David Keyt

GÖDEL'S COMPLETENESS THEOREM

0. Stated:
   a. 'sentence' = 'sentence of £'.
      'theorem' = 'theorem of £ by the rules P, T, C, US, UG, and E'.
   b. Every valid sentence is a theorem.
   c. That is, for any given sentence there is either
      (i) a derivation of it from the null set or
      (ii) a counterexample (that is, an interpretation under which it is false).
   d. Since the system is sound, these are exclusive alternatives: there cannot be *both* a derivation of a sentence from the null set *and* a counterexample to it.

1. The general strategy:
   a. The idea behind the proof is to devise a technique that for any given sentence will produce either a derivation of it from the null set or a counterexample to it.
   b. A sentence can be derived from the null set if a contradiction can be derived from its negation.
   c. An interpretation under which the negation of a given sentence is true is an interpretation under which the sentence itself is false.
   d. So it will suffice to find a technique by which either (i) a contradiction can be derived from the negation of a given sentence or (ii) an interpretation can be found under which the negation of the sentence is true.

2. The technique illustrated in two simple cases:
   a. Is '– (∃y)(x)(Fxy ↔ – –Fxx)' a theorem or is there a counterexample to it?

      1. Deny: {1} (1) – – (∃y)(x)(Fxy ↔ – Fxx) P
      2. Put in PNF: {1} (2) (∃y)(x)((Fxy & – Fxx) ∨ (– Fxy & Fxx)) 1, PNF
      3. Systematically remove quantifiers:

         {3} (3) (x)((Fxa_0 & – Fxx) ∨ (– Fxa_0 & Fxx)) P

         {3} (4) (Fa_0a_0 & – Fa_0a_0) ∨ (– Fa_0a_0 & Fa_0a_0) 3 US
      4. If a truth-functional inconsistency appears, enter 'Q & -Q':

         {3} (5) Q & – Q 4 T
      5. Discharge all premisses:

         {1} (6) Q & – Q 2,3,5 ES

         Λ (7) (1) → (Q & – Q) 1,6 C
      6. Infer the original sentence:

         Λ (8) – (∃y)(x)(Fxy) ↔ – Fxx) 7 T

b. Is ‘– (x)(∃y)Fxy’ a theorem or is there a counterexample to it?

1. Deny: – – (x)(∃y)Fxy
2. Put in PNF: (x)(∃y)Fxy   0,1,2,3, . . .
3. Systematically remove quantifiers: (∃y)Fa_0y

   Fa_0a_1

   (∃y)Fa_1y

   Fa_1a_2

   (∃y)Fa_2y

   Fa_2a_3

   .
   .
   .

4. If no truth-functional inconsistency appears, then there is a counterexample in the domain of natural numbers:

D	a_0	a_1	a_2	a_3 . . .	F
ω	0	1	2	3 . . .	{ <0,1>, <1,2>, <2,3>, . . . }
_				_	{ <x,y> | x ∈ ω & y = x + 1 }

3. *The principal metatheorem:* Any sentence that is not a theorem has a counterexample in the domain of natural numbers:
   a. If the sentence under consideration contains *n* individual constants, let these be ‘a_0’, ‘a_1’, ‘a_2’, … ‘a_{n-1}’. Obviously this does not restrict the generality of our proof.
   b. Negate the given sentence and put the negation in prenex normal form. (Lemma: For any given sentence there is an equivalent sentence in prenex normal form [PNF].)
   c. Construct a sequence of sentences in the following way:
      i. Every sentence in the sequence is either a QF or a general sentence, where a QF sentence is one that contains no quantifiers.
      ii. Begin with the sentence in PNF.
      iii. If a sentence of the sequence is of the form (∃α)ϕ, let the next sentence be ϕ α/β where β is the first individual constant in the series ‘a_0’, ‘a_1’, ‘a_2’, … not yet occurring in the sequence.
      iv. If a sentence of the sequence is a QF sentence or begins with a universal quantifier and if β is the first member of the series ‘a_0’, ‘a_1’, ‘a_2’, … for which there is a sentence of the form (α)ϕ not yet followed by ϕ α/β, enter ϕ α/β for the first such universally quantified sentence.

d. *Proof theory:* This sequence has been so constructed that the negation of its first sentence can be derived if a truth-functional inconsistency appears among the QF sentences after some finite number of sentences have been entered:

i. The sentences entered by (ii) and (iii) above are premisses of a derivation; those entered by (iv) are justified by US.

ii. After a truth-functional inconsistency appears, delete the remainder of the sequence and enter:

$(\phi \; \& - \phi)$ by T, where ϕ is a sentential letter.

$(\phi \; \& - \phi$ by ES until each premiss except the first has been discharged.

$(1) \rightarrow (\phi \; \& - \phi)$ by C

$-$ (1) by T

The last line (as well as the one before it) is derived from Λ and hence is a theorem.

e. Suppose, now, that our original sentence is not a theorem.

To show: that it has a counterexample in the domain of natural numbers.

f. It follows from (e) that a truth-functional inconsistency cannot be derived from the PNF of its negation.

g. Hence if the sequence of sentences we constructed is *finite*, the QF sentences it contains are truth-functionally consistent. (The sequence will be finite only if its first sentence contains no universal quantifier.)

h. If it is *infinite*, no initial segment of the sequence contains QF sentences that are truth-functionally inconsistent; for a derivation is a *finite* sequence of sentences. Hence, by the compactness theorem for the sentential calculus, the entire infinite subsequence of QF sentences is truth-functionally consistent.

i. Thus, there is an assignment of truth-values to the atomic constituents of the QF sentences under which all the QF sentences in the sequence are true simultaneously. (Those that are assigned truth under this assignment I shall call 'true'; those assigned falsity, 'false'.)

j. *Model theory:* Construct the following interpretation of £.

i. The domain of interpretation is the set of natural numbers.

ii. Assign to each individual constant the number named by its subscript: 0 to 'a_0', 1 to 'a_1', 2 to 'a_2', and so forth.

iii. The sentential letters occurring in our original sentence have already been assigned truth-values.

iv. Assign to each predicate of degree 1 that set whose members are exactly those numbers named by the subscripts of the individual constants in the true atomic sentences that contain the predicate.

v. Assign to each predicate of degree *n*, where $n > 1$, a set of *n*-tuples, one *n*-tuple for each true atomic sentence that contains the predicate. The *n*-tuple corresponding to the sentence contains as elements exactly the numbers named by the subscripts of the individual constants in the sentence. The k^{th} element of the *n*-tuple is the number named by the subscript of the k^{th} individual constant in the atomic sentence. (Thus one of the triples in the set assigned to 'F^3' when '$F^3a_0a_4a_7$' is true is <0,4,7>.)

k. All of the sentences in our sequence are true under this interpretation:
   i. The proof is by weak induction on the number of quantifiers in the sentence.
   ii. *Basis:* All sentences in the sequence with 0 quantifiers are true since they are QF sentences.
   iii. *Inductive step:*
      (1) Suppose that all sentences in the sequence with *n* quantifiers are true.
      (2) Suppose that $(\exists\alpha)\phi$ begins with $n + 1$ quantifiers. Then $(\exists\alpha)\phi$ is immediately succeeded in the sequence by $\phi\alpha/\beta$, which contains *n* quantifiers and hence, by the hypothesis of the induction, is true. So $(\exists\alpha)\phi$ is true.
      (3) Suppose that $(\alpha)\phi$ begins with $n + 1$ quantifiers. Then the sequence contains an infinite subsequence of sentences of the form $\phi\alpha/\beta$, where β is in turn 'a_0', 'a_1', 'a_2', … There is one such sentence for every member of the domain, and each is true by the hypothesis of the induction since it contains *n* quantifiers. So $(\alpha)\phi$ is true.
      (4) So all sentences with $n + 1$ quantifiers are true.
   iv. Thus all sentences in the sequence are true under the given interpretation.

l. The first member of the sequence is a sentence in PNF that is equivalent to the negation of our original sentence. Since the rules of derivation of £ are sound and since the first member of the sequence is true under the given interpretation, the negation of our original sentence must also be true under this interpretation.

m. So our original sentence has a counterexample in the domain of natural numbers.

4. Corollary: Every valid sentence is a theorem:
   a. Suppose that ϕ is valid.
   b. Then there is no interpretation under which ϕ is false.
   c. Hence there is no interpretation in the domain of natural numbers under which ϕ is false.
   d. So, by the principal metatheorem, ϕ is a theorem.

5. Corollary: Every sentence that is valid in the domain of natural numbers is universally valid (Löwenheim-Skolem theorem):
   a. Suppose that ϕ is valid in the domain of natural numbers.
   b. Then there is no interpretation in the domain of natural numbers under which ϕ is false.
   c. So, by the principal metatheorem, ϕ is a theorem.
   d. Therefore, (since the rules of derivation for £ are sound) ϕ is universally valid.

6. An equivalent formulation of the Löwenheim-Skolem theorem: Every sentence that is satisfiable is satisfiable in the domain of natural numbers.