ON A PROBLEM OF FORMAL LOGIC

By F. P. RAMSEY.

[Received 28 November, 1928.—Read 13 December, 1928.]

This paper is primarily concerned with a special case of one of the leading problems of mathematical logic, the problem of finding a regular procedure to determine the truth or falsity of any given logical formula*. But in the course of this investigation it is necessary to use certain theorems on combinations which have an independent interest and are most conveniently set out by themselves beforehand.

I.

The theorems which we actually require concern finite classes only, but we shall begin with a similar theorem about infinite classes which is easier to prove and gives a simple example of the method of argument.

THEOREM A. *Let Γ be an infinite class, and μ and r positive integers; and let all those sub-classes of Γ which have exactly r members, or, as we may say, let all r-combinations of the members of Γ be divided in any manner into μ mutually exclusive classes C_i ($i = 1, 2, \ldots, \mu$), so that every r-combination is a member of one and only one C_i; then, assuming the axiom of selections, Γ must contain an infinite sub-class Δ such that all the r-combinations of the members of Δ belong to the same C_i.*

Consider first the case $\mu = 2$. (If $\mu = 1$ there is nothing to prove.) The theorem is trivial when r is 1, and we prove it for all values of r by induction. Let us assume it, therefore, when $r = \rho - 1$ and deduce it for $r = \rho$, there being, since $\mu = 2$, only two classes C_i, namely C_1 and C_2.

* Called in German the *Entscheidungsproblem*; see Hilbert und Ackermann, *Grundzüge der Theoretischen Logik*, 72–81.

It may happen that Γ contains a member x_1 and an infinite sub-class Γ_1, not including x_1, such that the ρ-combinations consisting of x_1 together with any $\rho-1$ members of Γ_1, all belong to C_1. If so, Γ_1 may similarly contain a member x_2 and an infinite sub-class Γ_2, not including x_2, such that all the ρ-combinations consisting of x_2 together with $\rho-1$ members of Γ_2, belong to C_1. And, again, Γ_2 may contain an x_3 and a Γ_3 with similar properties, and so on indefinitely. We thus have two possibilities: either we can select in this way two infinite sequences of members of Γ $(x_1, x_2, \ldots, x_n, \ldots)$, and of infinite sub-classes of Γ $(\Gamma_1, \Gamma_2, \ldots, \Gamma_n, \ldots)$, in which x_n is always a member of Γ_{n-1}, and Γ_n a sub-class of Γ_{n-1} not including x_n, such that all the ρ-combinations consisting of x_n together with $\rho-1$ members of Γ_n, belong to C_1; or else the process of selection will fail at a certain stage, say the n-th, because Γ_{n-1} (or if $n=1$, Γ itself) will contain no member x_n and infinite sub-class Γ_n not including x_n such that all the ρ-combinations consisting of x_n together with $\rho-1$ members of Γ_n belong to C_1. Let us take these possibilities in turn.

If the process goes on for ever let Δ be the class $(x_1, x_2, \ldots, x_n, \ldots)$. Then all these x's are distinct, since if $r > s$, x_r is a member of Γ_{r-1} and so of $\Gamma_{r-2}, \Gamma_{r-3}, \ldots$, and ultimately of Γ_s which does not contain x_s. Hence Δ is infinite. Also all ρ-combinations of members of Δ belong to C_1; for if x_s is the term of such a combination with least suffix s, the other $\rho-1$ terms of the combination belong to Γ_s, and so form with x_s a ρ-combination belonging to C_1. Γ therefore contains an infinite sub-class Δ of the required kind.

Suppose, on the other hand, that the process of selecting the x's and Γ's fails at the n-th stage, and let y_1 be any member of Γ_{n-1}. Then the $(\rho-1)$-combinations of members of $\Gamma_{n-1}-(y_1)$ can be divided into two mutually exclusive classes C_1' and C_2' according as the ρ-combinations formed by adding to them y_1 belong to C_1 or C_2, and by our theorem (A), which we are assuming true when $r=\rho-1$ (and $\mu=2$), $\Gamma_{n-1}-(y_1)$ must contain an infinite sub-class Δ_1 such that all $(\rho-1)$-combinations of the members of Δ_1 belong to the same C_i'; *i.e.* such that the ρ-combinations formed by joining y_1 to $\rho-1$ members of Δ_1 all belong to the same C_i. Moreover, this C_i cannot be C_1, or y_1 and Δ_1 could be taken to be x_n and Γ_n and our previous process of selection would not have failed at the n-th stage. Consequently the ρ-combinations formed by joining y_1 to $\rho-1$ members of Δ_1 all belong to C_2. Consider now Δ_1 and let y_2 be any of its members. By repeating the preceding argument $\Delta_1-(y_2)$ must contain an infinite sub-class Δ_2 such that all the ρ-combinations got by joining y_2 to $\rho-1$ members of Δ_2 belong to the same C_i.

And, again, this C_i cannot be C_1, or, since y_2 is a member and Δ_2 a sub-class of Δ_1 and so of Γ_{n-1} which includes Δ_1, y_2 and Δ_2 could have been chosen as x_n and Γ_n and the process of selecting these would not have failed at the n-th stage. Now let y_3 be any member of Δ_2; then $\Delta_2-(y_3)$ must contain an infinite sub-class Δ_3 such that all ρ-combinations consisting of y_3 together with $\rho-1$ members of Δ_3, belong to the same C_i, which, as before, cannot be C_1 and must be C_2. And by continuing in this way we shall evidently find two infinite sequences $y_1, y_2, \ldots, y_n, \ldots$ and $\Delta_1, \Delta_2, \ldots, \Delta_n, \ldots$ consisting respectively of members and sub-classes of Γ, and such that y_n is always a member of Δ_{n-1}, Δ_n a sub-class of Δ_{n-1} not including y_n, and all the ρ-combinations formed by joining y_n to $\rho-1$ members of Δ_n belong to C_2; and if we denote by Δ the class $(y_1, y_2, \ldots, y_n, \ldots)$ we have, by a previous argument, that all ρ-combinations of members of Δ belong to C_2.

Hence, in either case, Γ contains an infinite sub-class Δ of the required kind, and Theorem A is proved for all values of r, provided that $\mu=2$. For higher values of μ we prove it by induction; supposing it already established for $\mu=2$ and $\mu=\nu-1$, we deduce it for $\mu=\nu$.

The r-combinations of members of Γ are then divided into ν classes C_i $(i=1, 2, \ldots, \nu)$. We define new classes C'_i for $i=1, 2, \ldots, \nu-1$ by

$$C'_i = C_i \quad (i = 1, 2, \ldots, \nu-2),$$

$$C'_{\nu-1} = C_{\nu-1}+C_\nu.$$

Then by the theorem for $\mu=\nu-1$, Γ must contain an infinite sub-class Δ such that all r-combinations of the members of Δ belong to the same C'_i. If, in this C'_i, $i \leqslant \nu-2$, they all belong to the same C_i, which is the result to be proved; otherwise they all belong to $C'_{\nu-1}$, *i.e.* either to $C_{\nu-1}$ or to C_ν. In this case, by the theorem for $\mu=2$, Δ must contain an infinite sub-class Δ' such that the r-combinations of members of Δ' either all belong to $C_{\nu-1}$ or all belong to C_ν; and our theorem is thus established.

Coming now to finite classes it will save trouble to make some conventions as to notation. Small letters other than x and y, whether Italic or Greek (*e.g.* n, r, μ, m) will always denote finite cardinals, positive unless otherwise stated. Large Greek letters (*e.g.* Γ, Δ) will denote classes, and their suffixes will indicate the number of their members (*e.g.* Γ_m is a class with m members). The letters x and y will represent members of the classes Γ, Δ, etc., and their suffixes will be used merely to distinguish them. Lastly, the letter C will stand, as before, for classes of combinations, and its suffixes will not refer to the number of members, but serve merely to distinguish the different classes of combinations considered.

Corresponding to Theorem A we then have

THEOREM B. *Given any r, n, and μ we can find an m_0 such that, if $m \geqslant m_0$ and the r-combinations of any Γ_m are divided in any manner into μ mutually exclusive classes C_i $(i = 1, 2, \ldots, \mu)$, then Γ_m must contain a sub-class Δ_n such that all the r-combinations of members of Δ_n belong to the same C_i.*

This is the theorem which we require in our logical investigations, and we should at the same time like to have information as to how large m_0 must be taken for any given r, n, and μ. This problem I do not know how to solve, and I have little doubt that the values for m_0 obtained below are far larger than is necessary.

To prove the theorem we begin, as in Theorem A, by supposing that $\mu = 2$. We then take, not Theorem B itself, but the equivalent

THEOREM C. *Given any r, n, and k such that $n+k \geqslant r$, there is an m_0 such that, if $m \geqslant m_0$ and the r-combinations of any Γ_m are divided into two mutually exclusive classes C_1 and C_2, then Γ_m must contain two mutually exclusive sub-classes Δ_n and Λ_k such that all the combinations formed by r members of $\Delta_n + \Lambda_k$ which include at least one member from Δ_n belong to the same C_i.*

That this is equivalent to Theorem B with $\mu = 2$ is evident from the fact that, for any given r, Theorem C, for n and k, asserts more than Theorem B for n, but less than Theorem B for $n+k$.

The proof of Theorem C must be performed by mathematical induction, and can conveniently be set out as a demonstration that it is possible to define by recursion a function $f(r, n, k)$ which will serve as m_0 in the theorem.

If $r = 1$, the theorem is evidently true with m_0 equal to the greater of $2n-1$ and $n+k$, so that we may define

$$f(1, n, k) = \max(2n-1, n+k) \quad (n \geqslant 1,\ k \geqslant 0).$$

For other values of r we define $f(r, n, k)$ by recursion formulae involving an auxiliary function $g(r, n, k)$. Suppose that $f(r-1, n, k)$ has been defined for a certain $r-1$, and all n, k such that $n+k \geqslant r-1$, then we define it for r by putting

$$\begin{aligned}
f(r, 1, k) &= f(r-1, k-r+2, r-2)+1 && (k+1 \geqslant r),\\
g(r, 0, k) &= \max(r-1, k),\\
g(r, n, k) &= f\{r, 1, g(r, n-1, k)\} && (n \geqslant 1),\\
f(r, n, k) &= f\{r, n-1, g(r, n, k)\} && (n > 1).
\end{aligned}$$

These formulae can be easily seen to define $f(r, n, k)$ for all positive values of r, n and k satisfying $n+k \geqslant r$, and $g(r, n, k)$ for all values of r greater than 1, and all positive values of n and k; and we shall prove that Theorem C is true when we take m_0 to be this $f(r, n, k)$. We know that this is so when $r=1$, and we shall therefore assume it for all values up to $r-1$ and deduce it for r.

When $n=1$, and $m \geqslant m_0 = f(r-1, k-r+2, r-2)+1$, we may take any member x of Γ_m to be sole member of Δ_1 and there remain at least $f(r-1, k-r+2, r-2)$ members of $\Gamma_m-(x)$; the $(r-1)$-combinations of these members of $\Gamma_m-(x)$ can be divided into classes C_1' and C_2' according as they belong to C_1 or C_2 when x is added to them, and, by our theorem for $r-1$, $\Gamma_m-(x)$ must contain two mutually exclusive classes Δ_{k-r+2}, Λ_{r-2} such that every combination of $r-1$ terms from $\Delta_{k-r+2}+\Lambda_{r-2}$ (since one of its terms must come from Δ_{k-r+2}, Λ_{r-2} having only $r-2$ members) belongs to the same C_i'. Taking Λ_k to be this $\Delta_{k-r+2}+\Lambda_{r-2}$ all combinations consisting of x, together with $r-1$ members of Λ_k, belong to the same C_i. The theorem is therefore true for r when $n=1$.

For other values of n we prove it by induction, assuming it for $n-1$ and deducing it for n. Taking

$$m \geqslant m_0 = f(r, n, k) = f\{r, n-1, g(r, n, k)\},$$

Γ_m must, by the theorem for $n-1$, contain a Δ_{n-1} and a $\Lambda_{g(r,n,k)}$ such that every combination of r members of $\Delta_{n-1}+\Lambda_{g(r,n,k)}$, at least one term of which comes from Δ_{n-1}, belongs to the same C_i, say to C_1. If, now, $\Lambda_{g(r,n,k)}$ contains a member x and a sub-class Λ_k not including x, such that every combination of x and $r-1$ members of Λ_k belongs to C_1, then, taking Δ_n to be $\Delta_{n-1}+(x)$ and Λ_k to be this Λ_k, our theorem is true. If not, there can be no member of $\Lambda_{g(r,n,k)}$ which has a sub-class of k members of $\Lambda_{g(r,n,k)}$ connected with it in this way. But since

$$g(r, n, k) = f\{r, 1, g(r, n-1, k)\},$$

$\Lambda_{g(r,n,k)}$ must contain a member x_1 and a sub-class $\Lambda_{g(r,n-1,k)}$, not including x_1, such that x_1 combined with any $r-1$ members of $\Lambda_{g(r,n-1,k)}$ gives a combination belonging to the same C_i, which cannot be C_1, or x_1 and any k members of $\Lambda_{g(r,n-1,k)}$ could have been taken as the x and Λ_k above. Hence the combinations formed by x_1 together with any $r-1$ members of $\Lambda_{g(r,n-1,k)}$ all belong to C_2. But now

$$g(r, n-1, k) = f\{r, 1, g(r, n-2, k)\},$$

and $\Lambda_{g(r,n-1,k)}$ must contain an x_2 and a $\Lambda_{g(r,n-2,k)}$, not including x_2, such that the combinations formed by x_2 and $r-1$ members of $\Lambda_{g(r,n-2,k)}$ all belong to the same C_i, which must, as before, be C_2, since x_2 and $\Lambda_{g(r, n-2, k)}$ are both contained in $\Lambda_{g(r, n, k)}$ and $g(r, n-2, k) \geqslant k$. Continuing in this way we can find n distinct terms $x_1, x_2, \ldots, x_n$ and a $\Lambda_{g(r, 0, k)}$ such that every combination of r terms from $(x_1, x_2, \ldots, x_n)+\Lambda_{g(r, 0, k)}$ belongs to C_2, provided that at least one term of the combination comes from $(x_1, x_2, \ldots, x_n)$. Since $g(r, 0, k) \geqslant k$ this proves our theorem, taking Δ_n to be $(x_1, x_2, \ldots, x_n)$ and Λ_k to be any k terms of $\Lambda_{g(r, 0, k)}$.

Theorem C is therefore established for all values of r, n, and k, with m_0 equal to $f(r, n, k)$. It follows that, if $\mu = 2$, Theorem B is true for all values of r and n with m_0 equal to $f(r, n-r+1, r-1)$, which we shall also call $h(r, n, 2)$.

For other values of μ we prove Theorem B by induction, taking m_0 to be $h(r, n, \mu)$, where

$$h(r, n, 2) = f(r, n-r+1, r-1)$$

$$h(r, n, \mu) = h\{r, h(r, n, \mu-1), 2\} \quad (\mu > 2).$$

For, assuming the theorem for $\mu-1$, we prove it for μ by defining new classes of combinations

$$C_1' = C_1,$$

$$C_2' = \sum_{i=2}^{\mu} C_i.$$

If then $m \geqslant h(r, n, \mu) = h\{r, h(r, n, \mu-1), 2\}$, by the theorem for $\mu = 2$, Γ_m must contain a $\Gamma_{h(r, n, \mu-1)}$ the r-combinations of whose members belong either all to C_1' or all to C_2'. In the first case there is no more to prove; in the second we have only to apply the theorem for $\mu-1$ to $\Gamma_{h(r, n, \mu-1)}$.

In the simplest case in which $r = \mu = 2$ the above reasoning gives m_0 equal to $h(2, n, 2)$, which is easily shown to be $2^{n(n-1)/2}$. But for this case there is a simple argument which gives the much lower value $m_0 = n\,!$, and shows that our value $h(r, n, \mu)$ is altogether excessive.

For, taking Theorem C first, we can prove by induction with regard to n that, for $r = 2$, we may take m_0 to be $k\,.\,(n+1)\,!$. (k is here supposed greater than or equal to 1.) For this is true when $n = 1$, since, if $m \geqslant 2k$, of the $m-1$ pairs obtained by combining any given member of Γ_m with the others, at least k must belong to the same C_i. Assuming it, then, for $n-1$, let us prove it for n.

If $m \geqslant k\,.\,(n+1)\,! = k(n+1)\,.\,n\,!$, Γ_m must, by the theorem for $n-1$, contain two mutually exclusive sub-classes Δ_{n-1} and $\Lambda_{k(n+1)}$ such that all pairs from $\Delta_{n-1}+\Lambda_{k(n+1)}$, at least one term of which comes from Δ_{n-1}, belong to the same C_i, say C_1. Now consider the members of $\Lambda_{k(n+1)}$; in the first place, there may be one of these, x say, which is such that there are k other members of $\Lambda_{k(n+1)}$ which combined with x give pairs belonging to C_1. If so, the theorem is true, taking Δ_n to be $\Delta_{n-1}+(x)$; if not, let x_1 be any member of $\Lambda_{k(n+1)}$. Then there are at most $k-1$ other members of $\Lambda_{k(n+1)}$ which combined with x_1 give pairs belonging to C_1, and $\Lambda_{k(n+1)}-(x_1)$ must contain a Λ_{kn} any member of which gives when combined with x_1 a pair belonging to C_2. Let x_2 be any member of Λ_{kn}, then, since x_2 and Λ_{kn} are both contained in $\Lambda_{k(n+1)}$, there are at most $k-1$ other members of Λ_{kn} which when combined with x_2 give pairs belonging to C_1. Hence $\Lambda_{kn}-(x_2)$ contains a $\Lambda_{k(n-1)}$ any member of which combined with x_2 gives a pair belonging to C_2. Continuing in this way we obtain $x_1, x_2, \ldots, x_n$ and Λ_k, such that every pair x_i, x_j and every pair consisting of an x_i and a member of Λ_k belongs to C_2. Theorem C is therefore proved.

Theorem B for n then follows, with the m_0 of Theorem C for $n-1$ and 1, *i.e.* with m_0 equal to $n!$*; and it is an easy extension to show that, if in Theorem B $r=2$ but $\mu \neq 2$, we can take m_0 to be $n!!!. \ldots$, where the process of taking the factorial is performed $\mu-1$ times.

II.

We shall be concerned with logical formulae containing variable propositional functions, *i.e.* predicates or relations, which we shall denote by Greek letters ϕ, χ, ψ, etc. These functions have as arguments individuals denoted by x, y, z, etc., and we shall deal with functions with any finite number of arguments, *i.e.* of any of the forms

$$\phi(x),\ \chi(x, y),\ \psi(x, y, z),\ \ldots.$$

In addition to these variable functions we shall have the one constant function of identity

$$x = y \quad \text{or} \quad = (x, y).$$

By operating on the values of ϕ, χ, ψ, ..., and $=$ with the logical operations

$\sim$	meaning	*not*,
$\vee$	,,	*or*,
.	,,	*and*,
(x)	,,	*for all x*.
(Ex)	,,	*there is an x for which*,

* But this value is, I think, still much too high. It can easily be lowered slightly even when following the line of argument above, by using the fact that if k is even it is impossible for every member of an odd class to have exactly $k-1$ others with which it forms a pair of C_1, for then twice the number of these pairs would be odd; we can thus start when k is even with a $\Lambda_{k(n+1)-1}$ instead of a $\Lambda_{k(n+1)}$

we can construct expressions such as

$$[(x, y)\{\phi(x, y) \vee x = y\}] \vee \{(Ez)\chi(z)\}$$

in which all the individual variables are made "apparent" by prefixes (x) or (Ex), and the only real variables left are the functions ϕ, χ, ... Such an expression we shall call a *first order formula*.

If such a formula is true for all interpretations* of the functional variables ϕ, χ, ψ, etc., we shall call it *valid*, and if it is true for no interpretations of these variables we shall call it *inconsistent*. If it is true for some interpretations (whether or not for all) we shall call it *consistent*†.

The *Entscheidungsproblem* is to find a procedure for determining whether any given formula is valid, or, alternatively, whether any given formula is consistent; for these two problems are equivalent, since the necessary and sufficient condition for a formula to be consistent is that its contradictory should not be valid. We shall find it more convenient to take the problem in this second form as an investigation of *consistency*. The consistency of a formula may, of course, depend on the number of individuals in the universe considered, and we shall have to distinguish between formulae which are consistent in every universe and those which are only consistent in universes with some particular numbers of members. Whenever the universe is infinite we shall have to assume the axiom of selections.

The problem has been solved by Behmann‡ for formulae involving only functions of one variable, and by Bernays and Schönfinkel§ for formulae involving only two individual apparent variables. It is solved below for the further case in which, when the formula is written in "normal form", there are any number of prefixes of generality (x) but none of existence (Ex)‖. By "normal form"¶ is here meant that all the prefixes stand at the beginning, with no negatives between or in front of them, and have scopes extending to the end of the formula.

---

* To avoid confusion we call a constant function substituted for a variable ϕ, not a value but an *interpretation* of ϕ; the *values* of $\phi(x, y, z)$ are got by substituting constant individuals for x, y, and z.

† German *erfüllbar*.

‡ H. Behmann, "Beiträge zur Algebra der Logik und zum Entscheidungsproblem", *Math. Annalen*, 86 (1922), 163-229.

§ P. Bernays und M. Schönfinkel, "Zum Entscheidungsproblem der mathematischen Logik", *Math. Annalen*, 99 (1928), 342-372. These authors do not, however, include identity in the formulae they consider.

‖ Later we extend our solution to the case in which there are also prefixes of existence provided that these all precede all the prefixes of generality.

¶ Hilbert und Ackermann, *op. cit.*, 63-4.

The formulae to be considered are thus of the form

$$(x_1, x_2, \ldots, x_n) F(\phi, \chi, \psi; \ldots, =, x_1, x_2, \ldots, x_n),$$

where the matrix F is a truth-function of values of the functions ϕ, χ, ψ, etc., and $=$ for arguments drawn from $x_1, x_2, \ldots, x_n$.

This type of formula is interesting as being the general type of an axiom system consisting entirely of "general laws"*. The axioms for order, betweenness, and cyclic order are all of this nature, and we are thus attempting a general theory of the consistency of axiom systems of a common, if very simple, type.

If identity does not occur in F the problem is trivial, since in this case whether the formula is consistent or not can be shown to be independent of the number of individuals in the universe, and we have only the easy task of testing it for a universe with one member only†.

But when we introduce identity the question becomes much more difficult, for although it is still obvious that if the formula is consistent in a universe U it must be consistent in any universe with fewer members than U, yet it may easily be consistent in the smaller universe but not in the larger. For instance,

$$(x_1, x_2)[x_1 = x_2 \vee \{\phi(x_1) . \sim \phi(x_2)\}]$$

is consistent in a universe with only one member but not in any other.

We begin our investigation by expressing F in a special form. F is a truth-function of the values of ϕ, χ, ψ, ..., and $=$ for arguments drawn from $x_1, x_2, \ldots, x_n$. If ϕ is a function of r variables there will be n^r‡ values of ϕ which can occur in F, and F will be a truth-function of Σn^r values of ϕ, χ, ψ, ..., and $=$, which we shall call *atomic propositions*. With regard to these Σn^r atomic propositions there are $2^{\Sigma n^r}$ possibilities of truth and falsity which we shall call *alternatives*, each alternative being a conjunction of Σn^r propositions which are either atomic propositions or their contradictions. In constructing the alternatives all the Σn^r atomic propositions are to be used whether or not they occur in F. F can then be expressed as a disjunction of some of these alternatives, namely those with which it is compatible. It is well known that such an

* C. H. Langford, "Analytic completeness of postulate sets", *Proc. London Math. Soc.* (2), 25 (1926), 115-6.

† Bernays und Schönfinkel, *op. cit.*, 359. We disregard altogether universes with no members.

‡ Here and elsewhere numbers are given not because they are relevant to the argument, but to enable the reader to check that he has in mind the same class of entities as the author.

expression is possible; indeed, it is the dual of what Hilbert and Ackermann call the "ausgezeichnete konjunktive Normalform"*, and is fundamental also in Wittgenstein's logic. The only exception is when F is a self-contradictory truth-function, in which case our formula is certainly not consistent.

F having been thus expressed as a disjunction of alternatives (in our special sense of the word), our next task is to show that some of these alternatives may be able to be removed without affecting the consistency or inconsistency of the formula. If all the alternatives can be removed in this way the formula will be inconsistent; otherwise we shall have still to consider the alternatives that remain.

In the first place an alternative may violate the laws of identity by containing parts of any of the following forms :—

$$x_i \neq x_i \dagger,$$

$$x_i = x_j \,.\, x_j \neq x_i \qquad (i \neq j),$$

$$x_i = x_j \,.\, x_j = x_k \,.\, x_i \neq x_k \quad (i \neq j,\ j \neq k,\ k \neq i),$$

or by containing $x_i = x_j$ $(i \neq j)$ and values of a function ϕ and its contradictory $\sim \phi$ for sets of arguments which become the same when x_i is substituted for x_j [*e.g.* $x_1 = x_2 \,.\, \phi(x_1, x_2, x_3) \,.\, \sim \phi(x_2, x_1, x_3)$].

Any alternative which violates these laws must always be false and can evidently be discarded without affecting the consistency of the formula. The remaining alternatives can then be classified according to the number of x's they make to be different, which may be anything from 1 up to n.

Suppose that for a given alternative this number is ν, then we can derive from it what we will call the corresponding y alternative by the following process :—

For x_1, wherever it occurs in the given alternative, write y_1; next, if in the alternative $x_2 = x_1$, for x_2 write y_1 again, if not for x_2 write y_2. In general, if x_i is in the given alternative identical with any x_j with j less than i, write for x_i the y previously written for x_j; otherwise write for x_i, y_{k+1}, where k is the number of y's already introduced. The expression which results contains ν y's all different instead of n x's, some of which are identical, and we shall call it the y alternative corresponding to the given x alternative.

---

* *Op. cit.*, 16.

† We write $x \neq y$ for $\sim (x = y)$.

Thus to the alternative

$$\phi(x_1) . \sim \phi(x_2) . \phi(x_3) . \sim \phi(x_4) . x_1 = x_3 . x_2 = x_4 . x_1 \neq x_2 {}^*$$

corresponds the y alternative

$$\phi(y_1) . \sim \phi(y_2) . y_1 \neq y_2 .$$

We call two y alternatives *similar* if they contain the same number of y's and can be derived from one another by permuting those y's, and we call two x alternatives *equivalent* if they correspond to similar (or identical) y alternatives.

Thus

$$\phi(x_1) . \sim \phi(x_2) . \phi(x_3) . \sim \phi(x_4) . x_1 = x_3 . x_2 = x_4 . x_1 \neq x_2, \qquad (\alpha)$$

is equivalent to

$$\sim \phi(x_1) . \phi(x_2) . \phi(x_3) . \phi(x_4) . x_1 \neq x_2 . x_2 = x_3 = x_4, \qquad (\beta)$$

since they correspond to the similar y alternatives

$$\phi(y_1) . \sim \phi(y_2) . y_1 \neq y_2$$

$$\sim \phi(y_1) . \phi(y_2) . y_1 \neq y_2$$

derivable from one another by interchanging y_1 and y_2, although (α) and (β) are not so derivable by permuting the x's.

We now see that we can discard any alternative contained in F unless F also contains all the alternatives equivalent to it; *e.g.* if F contains (α) but not (β), (α) may be discarded from it. For omitting alternatives clearly cannot make the formula consistent if it was not so before; and we can easily prove that, if it was consistent before, omitting *these* alternatives cannot make it inconsistent.

For suppose that the formula is consistent, *i.e.* that for some particular interpretation of ϕ, χ, ψ, ..., F is true for every set of x's, and let p be an alternative contained in F, q an alternative equivalent to p but not contained in F. Then for every set of x's one and only one alternative in F will (on this interpretation of ϕ, χ, ψ, ...) be the true one, and this alternative can never be p. For if it were p, the corresponding y alternative would be true for some set of y's, and the similar y alternative corresponding to q would be true for a set of y's got by permuting this last set. Giving the x's suitable values in terms of the y's, q would then

* We take one function of one variable only for simplicity; also to save space we omit expressions which may be taken for granted, such as $x_1 = x_1$, $x_1 \neq x_4$.

be true for a certain set of x's and F would be false for these x's contrary to hypothesis. Hence p is never the true alternative and may be omitted without affecting the consistency of the formula.

When we have discarded all these alternatives from F, the remainder will fall into sets each of which is the complete set of all alternatives equivalent to a given alternative. To such a set of x alternatives will correspond a complete set of similar y alternatives, and the disjunction of such a complete set of similar y alternatives (*i.e.* of all permutations of a given y alternative) we shall call a *form**. A form containing ν y's we shall denote by an Italic capital with suffix ν, *e.g.* A_ν, B_ν.

The force of our formula can now be represented by the following conjunction, which we shall call P.

$$\left.\begin{array}{ll} \text{For every } y_1, & A_1 \text{ or } B_1 \text{ or } \ldots \\ \text{For every distinct } y_1, y_2, & A_2 \text{ or } B_2 \text{ or } \ldots \\ \ldots \quad \ldots \quad \ldots \quad \ldots & \ldots \quad \ldots \quad \ldots \\ \text{For every distinct } y_1, y_2, \ldots, y_\nu, & A_\nu \text{ or } B_\nu \text{ or } \ldots \\ \ldots \quad \ldots \quad \ldots \quad \ldots & \ldots \quad \ldots \quad \ldots \\ \text{For every distinct } y_1, y_2, \ldots, y_n, & A_n \text{ or } B_n \text{ or } \ldots \end{array}\right\} \quad (P)$$

where A_ν, B_ν†, etc., are the forms corresponding to the x alternatives still remaining in F. If for any ν there are no such forms, *i.e.* if no alternatives with ν different x's remain in F, our formula implies that there are no such things as ν distinct individuals, and so cannot be consistent in a world of ν or more members.

We have now to define what is meant by saying that one form is *involved* in another. Consider a form A_ν and take one of the y alternatives contained in it. This y alternative is a conjunction of the values of ϕ, χ, ψ, ..., and their negatives for arguments drawn from $y_1, y_2, \ldots, y_\nu$. (We may leave out the values of identity and difference, since it is taken for granted that y's are always different.) If $\mu < \nu$ we can select μ of these y's in any way and leave out from the alternative all the terms in it which contain any of the $\nu-\mu$ y's not selected. We have left an alternative in μ y's which we can renumber $y_1, y_2, \ldots, y_\mu$, and the form E_μ to which this new alternative belongs we shall describe as being *involved* in the A_ν with which we started. Starting with one particular y alternative in A_ν we shall get a large number of different E_μ's by

* Cf. Langford, *op. cit.*, 116–120.

† The notation is partially misleading, since A_ν has no closer relation to A_μ than to B_μ.

choosing differently the μ y's which we select to preserve; and from whichever y alternative in A_ν we start, the E_μ's which we find to be involved in A_ν will be the same.

For example,

$$\{\phi(y_1, y_1) \cdot \phi(y_1, y_2) \cdot \phi(y_2, y_1) \cdot \sim \phi(y_2 \cdot y_2)\}$$
$$\vee \{\sim \phi(y_1, y_1) \cdot \phi(y_1, y_2) \cdot \phi(y_2, y_1) \cdot \phi(y_2, y_2)\}$$

is a form A_2 which involves the two E_1's

$$\phi(y_1, y_1)$$
$$\sim \phi(y_1, y_1).$$

It is clear that if for some distinct set of ν y's a form A_ν is true, then every form E_μ involved in A_ν will be true for some distinct set of μ y's contained in the ν.

We are now in a position to settle the consistency or inconsistency of our formula when N, the number of individuals in the universe, is less than or equal to n, the number of x's in our formula. In fact, if $N \leqslant n$, it is necessary and sufficient for the consistency of the formula that P should contain a form A_N together with all the forms E_μ involved in it for every μ less than N.

This condition is evidently necessary, since the N individuals in the universe must, taken as $y_1, y_2, \ldots, y_N$, have some form A_N in regard to any $\phi, \chi, \psi, \ldots$; and all forms involved in this A_N must be true for different selections of y's, and so contained in P if P is to be true for this $\phi, \chi, \psi, \ldots$.

Conversely, suppose that P contains a form A_N together with all forms involved in A_N; then, calling the N individuals in the universe $y_1, y_2, \ldots, y_N$, we can define functions $\phi, \chi, \psi, \ldots$ to make any assigned y alternative in A_N true; for any permutation of these N y's another alternative in A_N will be the true one, and for any subset of y's some y alternative in a form involved in A_N. Since all these y alternatives are by hypothesis contained in P, P will be true for these $\phi, \chi, \psi, \ldots$, and our formula consistent.

When, however, $N > n$ the problem is not so simple, although it clearly depends on the A_n's in P such that all forms involved in them are also contained in P. These A_n's we may call *completely contained* in P, and if there are no such A_n's a similar argument to that used when $N \leqslant n$ will show that the formula is inconsistent. But the converse argument, that if there is an A_n completely contained in P the formula must be consistent, no longer holds good; and to proceed further we have to introduce a new conception, the conception of a form being *serial*.

But before proceeding to explain this idea it is best to simplify matters by the introduction of new functions. Let ϕ be one of the variable functions in our formula, with, say, r arguments. Then, if $r < n$, ϕ will occur in P with all its arguments different [*e.g.* $\phi(y_1, y_2, \ldots, y_r)$] and also with some of them the same [*e.g.* $\phi(y_1, y_2, \ldots, y_{r-1}, y_1)$]; but we can conveniently eliminate values of the second kind by introducing new functions of fewer arguments than r, which, when all their arguments are different, take values equivalent to those of ϕ with some of its arguments identical.

E.g. we may put

$$\phi_1(y_1, y_2, \ldots, y_{r-1}) = \phi(y_1, y_2, \ldots, y_{r-1}, y_1).$$

In this way ϕ gives rise to a large number of functions with fewer arguments; each of these functions we define only for the case in which all its arguments are different, as is secured by these arguments being y's with different suffixes. If $r > n$, there is no difference except that ϕ can never occur with all its arguments different, and so is entirely replaced by the new functions.

If we do this for all the functions $\phi, \chi, \psi, \ldots$, and replace them by new functions wherever they occur in P with some of their arguments the same, P will contain a new set of variable functions (including all the old ones which have no more than n arguments), and these will never occur in P with the same argument repeated.

It is easy to see that this transformation does not affect the consistency of the formula, for, if it were consistent before, it must be consistent afterwards, since the new functions have simply to be replaced by their definitions. And if it is consistent afterwards it must have been so before, since any function of the old set has only to be given for any set of arguments the value of the appropriate function of the new set*.

---

* For instance, if $\phi(y_1, y_2, y_3)$ is a function of the old set, we have five new functions

$$\begin{aligned}
\phi_0(y_1, y_2, y_3) &= \phi(y_1, y_2, y_3), \\
\chi_0(y_1, y_2) &= \phi(y_1, y_1, y_2), \\
\psi_0(y_1, y_2) &= \phi(y_1, y_2, y_1), \\
\pi_0(y_1, y_2) &= \phi(y_1, y_1, y_2), \\
\rho_0(y_1) &= \phi(y_1, y_1, y_1),
\end{aligned}$$

and any value of ϕ is equivalent to a value of one and only one of the new functions. It must be remembered that the new functions are used only with all their arguments different; for otherwise they would not be independent, since we should have, for instance, $\chi_0(y_1, y_1)$ equivalent to $\rho_0(y_1)$. But $\chi_0(y_1, y_1)$ never occurs, and $\phi(y_1, y_1, y_1)$ is equivalent not to any value of χ_0 but only to $\rho_0(y_1)$.

In view of this fact we shall find it more convenient to take P in its new form, and denote the new set of functions by $\phi_0, \chi_0, \psi_0, \ldots$.

Suppose, then, that ϕ_0 is a function of r variables; there are

$$n(n-1)\ldots(n-r+1)$$

values of ϕ_0 with r different arguments drawn from $y_1, y_2, \ldots, y_n$ and every y alternative must contain each of these values or its contradictory. $r!$ of these values will have as arguments permutations of $y_1, y_2, \ldots, y_r$. Any other set of r y's can be arranged in the order of their suffixes as $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$, $s_1 < s_2 < s_3 \ldots < s_r$, and it may happen that a given alternative contains the values of ϕ_0 for those and only those permutations of $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ which correspond (in the obvious way) to the permutations of $y_1, y_2, \ldots, y_r$ for which it (the alternative) contains the values of ϕ_0; *e.g.* if the alternative contains $\phi_0(y_1, y_2, \ldots, y_r)$ and $\phi_0(y_r, y_{r-1}, \ldots, y_1)$, but for every other permutation of $y_1, y_2, \ldots, y_r$ contains the corresponding value of $\sim\phi_0$, then it may happen that the alternative contains $\phi_0(y_{s_1}, y_{s_2}, \ldots, y_{s_r})$ and $\phi_0(y_{s_r}, y_{s_{r-1}}, \ldots, y_{s_1})$, but for every other permutation of $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ contains the corresponding value of $\sim\phi_0$.

If this happens, no matter how the set of r y's, $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ is chosen from $y_1, y_2, \ldots, y_n$, then we say that the *alternative* is *serial in* ϕ_0*, and if an alternative is serial in every function of the new set we shall call it *serial* simply.

Consider, for example, the following alternative, in which we may imagine ϕ_0 and ψ_0 to be derived from one "old" function ϕ by the definitions

$$\phi_0(y_i, y_k) = \phi(y_i, y_k),$$

$$\psi_0(y_i) = \phi(y_i, y_i),$$

$$\phi_0(y_1, y_2) . \sim\phi_0(y_2, y_1) . \phi_0(y_1, y_3) . \sim\phi_0(y_3, y_1) . \phi_0(y_2, y_3) . \sim\phi_0(y_3, y_2),$$

$$\psi_0(y_1) . \sim\psi_0(y_2) . \psi_0(y_3).$$

This is serial in ϕ_0, since we always have $\phi_0(y_{s_1}, y_{s_2}) . \sim\phi_0(y_{s_2}, y_{s_1})$; but not in ψ_0, since we sometimes have $\psi_0(y_{s_1})$, but sometimes $\sim\psi_0(y_{s_1})$. Hence it is not a serial alternative.

We call a *form serial* when it contains at least one serial alternative, and can now state our chief result as follows.

---

* Thus, if ϕ_0 is a function of n variables, *all* alternatives are serial in ϕ_0.

THEOREM.—*There is a finite number m, depending on n, the number of functions* ϕ, χ, ψ, ..., *and the numbers of their arguments, such that the necessary and sufficient condition for our formula to be consistent in a universe with m or more members is that there should be a serial form* A_n *completely contained in P. For consistency in a universe of fewer than m members this condition is sufficient but not necessary.*

We shall first prove that, whatever be the number N of individuals in the universe, the condition is sufficient for the consistency of the formula. If $N \leqslant n$, this is a consequence of a previous result, since, if A_n is completely contained in P, so is any A_N involved in A_n.

If $N > n$, we suppose the universe ordered in a series by a relation R. (If N is infinite this requires the Axiom of Selections.) Let q be any serial alternative contained in A_n. If ϕ_0 is a function of r arguments, q will contain the values of either ϕ_0 or $\sim \phi_0$ (but not both) for every permutation of $y_1, y_2, \ldots, y_r$. Any such permutation can be written $y_{\rho_1}, y_{\rho_2}, \ldots, y_{\rho_r}$ where $\rho_1, \rho_2, \ldots, \rho_r$ are $1, 2, \ldots, r$ rearranged. We make a list of all those permutations $(\rho_1, \rho_2, \ldots, \rho_r)$ for which q contains the values of ϕ_0, and call this list Σ. We now give ϕ_0 the constant interpretation that $\phi_0(z_1, z_2, \ldots, z_r)$ is to be true if and only if the order of the terms $z_1, z_2, \ldots, z_r$ in the series R is given by one of the permutations $(\rho_1, \rho_2, \ldots, \rho_r)$ contained in Σ, in the sense that, for each i, z_i is the ρ_i-th of $z_1, z_2, \ldots, z_r$ as they are ordered by R.

Let us suppose now that $y_1, y_2, \ldots, y_n$ are numbered in the order in which they occur in R, *i.e.* that in the R series y_1 is the first of them, y_2 the second, and so on. Then we shall see that, if ϕ_0 is given the constant interpretation defined above, all the values of ϕ_0 and $\sim \phi_0$ in q will be true. Indeed, for values whose arguments are obtained by permuting $y_1, y_2, \ldots, y_r$ this follows at once from the way in which ϕ_0 has been defined. For $\phi_0(y_{\sigma_1}, y_{\sigma_2}, \ldots, y_{\sigma_r})$ is true if and only if the order of $y_{\sigma_1}, y_{\sigma_2}, \ldots, y_{\sigma_r}$ in the R series is given by a permutation $(\rho_1, \rho_2, \ldots, \rho_r)$ contained in Σ. But the order in the series of $y_{\sigma_1}, y_{\sigma_2}, \ldots, y_{\sigma_r}$ is in fact given (on our present hypothesis that the order of the y's is $y_1, y_2, \ldots, y_r$) by $(\sigma_1, \sigma_2, \ldots, \sigma_r)$, which is contained in Σ if and only if $\phi_0(y_{\sigma_1}, y_{\sigma_2}, \ldots, y_{\sigma_r})$ is contained in q. Hence values of ϕ_0 for arguments consisting of the first r y's are true when they are contained in q and false otherwise, *i.e.* when the corresponding values of $\sim \phi_0$ are contained in q.

For sets of arguments not confined to the first r y's our result follows from the fact that q is serial, *i.e.* that if $s_1 < s_2 < \ldots < s_r$, so that $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ are in the order given by the R series, q contains the values of ϕ_0 for just those permutations of $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ which correspond to the permutations of $y_1, y_2, \ldots, y_r$ for which it contains the values of ϕ_0, *i.e.* by the definition of ϕ_0 and the preceding argument, for just those permutations of $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ which make ϕ_0 true.

Hence all the values of ϕ_0 and $\sim \phi_0$ in q are true when $y_1, y_2, \ldots, y_n$ are in the order given by the R series.

If, then, we define analogous constant interpretations for χ_0, ψ_0, etc., and combine these with our interpretation of ϕ_0, the whole of q will be true provided that $y_1, y_2, \ldots, y_n$ are in the order given by the R series, and if $y_1, y_2, \ldots, y_n$ are in any other order the true alternative will be obtained from q by suitably permuting the y's, *i.e.* will be an alternative similar to q and contained in the same form A_n. Hence A_n is true for any set of distinct $y_1, y_2, \ldots, y_n$. Moreover, for any set of distinct $y_1, y_2, \ldots, y_\nu$ $(\nu < n)$ the true form will be one involved in A_n, and since A_n and all forms involved in it are contained in P, P will be true for these interpretations of $\phi_0, \chi_0, \psi_0, \ldots$, and our formula must be consistent.

Having thus proved our condition for consistency sufficient in any universe, we have now to prove it necessary in any infinite or sufficiently large finite universe, and for this we have to use the Theorem B proved in the first part of the paper.

Our line of argument is as follows: we have to show that, whatever $\phi_0, \chi_0, \psi_0, \ldots$ we take P will be false unless it completely contains a serial A_n. For this it is enough to show that, given any $\phi_0, \chi_0, \psi_0, \ldots$, there must be a set of n y's for which the true form is serial*, or, since a serial form is one which contains a serial alternative, that there must be a set of values of $y_1, y_2, \ldots, y_n$ for which the true alternative is serial.

Let us suppose that among our functions $\phi_0, \chi_0, \psi_0, \ldots$ there are a_1 functions of one variable, a_2 of two variables, ..., and a_n of n variables, and let us order the universe by a serial relation R.

The N individuals in the universe are divided by the a_1 functions of one variable into 2^{a_1} classes according to which of these functions they make true or false, and if $N \geqslant 2^{a_1} k_1$ we can find k_1 individuals which all belong to the same class, *i.e.* agree as to which of the a_1 functions they make true and which false, where k_1 is a positive integer to be assigned later. Let us call this set of k_1 individuals Γ_{k_1}.

Now consider any two distinct members of Γ_{k_1}, z_1 and z_2 say, and let z_1 precede z_2 in the R series. Then in regard to any of the a_2 functions of two variables, ϕ_0 say, there are four possibilities. We may either

* For then P can only be true for $\phi_0, \chi_0, \psi_0, \ldots$ by *completely* containing this true serial form.

have

(1) $\phi_0(z_1, z_2) . \phi_0(z_2, z_1)$,

or (2) $\phi_0(z_1, z_2) . \sim \phi_0(z_2, z_1)$,

or (3) $\sim \phi_0(z_1, z_2) . \phi_0(z_2, z_1)$,

or (4) $\sim \phi_0(z_1, z_2) . \sim \phi_0(z_2, z_1)$.

ϕ_0 thus divides the combinations two at a time of the members of Γ_{k_1} into four distinct classes according to which of these four possibilities is realised when the combination is taken as z_1, z_2 in the order in which its terms occur in the R series; and the whole set of a_2 functions of two variables divide the combinations two at a time of the members of Γ_{k_1} into 4^{a_2} classes, the combinations in each class agreeing in the possibility they realise with respect to each of the a_2 functions. Hence, by Theorem B, if $k_1 = h(2, k_2, 4^{a_2})$, Γ_{k_1} must contain a sub-class Γ_{k_2} of k_2 members such that all the pairs out of Γ_{k_2} agree in the possibilities they realise with respect to each of the a_2 functions of two variables.

We continue to reason in the same way according to the following general form :—

Consider any r distinct members of $\Gamma_{k_{r-1}}$; suppose that in the R series they have the order $z_1, z_2, \ldots, z_r$. Then with respect to any function of r variables there are $2^{r!}$ possibilities in regard to $z_1, z_2, \ldots, z_r$, and the a_r functions of r variables divide the combinations r at a time of the members of $\Gamma_{k_{r-1}}$ into $2^{r!a_r}$ classes. By Theorem B, if $k_{r-1} = h(r, k_r, 2^{r!a_r})$*, $\Gamma_{k_{r-1}}$ must contain a sub-class Γ_{k_r} of k_r members such that all the combinations r at a time of the members of Γ_{k_r} agree in the possibilities they realise with respect to each of the a_r functions of r variables.

We proceed in this way until we reach $\Gamma_{k_{n-1}}$, all combinations $n-1$ at a time of whose members agree in the possibilities they realise with respect to each of the a_{n-1} functions of $n-1$ variables. We then determine that k_{n-1} shall equal n, which fixes k_{n-2} as $h(n-1, n, 2^{(n-1)!a_{n-1}})$ and so on back to k_1, every k_{r-1} being determined from k_r.

If, then, $N \geqslant 2^{a_1} k_1$, the universe must contain a class $\Gamma_{k_{n-1}}$ or Γ_n (since $k_{n-1} = n$) of n members which is contained in Γ_{k_r} for every r, $r = 1, 2, \ldots, n-1$. Let its n members be, in the order given them by R, $y_1, y_2, \ldots, y_n$. Then for every r less than n, $y_1, y_2, \ldots, y_n$ are contained in Γ_{k_r} and all r combinations of them agree in the possibilities they realise

* If $a_r = 0$ we interpret $h(r, k_r, 1)$ as k_r and identify $\Gamma_{k_{r-1}}$ and Γ_{k_r}.

with respect to each function of r variables. Let $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ $(s_1 < s_2 < \ldots < s_r)$ be such a combination, and χ_0 a function of r variables. Then $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ are in the order given them by R, and so are $y_1, y_2, \ldots, y_r$; consequently the fact that these two combinations agree in the possibilities which they realise with respect to χ_0 means that χ_0 is true for the same permutations of $y_{s_1}, y_{s_2}, \ldots, y_{s_r}$ as it is of $y_1, y_2, \ldots, y_r$. The true alternative for $y_1, y_2, \ldots, y_n$ is therefore serial in χ_0, and similarly it is serial in every other function of any number r of variables*; it is therefore a serial alternative.

Our condition is, therefore, shown to be necessary in any universe of at least $2^{a_1} k_1$ members where k_1 is given by

$$k_{n-1} = n,$$

$$\left.\begin{aligned} k_{r-1} &= h(r, k_r, 2^{r!\,a_r}) \quad \text{if} \quad a_r \neq 0 \\ &= k_r \qquad\qquad\qquad \text{if} \quad a_r = 0 \end{aligned}\right\} \quad (r = n-1, n-2, \ldots, 2).$$

For universes lying between n and $2^{a_1} k_1$ we have not found a necessary and sufficient condition for the consistency of the formula, but it is evidently possible to determine by trial whether any given formula is consistent in any such universe.

III.

We will now consider what our result becomes when our formula

$$(x_1, x_2, \ldots, x_n)\, F(\phi, \chi, \psi, \ldots, =, x_1, x_2, \ldots, x_n)$$

contains in addition to identity only one function ϕ of two variables.

In this case we have two functions ϕ_0, ψ_0 given by

$$\phi_0(y_i, y_k) = \phi(y_i, y_k) \quad (i \neq k),$$

$$\chi_0(y_i) \quad\; = \phi(y_i, y_i),$$

so that $a_1 = 1$, $a_2 = 1$, $a_r = 0$ when $r > 2$. Consequently

$$k_2 = k_3 = \ldots = k_{n-1} = n \quad \text{and} \quad k_1 = h(2, n, 4);$$

but the argument at the end of I shows that we may take instead $k_1 = n\,!\,!\,!$, and our necessary and sufficient condition for consistency applies to any universe with at least $2 \,.\, n\,!\,!\,!$ individuals.

In this simple case we can present our condition in a more striking form as follows.

---

* We have shown this when $r < n$; we may also have $r = n$, but then there is nothing to prove since in a function of n variables every alternative is serial.

It is necessary and sufficient for the consistency of the formula that it should be true when ϕ is replaced by at least one of the following types of function :—

(1) The universal function $x = x \,.\, y = y$.

(2) The null function $x \neq x \,.\, y \neq y$.

(3) Identity $x = y$.

(4) Difference $x \neq y$.

(5) A serial function ordering the whole universe in a series, *i.e.* satisfying

(*a*) $(x) \sim \phi(x, x)$,

(*b*) $(x, y)[x = y \vee \{\phi(x, y) \,.\, \sim \phi(y, x)\} \vee \{\phi(y, x) \,.\, \sim \phi(x, y)\}]$,

(*c*) $(x, y, z)\{\sim\phi(x, y) \vee \sim \phi(y, z) \vee \phi(x, z)\}$.

(6) A function ordering the whole universe in a series, but also holding between every term and itself, *i.e.* satisfying

(*a'*) $(x)\,\phi(x, x)$

and (*b*) and (*c*) as in (5).

Types (1)–(4) include only one function each; in regard to types (5) and (6) it is immaterial what function of the type we take, since if one satisfies the formula so, we shall see, do all the others*.

We have to prove this new form of our condition by showing that P will completely contain a serial A_n if and only if it is satisfied by functions of at least one of our six types. Now an alternative in n y's is serial in χ_0 if it contains

either (i) $\chi_0(y_1) \cdot \chi_0(y_2) \cdot \ldots \chi_0(y_n)$ or, for short, $\prod\limits_r \chi_0(y_r)$,

or (ii) $\sim \chi_0(y_1) \cdot \ldots \sim \chi_0(y_n)$,, ,, $\prod\limits_r \sim \chi_0(y_r)$,

but not otherwise, and it will be serial in ϕ_0 if it contains

either (*a*) $\prod\limits_{r<s} \phi_0(y_r, y_s) \cdot \phi_0(y_s, y_r)$,

or (*b*) $\prod\limits_{r<s} \phi_0(y_r, y_s) \cdot \sim \phi_0(y_s, y_r)$,

or (*c*) $\prod\limits_{r<s} \sim \phi_0(y_r, y_s) \cdot \phi_0(y_s, y_r)$,

or (*d*) $\prod\limits_{r<s} \sim \phi_0(y_r, y_s) \cdot \sim \phi_0(y_s, y_r)$.

---

* A result previously obtained for type (5) by Langford, *op. cit.*

There are thus altogether eight alternatives serial in both ϕ_0 and χ_0 got by combining either of (i), (ii) with any of (*a*), (*b*), (*c*), (*d*); but these eight serial alternatives only give rise to six serial forms, since the alternatives (i) (*b*) and (i) (*c*) can be obtained from one another by reversing the order of the y's and so belong to the same form, and so do the alternatives (ii) (*b*) and (ii) (*c*).

It is also easy to see that any formula completely containing one of these six serial forms will be satisfied by all functions of one of the six types according to the scheme

Form	(i) (*a*)	(i) (*b* and *c*)	(i) (*d*)	(ii) (*a*)	(ii) (*b* and *c*)	(ii) (*d*)
Type of function	1	6	3	4	5	2

and that conversely a formula satisfied by a function of one of the six types must completely contain the corresponding form. For instance, a function of type 6 will satisfy the alternative (i) (*b*) when $y_1, y_2, \ldots, y_n$ are in their order in the series determined by the function, and when $y_1, y_2, \ldots, y_n$ are in any other order the function will satisfy an alternative of the same form.

In the language of the theory of postulate systems we can interpret our universe as a class K, and conclude that a postulate system on a base (K, R) consisting only of general laws involving at most n elements will be compatible with K having as many as $2 \, . \, n!!!$ members if and only if it can be satisfied by an R of one of our six types.

IV.

Let us, in conclusion, briefly indicate how to extend our method in order to determine the consistency or inconsistency of formulae of the more general type

$$(Ez_1, z_2, \ldots, z_m)(x_1, x_2, \ldots, x_n)\, F(\phi, \chi, \psi, \ldots, =, z_1, z_2, \ldots, z_m, x_1, x_2, \ldots, x_n)$$

which have in normal form both kinds of prefix, but satisfy the condition that all the prefixes of existence precede all those of generality.

As before, we can suppose F represented as a disjunction of alternatives and discard those which violate the laws of identity. Those left we can group according to the values of identity and difference for arguments drawn entirely from the z's. Such a set of values of identity and difference we can denote by $H_i\,(=, z_1, z_2, \ldots, z_m)$, and F can be put in the form

$$(H_1 \, . \, F_1) \vee (H_2 \, . \, F_2) \vee (H_3 \, . \, F_3) \vee \ldots,$$

and the whole formula is equivalent to a disjunction of formulae.

$$(Ez_1, z_2, \ldots, z_m)\{H_1(=, z_1, z_2, \ldots, z_m) \\ .\, (x_1, x_2, \ldots, x_n)\, F_1(\phi, \ldots, =, z_1, \ldots, z_m, x_1, \ldots, x_n)\}$$

$$\vee\, (Ez_1, z_2, \ldots, z_m)\{H_2(=, z_1, \ldots, z_m) \\ .\, (x_1, x_2, \ldots, x_n)\, F_2(\phi, \ldots, =, z_1, \ldots, z_m, x_1, \ldots, x_n)\}$$

$\vee$ etc.

Since if any one of these formulae is consistent so is their disjunction, and if their disjunction is consistent one at least of its terms must be consistent, it is enough for us to show how to determine the consistency of any one of them, say the first. In this $H_1(=, z_1, z_2, \ldots, z_m)$ is a consistent set of values of identity and difference for every pair of z's. We renumber the z's $z_1, z_2, \ldots, z_\mu$ using the same suffix for every set of z's that are identical in H_1, and our formula becomes

$$(Ez_1, z_2, \ldots, z_\mu)(x_1, x_2, \ldots, x_n)\, F_1(\phi, \chi, \ldots, =, z_1, z_2, \ldots, z_\mu, x_1, x_2, \ldots, x_n), \quad \text{(i)}$$

in which it is understood that two z's with different suffixes are always different.

Now supposing the universe to have at least $\mu+n$ members, we consider the different possibilities in regard to the x's being identical with the z's, and rewrite our formula

$$(Ez_1, z_2, \ldots, z_\mu)(x_1, x_2, \ldots, x_n) \\ \Big\{ \prod_{\substack{i=1,\ldots,n \\ j=1,\ldots,\mu}} x_i \neq z_j \rightarrow G(\phi, \chi, \ldots, =, z_1, \ldots, z_\mu, x_1, \ldots, x_n)\Big\},$$

in which $\rightarrow$ means "if, then" and

$$G(\phi, \ldots, x_n) = \Pi\, F_1(\phi, \chi, \ldots, =, z_1, \ldots, z_\mu, \theta_1, \theta_2, \ldots, \theta_n),$$

the product being taken for

$$\theta_1 = x_1, z_1, z_2, \ldots, z_\mu,$$
$$\theta_2 = x_2, z_1, z_2, \ldots, z_\mu,$$
$$\ldots \quad \ldots \quad \ldots$$
$$\theta_n = x_n, z_1, z_2, \ldots, z_\mu,$$

and in G any term $x_i = z_j$ is replaced by a falsehood (*e.g.* $x_i \neq x_i$) not involving any z.

Next we modify G by introducing new functions. In G occur values of, *e.g.* ϕ, with arguments some of which are z's and some x's; from these we define functions of the x's only by simply regarding the z's as constants, and call these new functions $\phi_0, \chi_0, \ldots$. Values of $\phi, \chi, \psi, \ldots$, which include no x's among their arguments, we replace by constant propositions $p, q, \ldots$. The only values of identity in G are of the form $x_i = x_j$ and these we leave alone. Suppose that by this process G turns into

$$L(\phi, \chi, \psi, \ldots, \phi_0, \chi_0, \ldots, p, q, \ldots, =, x_1, x_2, \ldots, x_n).$$

Then the consistency of formula (i) in a universe of N individuals is evidently equivalent to the consistency in a universe of $N-\mu$ individuals of the formula

$$(x_1, x_2, \ldots, x_n) L(\phi, \chi, \psi, \ldots, \phi_0, \chi_0, \ldots, p, q, \ldots, =, x_1, x_2, \ldots, x_n).$$

But this is a formula of the type previously dealt with, except for the variable propositions $p, q, \ldots$, which are easily eliminated by considering the different cases of their truth and falsity, the formula being consistent if it is consistent in one such case.