3

The Paradox of Analysis*

18. The Paradox

The paradox of analysis is an important and complex problem in the philosophy of logic and language. What makes it important are its deep implications for philosophy in the areas of philosophical methodology and philosophical psychology and for psychology in the areas of development, perception, decision, and perhaps psychoanalysis. Yet in recent years philosophers have all but forgotten the problem.

When I speak of the paradox of analysis I am referring to logical puzzles of the following sort. Take three formulas:

(1) x knows that whatever is a circle is a circle.
(2) x does not know that whatever is a circle is a locus of points in the same plane equidistant from some common point.
(3) Being a circle = being a locus of points in the same plane equidistant from some common point.

The paradox is this: (1) and (2) are simultaneously satisfiable; (3) is true, and yet the conjunction of (1) and (3) entails the negation of (2). Hence, a contradiction.[1]

To get the paradox of analysis squarely in mind, one must distinguish it from a superficially similar yet fundamentally different problem which I call *Mates' puzzle*.[2] Mates' puzzle is generated in an analogous way by formulas such as the following:

(4) x knows that whatever chews chews.
(5) x does not know that whatever masticates chews.
(6) Being something that masticates = being something that chews.

* The reader may skip this chapter without interrupting the larger line of development in the book.

It would seem that (4) and (5) can be satisfied at the same time, for even though x knows what chewing is, he can in some sense fail to know what masticating is. Such a situation might arise as follows. Our person x has the concept of chewing (and *eo ipso* the concept of masticating). In addition, x knows that the predicate 'chew' expresses this concept. However, x does not know what the predicate 'masticate' expresses. Indeed, x's contact with the historical information chain associated with the English predicate 'masticate' does not go beyond his knowledge that English speakers have such a chain. x's ignorance as reported in (5) originates in his ignorance of linguistic (or historical or social) matters.[3] In the case of a genuine paradox of analysis, by contrast, the sort of ignorance at work does not originate in linguistic (or historical or social) ignorance. In the above instance of the paradox of analysis, for example, we can imagine that, besides being fully aware of what circularity is, x also knows that the English word 'circle' expresses circularity. What x is ignorant of is neither the concept of circularity nor the semantics of his language nor the relation between historical information chains. Rather, x is ignorant of the *definition* of circularity itself. Herein lies the difference between the genuine paradox of analysis and Mates' puzzle. The challenge posed by the paradox of analysis then is to find a satisfactory way to represent the non-linguistic (non-historical, non-social) knowledge that one acquires when one learns a definition. This is no trvial affair.

A number of people have proposed informal resolutions to the paradox of analysis. But in my view none of these informal resolutions promises to be adequate. In the field of formal intensional logic the paradox of analysis has been virtually ignored. Alonzo Church is the only logician to have incorporated into a formal theory of propositions a serious attempt to resolve the paradox.[4] And so it is to this resolution that I now turn.

19. Difficulties in Church's Resolution

I begin this section by sketching Church's resolution. Then I will show why it too is inadequate. One purpose for engaging in this exercise is to gain a better understanding of what is required of a resolution of the paradox.

To understand Church's resolution of the paradox of analysis, one must be familiar with his theory of synonymy, which he first states in 'Intensional Isomorphism and the Identity of Belief' (pp. 66–7).[5] On Church's theory, expressions are synonymous if and only if they are synonymous isomorphic. And expressions are *synonymous isomorphic* iff_{df} one can be obtained from the other by a series of steps that consists of (1) alphabetic changes of a bound variable, (2) replacement of one individual constant by another that is synonymous with it, (3) replacement of one predicate constant by another that is synonymous to it, (4) replacement of an abstraction expression—i.e., an expression of the form $(\lambda x)(\ldots x \ldots)$—by a synonymous predicate constant, (5) replacement of a predicate constant by a synonymous abstraction expression, (6) replacement of an individual description by a synonymous individual constant, (7) replacement of an individual constant by a synonymous individual description.

If correct, Church's theory of synonymy yields a resolution of "one half" of the instances of the paradox of analysis. To see how this works, consider the instance of the paradox of analysis generated in the usual way by the following formulas:

(7) x knows that, for all y and z, if $y = z$, then $y = z$.

(8) x does not know that, for all y and z, if every property of y is a property of z and conversely, then $y = z$.

(9) Being y and z such that $y = z$ is the same as being y and z such that every property of y is a property of z and conversely.[6]

Church would resolve this instance of the paradox by denying (9). Since '$y = z$' cannot be converted by rules (1)–(7) into 'Every property of y is a property of z and conversely', they are not synonymous isomorphic.[7] So on Church's theory of synonymy they are not synonymous, and therefore, the concepts they express are different. Thus, (9) is false, and the contradiction is avoided.

But consider the following formulas adapted from an example discussed at length by Church:[8]

(10) x knows that whatever is a period lasting fourteen days is a period lasting fourteen days.

(11) x does not know whatever is a fortnight is a period lasting fourteen days.

(12) Being a fortnight = being a period lasting fourteen days.

These three formulas give rise to a contradiction in the usual way. Since 'x is a fortnight' can be converted into 'x is a period lasting fourteen days' by rule (5), these two expressions are synonymous isomorphic. Thus, on Church's theory of synonymy they are synonymous. So they express the same concept, and hence (12) is true. At the same time, given Church's axioms for the logic of sense and denotation,[9] the negation of (11) is derivable from the conjunction of (10) and (12). Thus, Church must deny that (10) and (11) are simultaneously satisfiable. This is what he does. But since we do have an intuition that (10) and (11) are simultaneously satisfiable, an explanation of why it is mistaken ought to be given. Church's explanation goes as follows.[10] Suppose that (10) is satisfied. Then the possibility that we take to be expressed by (11) is not in fact expressed by (11). Rather, it is expressed by the following formula containing metalinguistic terminology:

(11′) *x* does not know that whatever satisfies the English sentential matrix '*y* is a fortnight' satisfies the English sentential matrix '*y* is a period lasting fourteen days'.

Clearly, (10) and (11′) are simultaneously satisfiable. Indeed, (10), (11′), and (12) are jointly consistent. Extrapolating from the foregoing, I then arrive at the following conclusion. If there are instances of the paradox of analysis whose key expressions are synonymous isomorphic, Church is committed to resolving these instances by means of the above metalinguistic maneuver. We shall see below that there indeed are such instances of the paradox. And so it is that Church is led to resolve the "second half" of the instances of the paradox by means of the metalinguistic maneuver.

Now I turn to the assessment of Church's two types of resolutions of the paradox of analysis. The first type of resolution is inadequate because the theory of synonymy is mistaken. This can be shown in two ways. The first is simply by example. Intuitively, '*y* outweighs *z*' and 'The weight of *y* is greater than the weight of *z*' are synonymous; however, on Church's theory they could not be since they are not synonymous isomorphic.[11] This is not just an isolated counterexample; examples of this sort abound.[12] The second way to show the inadequacy of Church's theory is this. By radically limiting the synonym pairs, the theory acts as a two-edged sword: although in its way it does resolve certain instances of the paradox of analysis, it has the undesirable effect of artificially shrinking the logical-consequence relation for propositions.[13] To see this, consider the following two formulas:

> x believes that, if the weights of y and z differ and if y does not outweigh z, then z outweighs y.
>
> x believes that, if the weights of y and z differ and if y does not outweigh z, then the weight of z is greater than the weight of y.

There are readings of these formulas (perhaps the most natural readings) according to which the proposition expressed by the latter formula is a logical consequence of the proposition expressed by the former formula. However, Church's theory of synonymy makes these readings impossible. For this and the preceding reason I conclude that Church's theory is mistaken. After all, the only motivation for Church's theory appears to be that in its way it can help to resolve the paradox of analysis.

The major flaw in Church's second type of resolution, his metalinguistic resolution, is that it in effect rules out the possibility of *informative definitions* beyond those that concern mere linguistic facts. That is, it is a consequence of the metalinguistic resolution that what one learns when one discovers a correct definition is merely something about a language. Yet in view of Church's famous criticism of Carnap's analysis of statements of assertion and belief, it is surprising that Church should take this line, for a criticism akin to Church's criticism of Carnap can now be lodged against Church. To illustrate this criticism, let us consider the instance of the paradox of analysis given at the outset of the chapter. On analogy with 'y is a fortnight' and 'y is a period lasting fourteen days', the expressions 'y is a circle' and 'y is a locus of points in the same plane equidistant from some common point' are synonymous isomorphic. Hence, Church would hold that (3) is true. Therefore, on analogy with his treatment of the earlier 'fortnight' case, Church would be led to deny that (1) and (2) are simultaneously satisfiable. And in turn he would be led to hold that, if (1) is satisfied, then the possibility that we mistakenly take to be expressed by (2) is actually expressed by:

(2′) x does not know that whatever satisfies the English sentential matrix 'y is a circle' satisfies the English sentential matrix 'y is a locus of points in the same plane equidistant from some common point'.

Since (1), (2′), and (3) are jointly consistent, Church would hold that the paradox is resolved. However, let us consider a deaf-mute x

whom we know to know no English (and, we may suppose, no other language either). Suppose that we observe *x* sorting out circular objects from non-circular ones. On the basis of this and related evidence we infer that *x* has the concept of circularity. In turn, we infer the proposition expressed by (1). Now suppose I ask you, 'Does *x* know that whatever is a circle is a locus of points in the same plane equidistant from some common point?'. Shortly thereafter we observe *x* performing a variety of relevant geometric constructions with ruler and compass. On the basis of these observations, I make the inference that prompts me to assert, '*x* does know that whatever is a circle is a locus of points in the same plane equidistant from some common point'. Now according to Church's view, I have asserted the proposition expressed by the following metalinguistic formula:

> *x* knows that whatever satisfies the English sentential matrix '*y* is a circle' satisfies the English sentential matrix '*y* is a locus of points in the same plane equidistant from some common point'.

But this does not seem credible. I have not made an assertion about *x*'s knowledge of English; indeed, I know that *x* knows no English. Nor have I (*à la* Carnap) made an assertion about *x*'s knowledge of some other language somehow related to English; we may suppose that *x* knows no language at all. When people (regardless of which language, if any, they speak) learn that circles are loci of points in the same plane equidistant from some common point, they learn a fact about circles, not about language.[14] Here again we see that ignorance of conceptual definitions (analyses) is not a species of linguistic ignorance.

20. A New Resolution

The foregoing difficulties in Church's resolution suggest, first, that the paradox cannot be resolved by imposing more tight-fisted criteria of synonymy and, secondly, that the ignorance at the heart of the paradox is not linguistic in nature. This leads one to think that the paradox results instead from an *ambiguity* in intensional abstracts in natural language. On one way of reading intensional abstracts, substitutivity of synonyms is guaranteed; on another it is violated. The reading on which substitutivity is violated is that which is associated with ignorance of definitions, the source of the paradox. Consider the following intensional abstract:

the proposition that whatever is a circle is a locus of points in the same plane equidistant from some common point.

There is a reading of this abstract (reading I) according to which it denotes a different proposition from the one we would usually take to be denoted by:

the proposition that whatever is a circle is a circle.

And there is another reading (reading II) according to which it denotes the same proposition. Now we saw earlier that, considered one at a time, the following claims about sentences (1)–(3) seemed intuitively to hold:

(a) (1) and (2) are simultaneously satisfiable.
(b) (3) is true.
(c) The conjunction of (1) and (3) entails the negation of (2).

This leads to paradox since (a), (b), and (c) are jointly inconsistent. This paradox vanishes, though, when attention is paid to the two possible readings of the intensional abstract in (2). When it is given reading I, formulas (1) and (2) are simultaneously satisfiable. But the conjunction of (1) and (3) entails the negation of (2) only when the abstract in (2) is given reading II. The paradox thus is only a fallacy of equivocation.

It would be inelegant, perhaps even impossible, to treat each ambiguity of this sort as a semantic ambiguity, i.e., as an ambiguity resulting from a plurality of meanings of primitive non-logical constants. It is preferable to treat it as a structural (viz., syntactic) ambiguity, i.e., an ambiguity resulting from a plurality of syntactic deep structures that lead to intensional abstracts having the same surface syntactic structure.[15] This can be accomplished simply by introducing to L_ω a new syntactic operation according to which any open or closed sentence within an intensional abstract may be *underlined*. For heuristic purposes we may think of an underline as indicating that the intension expressed by an underlined open sentence is an *undefined concept* and the intension expressed by an underlined closed sentence is an *undefined thought*. (In these preliminary remarks it might be helpful to think of undefined concepts and thoughts as type 1 intensions, and defined concepts and thoughts as type 2 intensions; we shall see later that there is at least one further interesting way to conceive of defined and undefined concepts and thoughts.)

To see how this underlining device works, consider again the previous instance of the paradox that concerns identity. Formulas (7)–(9) are on one reading inconsistent with one another. This reading is now represented as follows:*

$$(7')\quad xK[(\forall y, z)(\underline{y = z} \supset \underline{y = z})]$$

$$(8')\quad \neg xK[(\forall y, z)(\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)} \supset \underline{y = z})]$$

$$(9')\quad [\underline{y = z}]_{yz} = [\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)}]_{yz}.$$

The inconsistency arises since (9′) entails

$$[(\forall y, z)(\underline{y = z} \supset \underline{y = z})] = [(\forall y, z)(\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)} \supset \underline{y = z})]$$

which, together with (7′), contradicts (8′). Before, we seemed to have a paradox on our hands because only an inconsistent reading of (7)–(9) could be represented; yet intuitively (7)–(9) were consistent. We can now avoid the paradox, though, for we can now represent the consistent reading by replacing (8′) with:

$$(8'')\quad \neg xK[(\forall y, z)((\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w) \supset \underline{y = z})].$$

The only difference between (8″) and (8′) is of course that $(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)$ is not underlined in (8″) whereas it is in (8′). Since $[\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)}]_{yz}$ denotes the undefined identity concept and since $[(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)]_{yz}$ denotes a defined identity concept, we know that

$$[\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)}]_{yz} \neq [(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)]_{yz}$$

and, hence, that

$$[(\forall y, z)(\underline{y = z} \supset \underline{y = z})] \neq [(\forall y, z)((\forall w)(y \,\Delta\, w \equiv y \,\Delta\, z) \supset \underline{y = z})].$$

This is what makes (7′), (8″), and (9′) consistent with one another, as the most natural reading of (7), (8), and (9) calls for. Thus, by enriching L_ω syntactically we successfully avoid the paradox.

It is easy to construct this enriched language, which will be called $\underline{L}_\omega$. The primitive symbols of $\underline{L}_\omega$ are those of L_ω plus the underline. The simultaneous inductive definition of term and formula for $\underline{L}_\omega$ goes as follows:

* Note Δ is a distinguished 2-place logical predicate that expresses the predication relation. For more on Δ, see the next chapter. Incidentally, there is an alternate inconsistent reading of (7)–(9) which can be represented by formulas that are like (7′)–(9′) except that all underlines are omitted.

(1) All variables are terms.
(2) If $t_1, \ldots, t_j$ are terms, then $F_i^j(t_1, \ldots, t_j)$ is a formula.
(3) If A and B are formulas and v_k is a variable, then $(A \,\&\, B)$, $\neg A$, and $(\exists v_i)A$ are formulas.
(4) If the expression A' is just like a formula A except perhaps that some subformulas occurring in A are underlined in A' and if $v_1, \ldots, v_p$ (for $p \geq 0$) are distinct variables, then $[A']_{v_1 \ldots v_p}$ is a term.
(5) If $[\underline{A}]_{v_1 \ldots v_p}$ (for $p \geq 0$) is a term, then so is $\underline{[A]_{v_1 \ldots v_p}}$.[16]

The indicated structural ambiguities in intensional abstracts in natural language can be unambiguously represented in $\underline{\mathrm{L}}_\omega$. Since these structural ambiguities are responsible for the paradox of analysis, the paradox should not arise in $\underline{\mathrm{L}}_\omega$. However, before this can be guaranteed, the right type of semantics for $\underline{\mathrm{L}}_\omega$ must be specified.

Since the aim is simply to characterize the logically valid formulas of $\underline{\mathrm{L}}_\omega$, a Tarski-style semantics will suffice. And what we are lacking at present is simply a method for modeling the heuristic distinction between undefined and defined ideas. There are two intriguing candidate methods, each of which calls for further philosophical study. The first invokes the theory of qualities and concepts which is developed in chapter 8. Specifically, undefined concepts would be identified with qualities (or connections), which are type 1 intensions, and defined concepts would be identified with complex concepts, which are type 2 intensions. This method is especially appealing since it meshes so nicely with the Platonic account of genuine forms and their analysis. The second method would just posit outright two primitive sorts of type 2 intensions, one corresponding to undefined concepts and the other to defined concepts. Either way, the semantics for $\underline{\mathrm{L}}_\omega$ is a straightforward affair once the appropriate types of model structures have been specified.[17] Relative to such semantics, it is easy to formulate logics for $\underline{\mathrm{L}}_\omega$ which can, it appears, be proven both sound and complete. Given this, a full resolution of the paradox of analysis is at hand.[18]

This finishes my study of intensional logic, the first stage in the study of PRPs. This complete foundation readies us for the second stage, the extension of intensional logic to include the predication relation.