TWO THEOREMS ON EXPERIMENTAL LOGICS

by

R. J. Jeroslow

Research Report 73-4

February, 1973

Abstract

A generalization of a formal system is considered, in which the axioms of the formal system can be withdrawn or supplemented, as mechanical experimentation proceeds through "time" and the consequences of various combinations of assumptions are realized. The "theorems" of these <u>experimental logics</u> are taken to be those assertions possessing a proof which remains valid for all "sufficiently large time."

Under very broad hypotheses on experimental logics, we obtain the following two results: (1) There is a (n infinitude of) true but unprovable Π_1^0 - sentence (s); (2) There is no mechanical procedure for uniformly finding any (of the infinitude of) true but unprovable Π_1^0 -sentence(s).

Our first result is analogous to Godel's First Incompleteness Theorem for formal systems (which it implies). Our second result differs sharply with that for formal systems, where a true but unprovable Π_1^0--sentence is mechanically obtained in the course of Godel's proof.

TWO THEOREMS ON EXPERIMENTAL LOGICS

by R. G. Jeroslow[1]

If one experiments through time with mechanical processes, and, on the basis of the outcome of these "computer experiments," one revises the axioms of a mathematical system, then the resulting time-dependent deductive system is called an experimental logic.

This paper initiates the study of experimental logics and gives two results concerning them. The first result states that, under rather broad hypotheses, there will be a true Π_1 sentence not provable in the experimental logic. This result is analogous to Godel's fundamental result [3], where the true Π_1 sentence not provable in a given (consistent) formal system happens also to be the system's own Consistency statement, a statement that can be constructed uniformly from a mechanical description of the formal system*

Our second result states that, even for a very "well-behaved" class of experimental logics, there does not exist any uniform mechanical procedure for finding any one of the infinitude of non-provable but true Π_1 sentences,, A fortiori, no uniform procedure exists for any extension of this "well-behaved" class of experimental logics*

Section 1 develops the basic concepts for experimental logics and gives examples of them. Section 2 presents the results of the paper.

Section 1: Definitions

An experimental logic can be identified with a computable predicate H(t,x,y) from some decidable class of computable predicates, with the intended intuitive interpretation:

---

[1] This research has been partially supported by NSF grant GP21067. The results of this paper were first reported in a working paper, "On Godel's Consistency Theorem," July, 1971.

(1) H(t,x,y) = at time t, the finite configuration with godel number y is sufficient grounds for asserting the formula with godel number x.

To make our discussion precise, let us limit the class of predicates H to the primitive recursive predicates.

The <u>experimental logic</u> at time t, designated H_t, has as its "theorems" { x | (3 y)H(t,x,y)}. It is not required that H_t be closed under deductions in the predicate calculus; it may be finite. The formulae which recur infinitely often in H_t are called the <u>recurring formulae,</u> and they are defined by the condition:

(2) Rec_H(x) = (Vt) (3 s ≥t) <^y) H(s,x,y).

Of special interest in the study of experimental logics are the <u>stable formulae,</u> by which we mean the formulae cp which, at some point of time,become provable via some proof II for all remaining time. I.e., not only is cp a theorem of H_t for all sufficiently large t, but there is a proof of cp which is never "subject to question" for all sufficiently large t. Stable formulae are defined by the condition:

(3) . Stbl(x) = (3t) (3y) (V$ £t) H(s,x,y).

By |^ cp (or, when H is clear from the discussion, by |»cp) we shall simply mean that cp is stable in the experimental logic defined by the predicate H(t,x,y): we shall also use the terminology that cp is a theorem of H, or is provable in H.

For furthur reference, we note that Rec (x) is a II_2 formula and Stbl(x) is a L_2 formula (see Chapter 13 of [7] for terminology). We shall use $\bar{cp}$ to denote the godel number of the formula cp.

It is characteristic of an experimental logic that one cannot detect its theorems in a mechanical manner, or else the experimental logic usually is equivalent to an ordinary formal logic, with axioms and rules of reasoning specified in advance.

In experimental logics, the implicit trust of the user of the logic lies <u>not</u> in the specific theorems which happen to be in H_t at some specific point t, but in the <u>method</u> for accepting and disgarding theorems which is described by H(t,x,y).

If this trust is to entail a belief that this method "converges" on some stable conceptual framework, be it true or false, then certainly this trust requires that the stable and recurring formulae are identical. When this latter condition is met, we say that the experimental logic converges.

If the experimental logic is to embody trust in the deductive method as a means of obtaining new conclusions, then certainly any purely logical deduction from a set of theorems valid in H_t for all sufficiently large t should eventually occur in H_s for sufficiently large s. This implies that the set of stable formulae shall be closed under deductive reasoning.

The discussion of the last two paragraphs is intended only to motivate our choice of hypotheses in Theorem 1 below.

An interesting subclass of the experimental logics are the finite extension experimental logics. Here some formal theory $\mathcal{T}$ is fixed, and as some primitive recursive procedure computes through time, a finite upper bound B is found. This bound insures that no more than B "additional axioms," i.e., axioms other than those in $\mathcal{T}$, will ever be considered. As the computations proceed, these axioms $\varphi_1, \ldots\ldots, \varphi_p$, with $p \leq B$, are generated (perhaps $p = 0$), and, as these axioms occur they are ranked (as to "desirability") relative to those already occuring, this ranking to be fixed for all future time.

As time proceeds, deductions from and the "currently active" additional axioms become longer and longer. If a contradiction is reached depending only on the axioms of $\mathcal{T}$, nothing is done. If a contradiction occurs in which some of the "currently active" additional axioms are used, then the least-ranked of these "currently active" addition axioms is "dropped" and becomes "inactive." This "dropped" axiom is called the "current error." Depending upon how one wishes to arrange things, one may simply consider continuing with the remaining "currently active" additional axioms, or one may follow the following alternate route. One may examine all previous proofs which led to contradictions and used the "current error," and then return to use, as "currently active," any lesser-ranked additional axiom (aside from the "current error") which was dropped when the earlier contradiction was found.

The finite extension experimental logics are clearly closed under predicate reasoning. They are also convergent. In fact, eventually there occurs an additional axiom which is highest-ranked among any additional axioms which ever occur. If this highest-ranked axiom is ever dropped, it can never be returned. Hence, after some finite point in time, this highest-ranked axiom is either always an axiom of H_t or always not an axiom of H_t. Arguing from this point in time on, by a downward induction on the ranking of the additional axioms which occur, one easily shows the existence of a point to in time such that, for $t \geq t_0$, the axioms of H_t are all identical. Then the formulae that stabilize in H are those which can be deduced from the axioms of H_t for any $t \geq t_0$, and these are also the formulae that recur in H.

It should be pointed out that, although finite extension logics permit only the addition of finitely many axioms, these can be from any of the formulae of the predicate calculus, and need not be in the language of the "base theory" $\mathcal{J}$. For instance, $\mathcal{J}$ might be elementary number theory, while one of the additional axioms is Godel-Bernays set theory. Therefore, while the bound B on the number of additional axioms is a restriction, it is not as severe by any means as restricting H to an extension of $\mathcal{J}$ which is finitely-axiomatizable over in the language of $\mathcal{J}$.

Slightly more general than the finite extension logics are the <u>ascending</u> experimental logics, in which, in addition to a finite bound B for a certain set of "additional axioms," one also considers adding a set of additional axioms of strictly increasing logical strength, where one mechanically knows the relative strengths. This "hierarchy of axioms" can be of recursive order type . Here, complex rules for adding in and taking out axioms in this hierarchy of axioms may be added, in addition to the simple "inconsistency tests" of the finite extension experimental logics. If one then knows convergence for the additional rules, then it is automatic for the overall ascending logic, because, once a single axiom from this hierarchy of axioms is dropped on account of an inconsistency, all more powerful axioms of the hierarchy are to be automatically dropped with it. This allows one to use the kind of simple reasoning appropriate to finite extension logics, in order to show that ascending logics will also converge.

It is a tedious technical exercise to write down the details of finite extension logics and ascending logics in terms of suitable primitive recursive predicates H(t,x,y), and those who have done any of this kind of exercise with primitive recursive calculations see immediately that it can be done. Other readers may wish to consult [1], or [6], or [7], to see the kind of techniques which are involved; we omit details, since they add nothing to our discussion.

It is worth remarking that consistent experimental logics can often prove their own consistency. In fact, the finite extension logic consisting of Peano $\mathcal{P}$ plus the additional statement "Con" asserting the consistency of Peano, proves its own consistency. For clearly $\mathcal{P} \cup \{Con\}$ is consistent, so the theorems of this experimental logic H are the theorems of the formal system $\mathcal{P} \cup \{Con\}$. Now in $\mathcal{P} \cup \{Con\}$ one can reason that, if there is an inconsistency in $\mathcal{P} \cup \{Con\}$, then H is "cut down" to $\mathcal{P}$, which is consistent. Therefore, for this H we have $\vdash_H Con_H$. Note, however, that for experimental logics H, Con_H is a π_2 and not a π_1 sentence.

We particularly wish to acknowledge the influence of the proof of Theorem 5.9 of Professor Feferman's paper [2] in the development of the concept of an experimental logic.

Section 2: Results

If one utilizes only the tools of recursion theory, which are known to be sufficient to obtain Godel's First Incompleteness Theorem (see particularly [8]), the requirement that an experimental logic converge simply places its theorems in $\pi_2 \cap \Sigma_2 = \Delta_2$, and a priori these theorems might include all true π_1 -sentences. However, by utilizing the reasoning formalized in deductive theories, one can show that this latter eventuality cannot arise. This is the essential content of our first result.

Theorem 1: Let H be a consistent, convergent experimental logic whose theorems contain those of first-order Peano arithmetic and whose theorems are closed under first-order classical predicate reasoning.

Then there is a true π_1 sentence which is not provable in H.

Proof: It is useful to remark that, if $(\exists x)\ (\forall y)\ R(x,y)$ is a true but unprovable Σ_2 sentence, then, for some n , $(\forall y)R(\ulcorner n \urcorner, y)$ is a true but unprovable π_1 sentence; so we need only find a true but unprovable Σ_2 sentence.

Via standard techniques, one easily obtains a formula which represents the assertion, " I do not recur," i.e., for which we have

(4) $\vdash \mathrm{Rec}\ (\bar{\varphi}) \iff \neg\varphi.$

(See, for instance, Lemma 5.1 of [2]). Note is (equivalent in classical predicate reasoning to) a Σ_2 sentence.

The proof now divides into two cases, and in either case we find a true but unprovable Σ_2 sentence.

Case 1: $\vdash \mathrm{Rec}(\bar{\varphi}) \longrightarrow \varphi$.

Then since (4) gives $\vdash \mathrm{Rec}(\bar{\varphi}) \longrightarrow \neg\varphi$, we obtain $\vdash \neg\ \mathrm{Rec}(\bar{\varphi})$, and hence by (4), we also obtain $\vdash \varphi$. Therefore Stbl $(\bar{\varphi})$ is a true Σ_2 sentence. It suffices to show that Stbl $(\bar{\varphi})$ is not provable. In fact, if $\vdash$ Stbl $(\bar{\varphi})$, since $\vdash$ Stbl $(\bar{\varphi}) \longrightarrow \mathrm{Rec}(\bar{\varphi})$, we would then have $\vdash \mathrm{Rec}(\bar{\varphi})$. In this case, H would be inconsistent, contradicting our hypothesis.

Case 2: We do not have $\vdash \mathrm{Rec}(\bar{\varphi}) \longrightarrow \varphi$.

In the classical predicate calculus, $\mathrm{Rec}(\bar{\varphi}) \longrightarrow \varphi$ is equivalent to a Σ_2 sentence. By the hypothesis of this case, it suffices to show that $\mathrm{Rec}(\bar{\varphi}) \longrightarrow \varphi$ is true.

However, if $\mathrm{Rec}(\bar{\varphi})$ is true, since H is convergent, Stbl $(\bar{\varphi})$ is also true, so $\vdash \varphi$. But then predicate reasoning gives $\vdash \mathrm{Rec}(\bar{\varphi}) \longrightarrow \varphi$, contradicting the case. Hence, Rec $(\bar{\varphi})$ is false; but then $\mathrm{Rec}_H(\bar{\varphi}) \longrightarrow \varphi$ is automatically true. Q.E.D.

Remark 1: We mention in passing some alternate sets of hypotheses which lead to the conclusion of Theorem 1, i.e., the existence of a true but unprovable π_1- sentence.

The hypothesis that all provable Σ_2-sentences are true (in the standard model) and that the experimental logic H is convergent, are by these themselves sufficient. Here one obtains a purely recursion-theoretic proof: (1) If all true Σ_2- sentences were provable, then by hypothesis the set of true Σ_2- sentences would be equivalent to the set of provable ones; hence (2) The complete Σ_2- set would be reducible to the set of theorems of H; but (3) The theorems of a convergent logic are a Δ_2- set; and hence (4) The assumption (1) is false, so that there is a true but unprovble Σ_1- sentence, and therefore a true but unprovable π_1- sentence.

However, the main hypothesis in the proof given in the last paragraph, i.e., that all provable Σ_2-sentences are true, however reasonable it may be when a formal system is given which is believed to embody "perfect knowledge," seems a particularly strong assumption when one is experimenting to try to obtain true theorems. After all, no finite set of mechanical trials is usually sufficient to detect the falsity of a Σ_2-assertion which is in fact false.

It is also possible to obtain the conclusion of Theorem 1 under these hypotheses on H: H is consistent, contains the theorems of Peano arithmetic, is closed under predicate reasoning, and moreover

(5) $\vdash \text{Stbl}(x) \longleftrightarrow \text{Rec}(x)$

with x a free variable. To see this, one simply repeats the proof of Theorem 1, but uses instead the sentence which represents the assertion ψ, "I do not stabilize." Then Case 1 is quite easy and Case 2 is that $\text{Stbl}(\bar{\Psi}) \rightarrow \Psi$ is not provable, although in this case it is clearly true. It remains only to show that $\text{Stbl}(\bar{\psi}) \longrightarrow \psi$ is equivalent to a Σ_2 sentence, and that is done by using (5) to change π_2- forms into Σ_2-forms whenever one needs to.

What is to be noted about <u>both</u> Theorem 1 and the result of the last paragraph is the fact that they do <u>not</u> use conditions such as:

(6) if $\vdash \varphi$, then $\vdash \text{Stbl}(\bar{\varphi})$.

For one thing (6) does not seem to be useful in obtaining Theorem 1. The condition (6) is, we feel, rather strong for experimental logics.

Of course, with $\text{Pr}(\bar{\varphi})$ (" φ is provable ") replacing $\text{Stbl}(\bar{\varphi})$, the condition (6) provides a primary technical result on which the usual proofs of Godel's First Incompleteness Theorem, for formal systems, are based. Conditions like (6) can be circumvented in the proof of Godel's result, by studying instead sentences of the form $\text{Pr}(\bar{\varphi}) \longrightarrow \varphi$, and our Theorem 1 is actually based upon exploiting this observation in experimental logics (see particularly 3.232 of [5], where this observation is implicit).

Remark 2: Our metamathematics in the proof of Theorem 1 is not intuitionistically acceptable; it utilizes a tertium non datur in the form of Case 1 and Case 2. However, our proof is formalizable in (classical) Peano arithmetic. Hence, it is not refutable even in the intuitionist theory of species, since this theory is a subtheory of classical analysis, the latter being a consistent extension of Peano arithmetic.

Remark 3: Our emphasis on true but unprovable π_1-sentences, as apart from an emphasis merely on any true but unprovable sentences, derives from Hilbert's own emphasis on this very special class of sentences (see particularly [4]). Hilbert found such sentences $(\forall x)P(x)$, with P a testable predicate, to be of particular significance, since they immediately give rise to the infinitude of "testable predictions," $P(\ulcorner 0 \urcorner)$, $P(\ulcorner 1 \urcorner)$, $P(\ulcorner 2 \urcorner)$,... etc. These sentences were, in this sense, "meaningful". As to the "meaning" of the rest of mathematical sentences, Hilbert questioned this notion as being relevant even to sentences as simple as $(\exists x)P(x)$, with the unbounded numerical existential quantifier (see [4]).

We say that a logic is 1-consistent if, whenever $(\exists x)P(x)$ is provable, with $P(x)$ primitive recursive, we do not have $\neg P(\ulcorner n \urcorner)$ provable for each numeral $\ulcorner n \urcorner$. For systems extending (actually small subsystems of) Peano arithmetic, this is equivalent to: whenever $(\exists x)P(x)$ is provable, then $P(\ulcorner n \urcorner)$ is provable for some numeral n.

An experimental logic H is said to be deductive if H_t is closed under (classical) predicate reasoning for each t.

Let K be the class of experimental logics possessing the following properties:

(i) H is 1-consistent.

(ii) H is deductive, and each H_t contains first-order Peano arithmetic.

(iii) (5) holds for $\vdash$ being $\vdash_H$.

(iv) H is a finite extension experimental logic.

(v) H proves its own consistency.

(vi) H is convergent.

As mentioned in the introduction, the class K is not chosen for any particular reason, other than that of defining a "very well-behaved" class of experimental logics. In the proof of Theorem 2 below, only property (i) for H e K, will play an essential technical role.

<u>Theorem 2:</u> There is no partial recursive function f having the following property:

If H e K, and q is the godel number of the primitive recursive predicate H(t,x,y), then f(q) converges and is the godel number of some rr_1 -sentence not provable in H.

<u>Proof:</u> Our discussion in Section 1 provided examples of H e K; i.e., K ^ 0. Fix any H e K for the discussion to follow.

Suppose, for the sake of contradiction, that f exists. Then the following partial recursive function g also exists. Given a number x, it is first determined whether or not x is the godel number $x = \bar{T}$ of a primitive recursive predicate of four free variables T(a,t,x,y). Then one first computes the godel number u of T($\bar{T}$, t,x,y), and one finally sets g(x) = f(u) .

Since g is partial recursive, its graph is recursively enumerable, and hence there exists a primitive recursive predicate G(a,r,w) of three free variables such that g(a) = r if and only if (5 w)G(a,r,w).

Let T(a,t,x,y) be the following primitive recursive predicate:

<u>"Case 1:</u> ($_3$ w $\leq$ t) (g r $\leq$ t) G(a,r,w)#

<u>Subcase 1.1</u>: There is a proof of godel number $\leq$ t, in H_t plus the axiom with godel number r, of an inconsistency.

In this subcase 1.1, T(a,t,x,y) is H(t,x,y).

<u>Subcase 1.2:</u> Case 1 holds; but subcase 1.1 fails. In this subcase, T(a,t,x,y) is the condition that y is a proof of x in H_t except for possibly some (finite number of) occurences of the formula with godel number r, used as an axiom.

<u>Case 2;</u> Case 1 fails

Then $T(a,t,x,y)$ is $H(t,x,y)$."

We first make the following claim:

(8) If $H^f(t,x,y) \equiv T(\bar{T}, t,x,y)$, then $H^f \in K$.

Note that (ii) and (iv) are immediate for H^f.

As to (vi), we reason as follows. Suppose that subcase 1.1, as described in the definition of $T(a,t,x,y)$, arises for infinitely many t. Then by hypothesis (ii) on H, for infinitely many t, $\neg(\forall x)P(x)$ is provable in H_t, where $(\forall x)P(x)$ is the Π_1-sentence with godel number r, cited in Case 1. Then hypothesis (vi) on H insures that $\neg(\forall x)P(x)$ stabilizes in H, and hence, that subcase 1.1 occurs for all sufficiently large t. Then, for sufficiently large t, H_t and H_t' are identical, and the hypothesis (vi) for H gives (vi) for H^1. Otherwise, subcase 1.1 does not arise infinitely often, so that there is a last point in time beyond which it arises. If Case 1 is then ever to arise for any t there is a point in time such that, at all later points in time, subcase 1.2 holds, and then H^1 is identical with what can be proven from H with the additional Π_1 - sentence $(\forall x)P(x)$ whose godel number is r, so H^f satisfies (vi). If Case 1 never arises, then H_t and H_t' are identical for all t, so the hypothesis (vi) for H gives (vi) for H^f. This proves (vi).

Note that the proof of the last paragraph, i.e., the proof that (vi) holds for H^1, can be proven in Peano arithmetic plus (5); hence (iii) holds for H^f, by virtue of the hypotheses (ii) and (iii) for H.

The verification of (v) for $H^!$ is trivial, using the fact that (v) holds for H, and the kind of reasoning exemplified in Section 1 for the experimental logic(?u{Con}.

Only the proof of (i) remains, but this is trivial, if one utilizes the same kind of analysis that we used to prove (vi) for H^f. For if subcase 1.2 arises from some point of time on, then by hypothesis (ii) on H, the Π_1-sentence with godel number r must be true' and it is a well-known result (an an easy exercise) that the addition of a true π_1- sentence to a 1-consistent logic extending Peano gives a 1-consistent logic. Otherwise, subcase 1.2 does not arise after some point in time, in which case, for sufficiently large t, H_t and H_t' are identical; and then the hypothesis (i) for H gives (i) for H'.

The claim (8) is established. Therefore, by the definition of the function g, $g(\bar{T})$ is the godel number of a true π_1-sentence $(\forall x)P(x)$ not provable in H'. This implies that, for sufficiently large t, $(\forall x)P(x)$ is inconsistent in the logic H_t, and hence that for sufficiently large t, $(\exists x)\neg P(x)$ is a theorem of H_t. Therefore, $\vdash_H (\exists x) \neg P(x)$, and so, by hypothesis (i) on H, $\vdash_H \neg P(\ulcorner n \urcorner)$ for some n. Then by hypotheses (i) and (ii) on H, $\vdash \neg P(\ulcorner n \urcorner)$ is true. Now we have the desired contradiction, since $(\forall x)P(x)$ is also true.

Therefore, f does not exist. Q. E.D.

Remark 4: Since known formalizations of constructive mathematics prove the existence of a function or partial function possessing a property only if a recursive or partial recursive function exists with that property, Theorem 2 makes it seem unlikely that Theorem 1 can be proven constructively. Therefore, in view of Remark 2, the results of this paper will probably remain independent over any formalism for constructive mathematics.

The Department of Mathematics
Carnegie-Mellon University
January 20, 1973

REFERENCES

1. M. Davis, Computability and Unsolvability, McGraw-Hill, New York, 1958. 210 pp.

2. S. Feferman, "The Arithmetization of Metamathematics in a General Setting," Fundamenta Mathematicae XLIV (1960). pp. 35-92.

3. K. Gödel, "Ueber formal unentscheidbare Satze der Principia Mathematica und verwanter Systeme," Monatshefte fur Mathematik und Physik 38 (1931), pp. 173-198.

4. D. Hilbert, "On the Infirfite," in From Frege to Godel, J. van Heijenoort, ed., Harvard University Press, Cambridge, 1967,. 660 pgs.

5. G. Kreisel, "Mathematical Logic," in Lectures on Modern Mathematics, vol. 3, T. L. Saaty, ed., Wiley and Sons, New York, 1963-65.

6. R. Péter, Recursive Functions, Academic Press, New York, 1967, 300 pp.

7. H., Rogers, Jr., Theory of Recursive Functions and Effective Computability, McGraw-Hill, New York, 1967. 482 pp.

8. R. Smullyan, Theory of Formal Systems, Princeton University Press, Princeton, New Jersey, 1961. 142 pgs.