exists x_n)$ $t_1(\mathbf{0}^{(e)}) = t_2(\mathbf{0}^{(e)})$, for some particular e. So not only is the decision problem for every r.e. set equivalent to the problem of solving some polynomial equation; we can also simultaneously bound the number of variables and the degree of the polynomial.

An immediate application of the normal form theorem is in the proof of the following result:

Theorem: If Γ is r.e., ω-consistent and contains Q, then Γ is complete and correct for the set of all formulae of the form $(\exists y)B$, where B is a formula of Lim without negation.
Proof: Completeness: If $(\exists y)B$ is true, then $B(\mathbf{0}^{(n)})$ is true, for some n. So $B(\mathbf{0}^{(n)})$ will be provable, since Q proves all the true sentences of Lim, and Γ contains Q. Therefore, $(\exists y)B$ will be provable too. Correctness: Suppose $(\exists y)B$ is provable but false. Then all the negations of its instances will be true: ~A(**0**), ~A(**0'**)... So these are all provable, again because Q is complete for the true sentences of Lim. But this contradicts ω-consistency.

Corollary: If Γ is r.e., ω-consistent and contains Q, then every r.e. set is nicely weakly representable in Γ.

Corollary: If Γ is r.e., ω-consistent and contains Q, then the theorems of Γ form a 1-1 complete set.

The hypothesis of ω-consistency in these results could have been replaced by that of consistency, but the corresponding proof is much trickier.

Exercises

1. Prove that if an r.e. set S_1 is 1-1 complete, there is an infinite r.e. set S_2 disjoint from S_1.

2. Prove that a nonempty set is r.e. iff it is the range of a total recursive function.
Note: this result is the origin of the term 'recursively enumerable'. That is, a nonempty set S is r.e. iff there is a recursive function ϕ that enumerates it, i.e. $S=\{\phi(0), \phi(1),...\}$.

3. The Goldbach conjecture is the statement that every even integer greater than 2 is the sum of two primes. Show that this conjecture can be written in the form (x)A, where A is in Lim. Suppose that the conjecture, written in this form, is undecidable in the system we have called Peano Arithmetic. What, if anything, would follow regarding the truth of the Goldbach conjecture itself? (Explain your answer; if nothing follows, explain why, or if something does follow, explain what follows and why.)

4. *Proper* substitution, as opposed to what we have called 'naive substitution', is the substitution of a term for a variable, subject to the following restrictions. Only *free* occurrences of the variable x_i are to be replaced by the term t; and the substitution is improper if any variable occurring in t becomes bound in the result. Define proper substitution in RE, that is, PSubst(m_1,m_2,v,m). where m_2 is the result of a proper substitution of the term m for free occurrences of the variable v in m_1. Use the following fact: an occurrence of a variable x_i within a term is bound in a formula iff the formula is a concatenation of three sequences m, n and p, where the occurrence in question is in the part corresponding to n, and n is (the Gödel number of) a formula beginning with (x_i). m and/or p are allowed to be empty. (This is a form of the usual definition.) Another treatment of proper substitution, which is perhaps more elegant, will be sketched later. It should be clear from the preceding why naive substitution is simpler, at least if this is the treatment adopted.

Lecture XI

An Effective Form of Gödel's Theorem

Recall that Γ is ω-consistent if we never have Γ fi (∃x)A(x) and Γ fi ~A(**n**) for all n. Γ is said to be *ω-complete* if whenever Γ fi A(**n**) for all n, Γ fi (x)A(x). Γ is *ω-inconsistent* iff it is not ω-consistent, and similarly for *ω-incomplete*.

Let us call a formula A *Σ_1* if A is of the form (∃y)B, where B is a formula of Lim, and *Π_1* if it is of the form (y)B for B a formula of Lim. Note that a negation ~(∃y)B of a Σ_1 formula is equivalent to (y)~B, which is Π_1, and that each Π_1 formula (y)B is equivalent to a negation ~(∃y)~B of a Σ_1 formula (and these equivalences are provable). We sometimes use the terms Σ_1 and Π_1 loosely to refer to formulae that are equivalent to formulae that are Σ_1 or Π_1 in the strict sense; we also refer to a *set* or *relation* as Σ_1 (or Π_1) if it is defined by some Σ_1 (or Π_1) formula. It follows from the normal form theorem for RE that the r.e. sets are precisely the Σ_1 sets and the complements of r.e. sets are precisely the Π_1 sets.

We sometimes write Σ^0_1 for Σ_1 and Π^0_1 for Π_1. The superscript zero indicates that the unbounded quantifier ranges over numbers. Other superscripts are possible; in general, when we talk about a Σ^m_n or Π^m_n formula, m indicates the type of the variables in the unbounded quantifiers, and the n indicates the number of alterations between unbounded universal and unbounded existential quantifiers. This will be made more precise later on in the course.

Suppose Γ extends Q. If B is a sentence of Lim, then as we saw in Lecture IX, if B is true, then B is a theorem of G. So let (∃y)B(y) be a true Σ_1 sentence. Since it is true, B($\mathbf{0}^{(n)}$) is true for some n and therefore is a theorem of Γ; but B($\mathbf{0}^{(n)}$) logically implies (∃y)B(y), so (∃y)B(y) is also a theorem of Γ. So every true Σ_1 sentence is a theorem of Γ; in short, Γ is Σ_1-complete. If Γ is also consistent, then it is Π_1-correct, i.e. every Π_1 sentence provable in Γ is true. To see this, let A be a Π_1 sentence provable in Γ. If A is false, then ~A is true; but ~A, being the negation of a Π_1 sentence, is provably equivalent to a Σ_1 sentence, and is therefore provable in Γ, since Γ is Σ_1-complete. But then both A and ~A are theorems of Γ, and so Γ is incomplete. So a consistent extension of Q is both Σ_1-complete and Π_1-correct.

Moreover, as we saw in the last lecture, every ω-consistent system extending Q is Σ_1-correct. Recall the argument: suppose Γ is such a system, and suppose it proves a false Σ_1 sentence (∃y)B(y). Since that sentence is false, B($\mathbf{0}^{(n)}$) is false for all n, and therefore, since Γ extends Q, Γ fi ~B($\mathbf{0}^{(n)}$) for all n, contradicting Γ's ω-consistency. So any ω-consistent extension of Q is Σ_1-complete and correct.

We can now prove an effective version of Gödel's theorem.

Effective Form of Gödel's Theorem: Let Γ be an r.e. extension of Q. Then we can find effectively a Π_1 formula A such that

(1) If Γ is consistent, then A is true but unprovable in Γ

and

(2) If Γ is ω-consistent, then ~A is also unprovable in Γ.

Proof: Since W(e,x) is r.e., it is definable by a Σ_1 formula ($\exists$y)L(e,x,y) which can be effectively found from the original RE formula and through the (effective) proof of the normal form theorem for RE. So K is r.e., and it is definable by ($\exists$y)L(x,x,y), which is Σ_1. -K is then defined by the Π_1 formula (y)~L(x,x,y). The set {n: (y)~L($\mathbf{0}^{(n)}$,$\mathbf{0}^{(n)}$,y) is provable from Γ} is r.e. (for the known reasons), and an RE formula that defines it can be found effectively; therefore also its Gödel number e can be found effectively. Then for all n, ($\exists$y)L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(n)}$,y) is true iff (y)~L($\mathbf{0}^{(n)}$,$\mathbf{0}^{(n)}$,y) is provable. So ($\exists$y)L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y) is true iff (y)~L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y) is provable. Then the Π_1 formula (y)~L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y) cannot be provable from Γ, because given that Γ is a consistent extension of Q, Γ is Π_1-correct, so (y)~L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y) would be true and so would be the equivalent formula ~($\exists$y)L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y), and on the other hand if (y)~L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y) were provable ($\exists$y)L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y) would be true. We therefore may take A to be (y)~L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y). A is not provable, and therefore ~($\exists$y)L($\mathbf{0}^{(e)}$,$\mathbf{0}^{(e)}$,y) and A itself are true.

Now suppose that Γ is ω-consistent. Then Γ is Σ_1-correct. ~A is logically equivalent to a false Σ_1 sentence, and is therefore not a theorem of Γ.

This is an informal argument in the sense that it appeals to the intuitive notion of computability or effectiveness. We could now give a more formal proof without making this appeal. It will be easier to give such a proof later, once we have some more results. We can note here that the effectiveness of the construction of A depends on the fact that we use K, for which the number e with appropriate properties can be effectively found from every Γ. Not every r.e. nonrecursive set would have served the purpose of effectiveness, since as we will show later, for some such sets the corresponding Gödel sentences cannot be effectively found..

The hypothesis of our effective form of Gödel's theorem is already quite weak; in fact, we can weaken it a bit more. In particular, the condition that Γ extends Q can be weakened. The only fact about Q needed in proving that Q is RE-complete and correct for negations of RE sentences is Fact 1, along with the fact that all sentences of School are provable in Q. So these are the only facts needed to show, using the normal form theorem, that Q is Σ_1-complete and Π_1-correct. So the theorem will still hold if Q is replaced by any theory containing School for which Fact 1 holds.

Gödel's Original Proof.

The following is, nearly enough, Gödel's own presentation of the first incompleteness theorem. Let Γ be an r.e. system containing Q. Consider the relation Prov(x, y) which holds if y is provable of x, i.e. if the result of replacing all (free) occurrences of x_1 in the formula (coded by) y by the numeral for x is provable. We can take Prov to be Σ_1, since it can be written out in RE; so it is of the form $(\exists z)L(x, y, z)$ for some formula L(x, y, z) of Lim. Consider the formula $\sim Prov(x_1, x_1)$; it has some Gödel number m. Let G be the sentence $\sim Prov(\mathbf{0}^{(m)}, \mathbf{0}^{(m)})$, i.e. $\sim(\exists z)L(\mathbf{0}^{(m)}, \mathbf{0}^{(m)}, z)$. Suppose G is provable; that is, suppose the formula $\sim Prov(x_1, x_1)$ is provable of m. Then $L(\mathbf{0}^{(m)}, \mathbf{0}^{(m)}, \mathbf{0}^{(k)})$ holds for some k; since this is a true sentence of Lim, it is provable, and so $(\exists z)L(\mathbf{0}^{(m)}, \mathbf{0}^{(m)}, z)$ is provable. But G, which we are supposing to be provable, is just the sentence $\sim(\exists z)L(\mathbf{0}^{(m)}, \mathbf{0}^{(m)}, z)$. So if our system is consistent, G is not provable after all, i.e. $\sim Prov(x_1, x_1)$ is not provable of m. But what G says is that the formula with Gödel number m, namely $\sim Prov(x_1, x_1)$, is not provable of m; so G is true. Therefore G is true but unprovable. As long as the system is Σ_1-correct, its negation is not provable either, and as we have seen, it suffices for this that the system be ω-consistent.

Presented in this way, the proof seems rather tricky, and the undecidable sentence is produced in a very devious way. As we have previously presented it, Gödel's theorem should seem more like the inevitable outcome of the Russell paradox. In fact, there is a way of viewing Gödel's original proof which makes it look this way.

Recall the proof that any fully classical language lacks its own satisfaction predicate. If L, for example, has a predicate Sat(y,x) which defines the relation {<x, y>: y codes a formula which x satisfies}, then L has a predicate Het(x) = ~Sat(x, x), which defines the set of (Gödel numbers of) heterological formulae. But then if we ask whether Het(x) is itself heterological, we can derive a contradiction. (Indeed, that there is no formula defining the set of heterological formulae follows directly from the inconsistency of the instance $(\exists y)(x)(x \in y \equiv x \notin x)$ of the unrestricted comprehension scheme, as we saw before.) It follows from the indefinability of satisfaction that the formula Prov does not define satisfaction.

We can show directly that the Gödel sentence G is true but unprovable, in a way that imitates the reasoning of the last paragraph. Call a formula *Gödel heterological* if is not provable of itself; the formula $\sim Prov(x_1, x_1)$ defines the set of Gödel heterological formulae. Let us write this formula as $GHet(x_1)$. Now we ask, is "Gödel heterological" Gödel heterological? The statement that "Gödel heterological" is Gödel heterological is simply the statement $GHet(\mathbf{0}^{(m)})$, where m is the Gödel number of $GHet(x_1)$. Rather than leading to a contradiction, our question has a definite answer "yes". Suppose "Gödel heterological" were not Gödel heterological, i.e. that $GHet(x_1)$ were provable of m. If $GHet(x_1)$ is provable of m, then $Prov(\mathbf{0}^{(m)}, \mathbf{0}^{(m)})$ is a theorem of Q and therefore of any system extending Q; note that $GHet(\mathbf{0}^{(m)})$ is simply the negation of $Prov(\mathbf{0}^{(m)}, \mathbf{0}^{(m)})$. So if $GHet(x_1)$ is provable of itself, then our system is inconsistent, since both $Prov(\mathbf{0}^{(m)}, \mathbf{0}^{(m)})$ and its negation are provable; so if our system is consistent, then $GHet(x_1)$ is not provable of itself, i.e. is Gödel heterological. This is simply to say that $GHet(\mathbf{0}^{(m)})$ is true but not provable. A similar argument shows that ~GHet($\mathbf{0}^{(m)}$) is also unprovable, provided that the system is ω-consistent. Finally, note that GHet($\mathbf{0}^{(m)}$) is simply the sentence G of the last paragraph. So we have really presented Gödel's own proof, but with a different exposition than is usual.

An analogy is often drawn between the unprovability of the Gödel sentence and the liar paradox. From the present exposition, we see that the analogy with the heterological paradox is even closer. In fact, all we really need in order to see that ~Prov($\mathbf{0}^{(m)}$, $\mathbf{0}^{(m)}$) is true but unprovable is to notice that ~Prov($\mathbf{0}^{(m)}$,$\mathbf{0}^{(m)}$) says '"Is not provable of itself" is not provable of itself', i.e. '"Gödel heterological" is Gödel heterological': it is not essential to our proof (though we may observe this afterwards) that it says "I am not provable".

That there is an analogy both to the heterological paradox and to the liar paradox is no accident, since the heterological paradox is really a special case of the liar paradox. The heterological paradox involves the sentence '"Is not true of itself" is not true of itself'. To say that "is not true of itself" is not true of itself is simply to say that the sentence '"Is not true of itself" is not true of itself' is not true, so this sentence says of itself that it is not true — that is, it is a liar sentence.

The Uniformization Theorem for r.e. Relations.

Definition: A *uniformization* of a binary relation R is a relation S such that:

(i) $S \subseteq R$

(ii) S and R have the same domain, i.e. for any x, there is a y such that R(x, y) iff there is a y such that S(x, y)

(iii) S is single valued, i.e. every x bears S to at most one y (i.e. S is the graph of a partial function).

We can think of a relation R(x, y) as a *many valued* function, with x as the argument and any y such that R(x, y) as one of the values for the argument x. Then, for example, an r.e. relation is a partial recursive many-valued function. A uniformization of a many valued function is a single valued function with the same domain.

S can be regarded as a choice function, i.e. S chooses, for each x, something that x bears R to. This definition can be extended to n+1-place relations in an obvious way.

A uniformization theorem in general says that any relation in a particular class C can be uniformized by a relation in C. If this is so, then C is said to have the *uniformization property*. Note that the class of all relations on the natural numbers has the uniformization property. Taking the 2-place case for simplicity, any relation R(x, y) is uniformized by the relation $R(x, y) \wedge (z < y){\sim}R(x, z)$ (i.e. the relation $y = \mu zR(x, z)$). This also shows that the class of recursive relations has the uniformization property, since the relation $y = \mu zR(x, z)$

is recursive if R is; the same applies to the class of relations definable in Lim and to the arithmetical relations. And this argument can easily be generalized to relations of more than two places.

A trickier case is the class of r.e. relations. The above argument will not show that this class has the uniformization property, since $y = \mu zR(x, z)$ will not in general be r.e. when R is. To see this, let X be any r.e. set which is not recursive, and let R be the r.e. relation $\{<x, y>: (y = 0 \wedge x \in X) \vee y = 1\}$. Let S be the relation $y = \mu zR(x, z)$. If $x \in X$, then $S(x) = 0$, but if $x \notin X$, then $S(x) = 1$. So if S is r.e., then -X can be defined in RE by $S(x, \mathbf{0'})$. But by hypothesis X is nonrecursive, so S is not r.e.

However, we can use a somewhat trickier proof to show that uniformization holds for the r.e. relations. Let R be any 2-place r.e. relation, and let F(x, y) be some formula of RE that defines it. By the normal form theorem, we can take F to be $(\exists z)L(x, y, z)$ for some formula L of Lim. $(\exists z)L(x, y, z)$ is equivalent to $(\exists w)(y = K_1(w) \wedge L(x, K_1(w), K_2(w)))$. We can now define a uniformizing relation S by $(\exists w)(y = K_1(w) \wedge L(x, K_1(w), K_2(w)) \wedge (u < w){\sim}L(x, K_1(u), K_2(u)))$ (intuitively, w is the smallest code of a pair [y,z] for which L(x,y,z) holds). Since L is a formula of Lim, the formula defining the uniformizing relation is a Σ_1 formula and so S is an r.e. relation. This can be generalized to n+1-place r.e. relations in a fairly obvious way. So we have proved the

Uniformization Theorem for r.e. Relations: The class of r.e. relations has the uniformization property.

Corollary: Every r.e. relation can be uniformized by a partial recursive function with the same domain.

The Normal Form Theorem for Partial Recursive Functions.

An application of the proof of the uniformization theorem for r.e. relations is a normal form theorem for partial recursive functions, due to Kleene. Let ϕ be any partial recursive function, and let R be its graph. R is defined by some Σ_1 formula $(\exists z)L(x, y, z)$. As in the proof of the uniformization theorem, we see that R is defined by $y = K_1(\mu wL(x, K_1(w), K_2(w)))$. Since $L(x, K_1(w), K_2(w))$ is a formula of Lim, and K_1 is a function whose graph is definable in Lim, we have the following:

Normal Form Theorem for Partial Recursive Functions: Every n-place partial recursive function is of the form $U(\mu wR(x_1, ..., x_n, w))$ for some relation R definable in Lim and some function U whose graph is definable in Lim.
Proof: The case n = 1 was just proved, and the general case is proved similarly.

This is not exactly what Kleene originally proved; he only required R and U to be primitive recursive.

It is important not to forget the U; it is not true that every partial recursive function is of the form $\mu wR(x_1, ..., x_n, w)$ for some R definable in Lim. Even if we allow R to be an arbitrary recursive relation, this is still wrong. If, on the other hand, we require the function ϕ to be total, then $\phi(x_1, ..., x_n)$ is $\mu y(\phi(x_1, ..., x_n) = y)$, so we can drop the U by taking R to be ϕ's graph; but then we cannot require R to be definable in Lim.

Lecture XII

An Enumeration Theorem for Partial Recursive Functions

We can use the uniformization theorem for r.e. relations to prove an enumeration theorem for partial recursive functions. First we recall that we have a general version of the Enumeration Theorem for n-place r.e. relations. That is, there is an n+1-place relation $W^{n+1}(e, m_1, \ldots, m_n)$ that enumerates the n-place r.e. relations. (So $W^2(e, m_1)$ is just our previous relation W(e, y). We will usually omit the superscript when the context makes it clear which one is intended.) The easiest way to prove this is by defining $W^{n+1}(e,m_1,...,m_n)$ as $W^2(e, [m_1, \ldots, m_n])$. It is clear that this enumerates the n-place r.e. relations. We now have:

Theorem: For all n, there is an n+1-place partial recursive function Φ^{n+1} which enumerates the n-place partial recursive functions, i.e. for each n-place partial recursive function ϕ there is a number e such that $\Phi^{n+1}(e, x_1, ..., x_n) = \phi(x_1, ..., x_n)$ for all $x_1, ..., x_n$ for which ϕ is defined, and is undefined on $e, x_1, ..., x_n$ when ϕ is undefined on $x_1, ..., x_n$.
Proof: We only prove the theorem in the case n = 1. (The general case can be proved either by imitation of this case, or via the pairing function.) Consider the relation W^3 which enumerates the 2-place r.e. relations. Being an r.e. relation itself, it is uniformized by some 2-place partial recursive function Φ. Now let ϕ be any 1-place partial recursive function, and let R be ϕ's graph. R is W^3_e for some e, i.e. for some e, $W^3(e, x, y)$ holds iff $\phi(x) = y$. Since Φ uniformizes W^3, $\Phi(e, x) = y$ iff $\phi(x) = y$; moreover, if $\phi(x)$ is undefined, then $W^3(e, x, y)$ does not hold for any y, and so $\Phi(e, x)$ is undefined.

The number e is called an *index* of the function ϕ. Kleene's notation for $\Phi(e, x)$ is $\{e\}(x)$; so $\{e\}$ denotes the partial recursive function with index e.

Just as no recursive relation enumerates the recursive sets, no total recursive function enumerates the total recursive functions. To see this, suppose Ψ did enumerate the total recursive functions. Let ϕ be the total recursive function $\Psi(x, x) + 1$; then there is an e such that $\phi(x) = \Psi(e, x)$ for all e. So in particular, $\phi(e) = \Psi(e, e)$. But $\phi(e) = \Psi(e, e) + 1$, so we have $\Psi(e, e) = \Psi(e, e) + 1$, which is impossible.

Why doesn't this show that an enumeration of the *partial* recursive functions is impossible? Let $\phi(x) = \Phi(x, x) + 1$. ϕ is a partial recursive function, so it has an index e; so $\Phi(e, x) = \Phi(x, x) + 1$ for all x, and in particular $\Phi(e, e) = \Phi(e, e) + 1$. But this is not a contradiction, for it only shows that $\Phi(e, e)$ is undefined. It is this fact about partial recursive functions, that they can be undefined, that allows there to be an enumeration theorem for partial recursive functions, and indeed this was the point of studying partial recursive functions (as opposed to just total recursive functions) in the first place.

(In some presentations, a new "number" u is introduced to represent an undefined value, i.e. we declare that $\phi(x) = u$ when $\phi(x)$ is undefined. Then every function we care to deal with has a value, of sorts, for every argument. The argument of the last paragraph shows that we must have $u = u + 1$: we showed that $\Phi(e, e) = \Phi(e, e) + 1$, from which it follows that $\Phi(e, e) = u$ ($n \neq n + 1$ for all n other than u), and therefore that $u = u + 1$. A similar argument shows that $\phi(u) = u$ for all partial recursive functions ϕ, i.e. u is a fixed point of all partial recursive functions. We will not use u in this course, however.)

This also provides an example of a partial recursive function which is not totally extendible, i.e. which is not extended by any total recursive function. Specifically, Φ is such a function. For suppose Ψ is a total recursive function extending Φ, and let $\psi(x) = \Psi(x, x) + 1$. ψ is a total recursive function with some index e. Then $\Phi(e, x) = \psi(x) = \Psi(x, x) + 1$ for all x on which y is defined, which is all x, since ψ is total. So $\Phi(e, e) = \psi(e) = \Psi(e, e) + 1$. Since Ψ extends Φ, and therefore agrees on Φ whenever Φ is defined, $\Psi(e, e) = \psi(e) = \Psi(e, e) + 1$, which is impossible.

Given a 2-place partial recursive function which is not totally extendible, we can find a 1-place partial recursive function which is not totally extendible via the pairing function: if $\phi(x, y)$ is such a 2-place partial function, let ψ be a partial recursive 1-place function such that $\psi([x, y]) = \phi(x, y)$ whenever $\phi(x, y)$ is defined. If ψ were totally extendible to some function ψ', then we could let $\phi'(x, y) = \psi'([x, y])$, and ϕ' would be a total recursive function extending ϕ. Alternatively, we could simply observe that the function $\phi(x) = \Phi(x, x) + 1$ is not totally extendible, using the argument of the last paragraph.

Reduction and Separation.

Let C be any class of sets. C is said to have the *separation property* if for any disjoint S_1 and $S_2 \in C$, there is an $S \in C$ such that $-S \in C$, $S_1 \subseteq S$, and $S_2 \subseteq -S$. S is said to separate S_1 and S_2.

Separation fails for the r.e. sets. A pair of r.e. sets which is not separated by any recursive set is called a *recursively inseparable pair*. The proof that there are recursively inseparable pairs of r.e. sets is due to Kleene, using Φ. Let $S_1 = \{m: \Phi(m, m) = 0\}$, and let $S_2 = \{m: \Phi(m, m)$ is defined and $> 0\}$. Clearly, S_1 and S_2 are disjoint r.e. sets. If separation held for the r.e. sets, then there would be a recursive S with $S_1 \subseteq S$ and $S_2 \subseteq -S$. But we can easily derive a contradiction by considering the characteristic function of S, ψ. If $\Phi(m, m) = 0$ then $\psi(m)=1$; if $\Phi(m, m) > 0$ then $\psi(m)=0$. Since S is recursive, ψ is recursive, and so partial recursive, and therefore, for all m, $\psi(m)=\Phi(e, m)$ for some e. Then if $\Phi(e, e) = 0$, $\psi(e)=1=\Phi(e, e)$; and if $\Phi(e, e)>0$, $\psi(e)=0=\Phi(e, e)$. Contradiction.

Let C be any class of sets. C is said to have the *reduction property* if for any $S_1, S_2 \in C$, there are disjoint sets S_1' and $S_2' \in C$ such that $S_1' \subseteq S_1$, $S_2' \subseteq S_2$, $S_1' \cup S_2' = S_1 \cup S_2$.

The class of r.e. sets has the reduction property. This can be seen by an application of the uniformization theorem. Specifically, let S_1 and S_2 be r.e. sets, and let R be the many-valued function that takes the value 1 on S_1 and 0 on S_2 (and therefore takes both values on $S_1 \cap S_2$); apply uniformization to shrink R to a single-valued function R' with the same domain. Then we let $S_1'=\{m: R'(m,1)\}$ and $S_2'=\{m: R'(m,0)\}$. S_1' and S_2' obviously have the desired properties.

If a class C has the reduction property, the corresponding class $C_1=\{-X: X\in C\}$ has the separation property. For suppose we have two disjoint sets S_1 and S_2 in C_1. Then $-S_1 \cup -S_2=\mathbf{N}$. Applying to $-S_1$ and $-S_2$ the reduction property of C, we know that there are S_1' and S_2' in C such that $-(S_1') \subseteq -S_1$, $-(S_2') \subseteq -S_2$, and such that $-(S_1') \cup -(S_2') = -S_1 \cup -S_2 =\mathbf{N}$, and $-(S_1')$ and $-(S_2')$ are disjoint. So $S_1' \cup S_2'=\mathbf{N}$ and S_1' and S_2' are disjoint and therefore $S_2'=-(S_1')$; and moreover $S_1 \subseteq S_1'$ and $S_2 \subseteq S_2'$. So S_1' separates S_1 and S_2.

This fact can be used to prove that separation holds for the Π_1 sets (which are the co-r.e. sets). On the other hand, they cannot have the reduction property, because that would imply that the Σ_1 sets (i.e., the r.e. sets) have the separation property, and they don't. We may also note that the Π_1 relations do not have, unlike the Σ_1 relations, the uniformization propety: if they did, we could imitate the proof we gave for the r.e. sets to prove that the Π_1 sets have the reduction property.

Functional Representability.

We have said what it is for a relation to be weakly or strongly representable in a theory; we now define a notion of representability in a theory for partial functions.

Definition: A partial function ϕ is *represented* in a theory Γ by a formula $A(x_1, ..., x_n, y)$ iff whenever $\phi(a_1, ..., a_n) = b$, Γ fi $A(\mathbf{0}^{(a_1)}, ..., \mathbf{0}^{(a_n)}, \mathbf{0}^{(b)}) \wedge (y)(A(\mathbf{0}^{(a_1)}, ..., \mathbf{0}^{(a_n)}, y) \supset y = \mathbf{0}^{(b)})$. ϕ is *representable* in Γ iff some formula represents it in Γ.

Notice that in our definition we do not say what happens when $\phi(a_1, ..., a_n)$ is undefined. In particular, we do not require that $A(\mathbf{0}^{(a_1)}, ..., \mathbf{0}^{(a_n)}, \mathbf{0}^{(b)})$ not be a theorem. So whenever a formula A represents a function ϕ in Γ, A also represents each subfunction of ϕ; in particular, every formula represents the completely undefined function in every theory. Also, if A represents ϕ in Γ and ϕ has an infinite domain, then ϕ has $2^{\aleph_0}$ subfunctions, and so A represents $2^{\aleph_0}$ functions in Γ. It follows that not every function that A represents is partial recursive, since there are only $\aleph_0$ partial recursive functions. Notice also that in an inconsistent theory, every formula represents every function.

Notice that representability is different from definability: a formula can represent a function without defining it, and vice versa. Notice also that if Δ extends Γ, then every function representable in Γ is representable in Δ, since if $A(\mathbf{0}^{(a_1)}, ..., \mathbf{0}^{(a_n)}, \mathbf{0}^{(b)}) \wedge (y)(A(\mathbf{0}^{(a_1)}, ..., \mathbf{0}^{(a_n)}, y) \supset y = \mathbf{0}^{(b)})$ is a theorem of Γ, then it is also a theorem of Δ.

We now set out to prove the theorem that every partial recursive function is representable in Q. To this effect we prove the following two lemmas.

Lemma 1: If two partial 1-place functions ϕ_1 and ϕ_2 are both representable in a theory Γ so is their composition $\phi_2(\phi_1(x))$.
Proof: Let $R_1(x,y)$ and $R_2(y,z)$ represent ϕ_1 and ϕ_2 respectively. Then it is not difficult to verify that the formula $(\exists y)(R_1(x,y) \wedge R_2(y,z))$ represents their composition.

Lemma 2: Any partial function whose graph is definable in Lim is representable in Q.
Proof: Let ϕ be any 1-place partial function whose graph is defined by the formula A(x, y) of Lim. Let B(x, y) be the formula $A(x, y) \wedge (z < y){\sim}A(x, z)$. We claim that B represents ϕ in Q. Suppose $\phi(a) = b$. We have to verify two things: namely, that $B(\mathbf{0}^{(a)}, \mathbf{0}^{(b)})$ is a theorem of Q, and that $(y)(B(\mathbf{0}^{(a)}, y) \supset y = \mathbf{0}^{(b)})$ is a theorem of Q. Clearly, $A(\mathbf{0}^{(a)}, \mathbf{0}^{(b)})$ is a theorem of Q, since $A(\mathbf{0}^{(a)}, \mathbf{0}^{(b)})$ is a true RE sentence. To show that Q fi $B(\mathbf{0}^{(a)}, \mathbf{0}^{(b)})$, we must also show that Q fi $(z < \mathbf{0}^{(b)}){\sim}A(\mathbf{0}^{(a)}, z)$. But again, this is a true sentence of RE, and is therefore a theorem of Q.

Next, we must show that Q fi $(y)(B(\mathbf{0}^{(a)}, y) \supset y = \mathbf{0}^{(b)})$. Here we use Fact 2 about Q from Lecture IX, i.e. for all n, Q fi $(x_1)(x_1 = \mathbf{0}^{(n)} \vee x_1 < \mathbf{0}^{(n)} \vee \mathbf{0}^{(n)} < x_1)$. Using this fact, we establish $(y)(B(\mathbf{0}^{(a)}, y) \supset y = \mathbf{0}^{(b)})$ by reasoning within Q. Suppose $B(\mathbf{0}^{(a)}, y)$, i.e. $A(\mathbf{0}^{(a)}, y)$ and $(z < y){\sim}A(\mathbf{0}^{(a)}, z)$. We want to show that $y = \mathbf{0}^{(b)}$. By Fact 2, there are three possibilities: $y = \mathbf{0}^{(b)}$, or $y < \mathbf{0}^{(b)}$, or $\mathbf{0}^{(b)} < y$. If $\mathbf{0}^{(b)} < y$, then ${\sim}A(\mathbf{0}^{(a)}, \mathbf{0}^{(b)})$, since ${\sim}A(\mathbf{0}^{(a)}, z)$ for all $z < \mathbf{0}^{(b)}$. So suppose $y < \mathbf{0}^{(b)}$. We know that $B(\mathbf{0}^{(a)}, \mathbf{0}^{(b)})$, and so $(z < \mathbf{0}^{(b)}){\sim}A(\mathbf{0}^{(a)}, z)$. So in particular ${\sim}A(\mathbf{0}^{(a)}, y)$, contradiction. So neither $\mathbf{0}^{(b)} < y$ nor $y < \mathbf{0}^{(b)}$ holds, and so $y = \mathbf{0}^{(b)}$. This reasoning can be carried out formally in Q, as can easily be verified, and so Q fi $(y)B(\mathbf{0}^{(a)}, y) \supset y = \mathbf{0}^{(b)})$. This completes the proof that B represents ϕ in Q.

We can now prove the desired

Theorem: Every partial recursive function is representable in Q (and therefore in any axiom system extending Q).
Proof: For simplicity we only prove the theorem for 1-place functions. Let ϕ be a partial recursive function. Then by the normal form theorem for partial recursive functions, $\phi(x) = U(\mu y R(x, y))$ for some relation R definable in Lim and some U whose graph is definable in Lim. (In fact, of course, we can take U to be K_1.) Then the functions U and $\mu y R(x, y)$ both have graphs definable in Lim, so by Lemma 2, both are representable in Q; by Lemma 1, their composition, which is ϕ, is representable in Q.

Corollary: Every recursive set is strongly representable in every consistent extension of Q.
Proof: Let Γ be some consistent extension of Q, and let S be any recursive set. Let ϕ be S's

characteristic function, and let R(x, y) be some formula which represents ϕ in Q, and therefore in Γ. (Such an R exists by the preceding theorem.) Let B(x) be the formula R(x, $\mathbf{0'}$). If n $\in$ S, then ϕ(n) = 1, so Γ fi R($\mathbf{0}^{(n)}$, $\mathbf{0'}$). If, on the other hand, n $\notin$ S, then ϕ(n) = 0, so Γ fi (y)(R($\mathbf{0}^{(n)}$, y) $\supset$ y = $\mathbf{0}$), so Γ fi ~R($\mathbf{0}^{(n)}$, $\mathbf{0'}$) (since $\mathbf{0} \neq \mathbf{0'}$ is a theorem of Q). So R(x, $\mathbf{0'}$) strongly represents S in Γ.

This corollary extends our previous result: before, we only knew that every set definable in Lim is strongly representable in Q (and therefore in any consistent extension of Q).

We can use our results to prove Rosser's form of Gödel's theorem:

Rosser's Theorem: If Γ is a consistent r.e. extension of Q, then Γ is incomplete.
Proof: We can give two different proofs using results we have proved. The first, closer in spirit to Rosser's is this. Consider the function Φ(x, x). We know that the sets S_1={m: Φ(m, m) = 0}, and S_2={m: Φ(m, m) is defined and > 0} are recursively inseparable. Let A(x, y) be a formula that represents the function Φ(x, x) in Γ. So we have that if Φ(m, m) = 0 then Γ fi A($\mathbf{0}^{(m)}$, $\mathbf{0}$) $\wedge$ (y)(A($\mathbf{0}^{(m)}$, y) $\supset$ y = $\mathbf{0}$); and if Φ(m, m) is defined and =n> 0 then Γ fi A($\mathbf{0}^{(m)}$, $\mathbf{0}^{(n)}$) $\wedge$ (y)(A($\mathbf{0}^{(m)}$, y) $\supset$ y = $\mathbf{0}^{(n)}$). By the second conjunct in the last formula, if Φ(m, m) is defined and > 0, Γ fi ~A($\mathbf{0}^{(m)}$, $\mathbf{0}$). Since Γ is consistent, it is not the case that Γ fi A($\mathbf{0}^{(m)}$, $\mathbf{0}$) and Γ fi ~A($\mathbf{0}^{(m)}$, $\mathbf{0}$). Let R_1={m: Γ fi A($\mathbf{0}^{(m)}$, $\mathbf{0}$)} and R_2={m: Γ fi ~A($\mathbf{0}^{(m)}$, $\mathbf{0}$)}. These are disjoint (since Γ is consistent) and, if Γ were complete, they would be the complement of each other (and so exhaust **N**). They are r.e., for the usual reasons. So if they were the complement of each other, they would be recursive, and then R_1 would be a recursive set that would separate S_1 and S_2, and we prove that no set does that. So we can conclude that Γ is not complete.

A second way of proving the theorem is the following. Suppose, for a contradiction, that Γ is a consistent r.e. extension of Q that is complete. Since Γ is complete, the set of theorems of Γ is recursive. Consider the relation R = {<e, m>: e is a Gödel number of a formula A(x_1), and A($\mathbf{0}^{(m)}$) is a theorem of Γ}. Γ being recursive, R is a recursive relation. Moreover, R enumerates the recursive sets, in the sense that each recursive set is R_e for some e. To see this, let S be a recursive set, and let A(x_1) be a formula that strongly represents it in Γ; then if e is a Gödel number of A(x_1), S = {m: Γ fi A($\mathbf{0}^{(m)}$)} = R_e. So R is a recursive enumeration of the recursive sets. But as we saw in Lecture IX, this is impossible. Therefore, no such Γ can exist, and so any r.e. consistent Γ extending Q is incomplete.

Exercises

1. (a) Prove that an infinite set S of natural numbers is r.e. iff it is the range of a 1-1 total recursive function.

(b) Prove that an infinite set S of natural numbers is recursive iff it is the range of a 1-1 monotone increasing total recursive function.

(c) Prove that every infinite r.e. set has an infinite recursive subset.

2. *Reduction property within Q.* (a) If S_1 and S_2 are two r.e. sets, prove that there are two r.e. sets S_1' and S_2' such that $S_1' \subseteq S_1$, $S_2' \subseteq S_2$, and $S_1' \cup S_2' = S_1 \cup S_2$, such that S_1' is weakly represented by a formula A(x) and S_2' by a formula B(x) and $(x){\sim}(A(x) \wedge B(x))$ is a theorem of Q.

(b) Hence, if two r.e. sets S_1 and S_2 are in fact disjoint, they can be weakly represented by two formulae A(x) and B(x) such that $(x){\sim}(A(x) \wedge B(x))$ is a theorem of Q.

3. (a) Show that the following instance of the naive comprehension scheme is inconsistent: $(\exists y)(x)(x \in y \equiv {\sim}(\exists w)(x \in w \wedge w \in x))$.

(b) Analogous to the construction of K using Russell's paradox, use the result in (a) to obtain a corresponding r.e. set which is not recursive.

(c) Given an r.e. axiom system Γ extending Q, define a number n to be *Gödel-unreciprocated* if m is a Gödel number of a formula $A(x_1)$ and there is no n such that n is the Gödel number of a formula $B(x_1)$ with $A(\mathbf{0}^{(n)})$ and $B(\mathbf{0}^{(m)})$ both provable in Γ. (Otherwise, m is *Gödel-reciprocated*.) Show, analogously to the treatment of 'Gödel-heterological', that the sentence '"Gödel-unreciprocated" is Gödel-unreciprocated' has the properties of the Gödel statement, i.e. it is a Π_1 statement that is true but unprovable if Γ is consistent and not disprovable if Γ is ω-consistent. (Note: this is the Gödelian analog of the paradox of part (a), and is meant to illustrate the theme that set-theoretic paradoxes can be turned into proofs of Gödel's theorem.)

4. (a) Show that there is a recursive function ψ(m,n) such that ψ(m,n) is a code of the n-term sequence all of whose terms are m.

(b) The Upwards Generated Sets Theorem says that if G is a set generated by a recursive basis set and some recursive generating relations such that for each generating relation R, the conclusion of R is greater than or equal to all of the premises, then G is recursive. Prove this theorem. [Hint: prove that every element m of G occurs in a proof sequence for G such that all elements preceding m in the sequence are strictly less than m. Then use (a).]

(c) Use (b) to prove that the set of Gödel numbers of formulae of the narrow first order language of arithmetic is recursive.

(d) Extend this result to any first order language (in the narrow formulation) with finitely many function letters and primitive predicate letters and constants.

5. **Gödel's Theorem via a language with self-reference and extra constants.**

The following is a method of proving the Gödel theorem that directly captures the idea that the Gödel sentence says "I am not provable". It goes by adding additional constants to the narrow first order language of arithmetic; as we have formulated that language, it has only a single constant a_1 standing for zero. We now add all of the others (a_2, a_3, ...) which will denote various numbers. Call the expanded language L*. If we have a set Γ of axioms in L, once we know what we want the extra constants to denote, Γ* will be obtained by adding to Γ all axioms of the form $a_{n+1} = \mathbf{0}^{(m_n)}$, where m_n is the number we want a_{n+1} to denote. (We may not care what certain of the a_{n+1}'s denote, in which case we do not add any axiom involving a_{n+1} to Γ*.) Notice that the language L* and the axiom system Γ* are a mere variant of L and Γ, since all we've done is to add special names for various particular numbers, and nothing can be expressed or proved that couldn't be expressed or proved already.

(a) Use the last remark to prove that if Γ is expanded to an axiom set Γ* with at most one axiom of the given form for each constant, then any proof in Γ* can be transformed into a proof in Γ by replacing each constant by the corresponding numeral and using the axiom (x)(x=x).

(b) Hence, show that every theorem of Γ* becomes a theorem of Γ when constants in the theorem, if any, are replaced by the corresponding numerals. Also show that Γ* is consistent iff Γ is, and that the same holds for ω-consistency.

Now let us make a particular choice of m_n, as follows: if n is a Gödel number of a formula A of L in which x_1 does not occur bound (but in which variables other than x_1 may occur free), let m_n be the least Gödel number of the formula $A(a_{n+1})$ obtained from A by naive substitution of x_1 by a_{n+1} throughout, and include the sentence $a_{n+1} = \mathbf{0}^{(m_n)}$ in Γ*. (Notice that intuitively, if A says something $A(x_1)$, then under our interpretation of the meaning of a_{n+1}, $A(a_{n+1})$ says "I have property $A(x_1)$". Observe that what numbers are the Gödel numbers of a given formula is independent of which interpretation we give to the extra symbols.)

(c) Show that if Γ r.e., then so is Γ* and therefore so is the set of theorems of Γ*.

(d) Show that there is therefore a Π_1 formula $(x_2)B(x_1, x_2)$, where $B(x_1, x_2)$ is a formula of Lim, and which is satisfied by precisely those numbers that are not Gödel numbers of theorems of Γ. We may assume that in this formula x_1 does not occur bound. Let n be the smallest Gödel number of this formula. Assume that Γ extends Q. Prove that if Γ is consistent, then $(x_2)B(a_{n+1}, x_2)$ is true but not provable from Γ*, and therefore that $(x_2)B(\mathbf{0}^{(m_n)}, x_2)$ is also true but unprovable from Γ.

(e) Show that if Γ* is ω-consistent, then $\sim(x_2)B(a_{n+1}, x_2)$ is not provable from Γ* and that if Γ is ω-consistent, $\sim(x_2)B(\mathbf{0}^{(m_n)}, x_2)$ is not provable from Γ.

Remark: (d) and (e) prove Gödel's theorem both for Γ* and for the original system Γ. The point of this exercise is to show that the use of "self-reference" in Gödelian arguments, usually obtained by a rather indirect method, can be obtained by directly constructing a formula of the form "a is not provable", where a is a *name* of the formula itself. Gödel himself may have been under a certain amount of misapprehension about this point. See his *Collected Works,* vol. I (Oxford, 1986), p. 151, n. 15: "Contrary to appearances, such a proposition involves no faulty circularity, for initially it [only] asserts that a certain well-defined formula (namely, the one obtained from the *q*th formula in the lexicographic order by a certain substitution) is unprovable. *Only subsequently (and so to speak by chance) does it turn out that this formula is precisely the one by which the proposition itself was expressed*." (Emphasis added) In the present construction, this is not at all "by chance". On the contrary, we have deliberately set up the denotation so that the formula refers to itself. Nonetheless, there is no "faulty circularity", because the constant a denotes the (smallest) Gödel number of a definite string of symbols, and this number is determined independently of any interpretation of a. We can then assign that number to a as denotation. There are other ways of accomplishing this type of 'direct' self-reference.

(f) In this version of the construction, why are infinitely many constants introduced? Only one constant is used in the undecidable formula.

6. Let ϕ be a uniformization of the relation defined by $W(x,y) \wedge y > 2x$. Let S be the range of ϕ.

(a) Prove that S is r.e.

(b) Prove that S intersects every infinite r.e. set.

(c) Prove that the complement of S is infinite.

(d) Prove that S is neither recursive nor 1-1 complete, citing a previous exercise.

Remark: This is the promised example of an r.e. set that is neither recursive nor 1-1 complete. As I have said, such sets rarely arise in practice unless we are trying to construct them. Later it will be proved that K *is* 1-1 complete.