DEFINITION OF SUBSTITUTION

BY W. V. QUINE*

1. *Basis*. The *elements* of this study, denoted by italic capitals, comprise *atoms*, at least two and perhaps infinite in number, and all finite *sequences* of such atoms. (The atoms are interpretable as signs, for example, and the sequences as rows of signs.) Thus each element E is composed successively of possibly duplicative atoms $A_1, A_2, \cdots$, and A_m, for some positive integer m, called the *length* of E; and an element F composed successively of atoms B_1 to B_n will be identical with E if and only if $m=n$ and $A_i=B_i$ for each i to m.

Further terminology is self-explanatory. Thus we may speak of the kth *place* of an element; this, in the case of E above, is occupied by the atom A_k. We may speak of one element as *occurring in* another (as a connected segment thereof), and more particularly as occurring *initially*, *internally*, or *terminally* therein; of two elements as occurring *overlapped* in a third; of the number of occurrences of one element in another; and so on.

Juxtaposition will be used to express that binary operation of *concatenation* whereby any elements E and F, composed as above, are put end to end to form that element EF which is composed successively of the atoms $A_1, A_2, \cdots, A_m, B_1, B_2, \cdots$, and B_n. The element E is itself describable, in terms of concatenation of its atoms, as $A_1A_2 \cdots A_m$; parentheses are suppressed, as here, in view of the obvious associativity of concatenation.

It is clear that the length of EF, for any E and F, exceeds the lengths of E and F and equals their sum; also that E occurs initially and F terminally in EF, while EF occurs neither in E nor in F; also that atoms are of length 1, and that an element G is an atom if and only if there are no elements E and F such that $G=EF$.

2. *Substitution*. The purpose of this paper is a formal definition of *substitution* in terms exclusively of concatenation and the following elementary logical devices: identity, applied to

* Society of Fellows, Harvard University.

elements; the truth functions; and quantification with respect to elements. The notation will be as in *Principia Mathematica*: the sign '$=$' for identity, the signs '$\sim$', '$\supset$', $\cdots$, for the truth functions, and prefixes of the forms '(X)' and '$(\exists X)$' for quantification.

The proposition to be formulated is expressible verbally thus: W is the result of substituting X for Y throughout Z; briefly, $sub(W, X, Y, Z)$. When the elements are interpreted as signs and rows of signs, the notion under consideration is the notational substitution which figures so prominently in metamathematics.

In the general form in which substitution is here conceived, its formulation is complicated by the fact that the element Y for which substitution is made need not be an atom and hence may have overlapping occurrences in Z. Obviously we cannot in general replace each of two overlapping occurrences of Y by X, since replacement of one occurrence will mutilate the other occurrence. To this extent the notion of substitution is ambiguous. The ambiguity is resolved by stipulating that in case of overlapping occurrences left is to prevail over right; thus the result of substituting X for TT in TTT is to be XT rather than TX. The result of substituting X for Y throughout Z is then describable, in general, as the result Z' of putting X for each of these occurrences of Y in Z: the first (left-most); the first which begins after the end of that first; the first which begins after the end of this second; and so on. In the trivial case where Y does not occur in Z, Z' is of course Z.

3. *Formal Definitions.* The following abbreviations are adopted:

D1. $\qquad U\ init\ V\ .=_{\mathrm{df}}\ U = V\ .\vee:(\exists T).\ UT = V.$

D2. $\qquad U\ in\ V\ .=_{\mathrm{df}}\ U\ init\ V\ .\vee:(\exists T).\ TU\ init\ V.$

D3. $\qquad U \sim in\ V\ .=_{\mathrm{df}}.\ \sim.\ U\ in\ V.$

D4. $\qquad \Theta(U, V, M, N)\ .=_{\mathrm{df}}.\ MUMVM\ in\ N\ .\ M \sim in\ UV.$

D5. $sub1(U, X, Y, V)\ .=_{\mathrm{df}}: U$
$$=X\ .\ V=Y\ .\vee:(\exists T).\ U=TX\ .\ V=TY:\sim(\exists S).\ YS\ in\ V.$$

Obviously '$U\ init\ V$' may be read 'U occurs initially in V', and

'U *in* V' may be read 'U occurs in V'. Again, '*sub* $1(U, X, Y, V)$' tells us that Y occurs in V terminally and only so and that U is the result of putting X for Y in V; this is seen as follows. For Y to occur terminally in V it is necessary and sufficient that either $V=Y$ or $(\exists T)\cdot V=TY$. Where $V=Y$ it is clear further that Y does not occur in V otherwise than terminally; where $(\exists T)\cdot V=TY$, on the other hand, in order that Y not occur in V otherwise than terminally it is obviously necessary and sufficient to add that $\sim(\exists S)\cdot YS\ in\ V$. In general, therefore, for Y to occur in V terminally and only so it is necessary and sufficient that

$$V = Y\ .\vee:(\exists T).\ V = TY : \sim (\exists S).\ YS\ in\ V.$$

Now if U is the result of putting X for Y in V, U will be X or TX according as V is Y or TY. D5 thus yields the described meaning.

The definition of substitution follows:

$$\begin{aligned}\text{D6.}\quad & sub(W, X, Y, Z)\ . =_{\text{df}}\!:::.\ Y \sim in Z\ .\ W = Z\ .\vee:::.\\ & (\exists G)(\exists H)::: (M)(N)::.\ (U)(V):: sub1(U, X, Y, V).\supset\\ & :.\ V\ init Z\ .\supset.\ \Theta(U, V, M, N):.\ (S)(T):\Theta\ (S, T, M, N)\\ & .\ TV\ init Z\ .\supset.\ \Theta(SU, TV, M, N) ::\supset.\ \Theta(G, H, M, N)\\ & ::: W = G\ .\ Z = H\ .\vee:(\exists K).\ Y \sim in\ K\ .\ W = GK\\ & .\ Z = HK.\end{aligned}$$

When abbreviations introduced by D1–D5 are eliminated in favor of their definientia, the definiens in D6 is seen to involve only concatenation and the elementary logical devices mentioned in §2.

4. *Demonstrandum.* It remains to show that D6 yields substitution in the sense of §2; that is, that $sub(W, X, Y, Z)$, in the sense of D6, if and only if W is Z' as of §2.

Supposing W, X, Y, and Z given as constants, we define as follows:

$$\begin{aligned}\text{Dt1.}\quad & \Phi(M, N)\cdot =_{\text{df}}\!:: (U)(V):: sub1(U, X, Y, V)\cdot\supset:\cdot\\ & V\ init Z\cdot\supset\cdot\ \Theta(U, V, M, N) :\cdot\ (S)(T): \Theta(S, T, M, N)\\ & \cdot\ TV\ init Z\cdot\supset\cdot\ \Theta(SU, TV, M, N).\end{aligned}$$

Dt2. $\Psi(F, G, H) \cdot =_{\mathrm{df}} :\cdot (M)(N): \Phi(M, N) \cdot \supset \cdot \Theta(G, H, M, N)$
$:\cdot F = G \cdot Z = H \cdot \vee : (\exists K) \cdot Y \sim in\, K \cdot F = GK$
$\cdot Z = HK.$

By §1, $(U)(V) \cdot XUXVX \sim in\ X$; by D4, then, $(U)(V) \cdot \sim\Theta(U, V, X, X)$, so that

$$(1) \quad (U)(V) :\cdot V\ init\, Z \cdot \supset \cdot \Theta(U, V, X, X) : \equiv \cdot \sim \cdot V\ init\, Z,$$

$$(2) \quad (G)(H) :\cdot \Phi(X, X) \cdot \supset \cdot \Theta(G, H, X, X) : \equiv \cdot \sim \Phi(X, X),$$

and, trivially,

$$(3) \quad (U)(V)(S)(T): \Theta(S, T, X, X) \cdot TV\ init\, Z \cdot \supset \cdot \Theta(SU, TV, X, X).$$

By Dt1, (1), and (3),

$$\Phi(X, X) \cdot \equiv : (U)(V): sub1(U, X, Y, V) \cdot \supset \cdot \sim \cdot V\ init\, Z,$$

whence $\sim\Phi(X, X) \cdot \supset \cdot (\exists U)(\exists V) \cdot sub1(U, X, Y, V) \cdot V\ init\ Z$, and therefore, by D5,

$$\sim \Phi(X, X) \cdot \supset :\cdot (\exists V) :\cdot V = Y \cdot \vee : (\exists T) \cdot V = TY :\cdot V\ init\, Z,$$

that is, $\sim\Phi(X, X) \cdot \supset : Y\ init\ Z \cdot \vee : (\exists T) \cdot TY\ init\ Z$, which is to say, by D2,

$$(4) \quad \sim \Phi(X, X) \cdot \supset \cdot Y\ in\, Z.$$

By Dt2, $(G)(H) : \cdot \Psi(W, G, H) \cdot \supset : \Phi(X, X) \cdot \supset \cdot \Theta(G, H, X, X)$ whence, by (2) and (4),

$$(G)(H): \Psi(W, G, H) \cdot \supset \cdot Y\ in\, Z,$$

that is, $(\exists G)(\exists H) \cdot \Psi(W, G, H) \cdot \supset \cdot Y\ in\ Z$, or, equivalently,

$$(5) \quad (\exists G)(\exists H) \cdot \Psi(W, G, H) \cdot \equiv : Y\ in\, Z : (\exists G)(\exists H) \cdot \Psi(W, G, H).$$

If Y does not occur in Z, then, by §2, Z' is Z. Hence

$$(6) \quad Y \sim in\, Z \cdot W = Z \cdot \equiv \cdot Y \sim in\, Z \cdot W = Z'.$$

By D6, Dt1, and Dt2,

$$sub(W, X, Y, Z) \cdot \equiv : Y \sim in\, Z \cdot W = Z \cdot \vee : (\exists G)(\exists H) \cdot \Psi(W, G, H).$$

Hence, by (5) and (6),

$$\text{(7)}\qquad \begin{aligned} sub(W, X, Y, Z) \cdot\equiv: Y \sim in\, Z \cdot W = Z' \cdot\vee: Y\ in\, Z \\ : (\exists G)(\exists H)\cdot \Psi(W, G, H). \end{aligned}$$

Now if we can prove that

$$\text{(I)}\qquad Y\ in\, Z \cdot\supset: (\exists G)(\exists H)\cdot \Psi(W, G, H)\cdot\equiv\cdot W = Z',$$

so that $Y\ in\, Z: (\exists G)(\exists H)\cdot\Psi(W, G, H): \equiv\cdot Y\ in\, Z\cdot W = Z'$, then from (7) we shall have

$$sub(W, X, Y, Z) \cdot\equiv\cdot Y \sim in Z \cdot W = Z' \cdot\vee\cdot Y\ in Z \cdot W = Z',$$

that is, $sub\ (W, X, Y, Z)\cdot\equiv\cdot W = Z'$ which was to be proved. It thus remains only to establish (I).

5. *Proof of* (I). Given that Y occurs in Z, it is to be proved that

$$(\exists G)(\exists H)\cdot \Psi(W, G, H) \cdot\equiv\cdot W = Z'.$$

Consider the following segments Z_i of Z; Z_1 extends from the beginning of Z to the end of the first occurrence of Y in Z; Z_2 extends from the beginning of Z^1 to the end of the first occurrence of Y in Z^1, where Z^1 is Z deprived of its initial segment Z_1; and, in general, Z_{i+1} extends from the beginning of Z^i to the end of the first occurrence of Y in Z^i, where Z^i is Z deprived of its initial segment $Z_1Z_2 \cdots Z_i$. By construction, Y occurs in each Z_i terminally and only so; in view of §3, then, where the Z_i' are the results of putting X for Y in the respective Z_i,

$$\text{(8)}\qquad (i)\cdot\ sub1(Z_i', X, Y, Z_i).$$

Let '$Z_1Z_2 \cdots Z_i$' and '$Z_1'Z_2' \cdots Z_i'$' be written '$Z_i!$' and '$Z_i'!$'. Thus

$$\text{(9)}\qquad Z_1! = Z_1 \cdot Z_1'! = Z_1',$$

$$\text{(10)}\qquad (i)\cdot Z_{i+1}! = Z_i!Z_{i+1} \cdot Z_{i+1}'! = Z_i'!Z_{i+1}',$$

and, by construction,

$$\text{(11)}\qquad (i)\cdot Z_i!\ init\, Z.$$

We exhaust the segments Z_i only when we reach a point in Z beyond which Y occurs no more. Thus, where Z_n is the last of the Z_i,

(12) $$(K): Z = Z_n!K \cdot \supset \cdot Y \sim in\ K.$$

Quantification with respect to subscripts, as in (8)–(11), refers of course only to the n or fewer significant values; for example, '$(i) \cdot \phi(Z_i, Z_i')$', '$(\exists i) \cdot \phi(Z_i, Z_i')$', and '$(i) \cdot \phi(Z_i, Z_{i+1})$' are short for

$$\text{`}\phi(Z_1, Z_1') \cdot \phi(Z_2, Z_2') \cdot \ldots \phi(Z_n, Z_n')\text{'},$$

$$\text{`}\phi(Z_1, Z_1') \vee \phi(Z_2, Z_2') \vee \ldots \phi(Z_n, Z_n')\text{'},$$

and

$$\text{`}\phi(Z_1, Z_2) \cdot \phi(Z_2, Z_3) \cdot \ldots \phi(Z_{n-1}, Z_n)\text{'}.$$

By §1, there are distinct atoms; let A and B be any two such, and, where k is the length of $Z_n'!Z_n!$, let $C = ABB \cdots B$ to k occurrences of B. Clearly

(13) $$(i) \cdot C \sim in\ Z_i'!Z_i!,$$

since the length of C is greater by one even than that of $Z_n'!Z_n!$. Now it will be shown that, for any elements G and H such that $C \sim in\ GH$, there are no occurrences of C in $CGCHC$ except the three indicated ones. First, no two occurrences of C can overlap; for, if they are distinct, one must start later than the other and hence must start at a non-initial place of the other; but all these non-initial places are occupied by B, whereas C starts with A. Hence if there is an occurrence of C in $CGCHC$ other than the three indicated ones, it overlaps none of the latter; it therefore lies wholly within G or H. But it cannot, since, by hypothesis, $C \sim in\ GH$. Consequently there are none but the three occurrences of C in $CGCHC$. In particular, then, it follows from (13) that there are none but the three occurrences of C in $CZ_i'!CZ_i!C$.

Since, where

(14) $$D = CZ_1'!CZ_1!CCZ_2'!CZ_2!C \cdots CZ_n'!CZ_n!C,$$

D is made up wholly of segments of the form $CZ_i'!CZ_i!C$, any occurrence of C in D must lie either wholly within or partly within and partly beyond such a segment. But in the latter case the occurrence of C in question would overlap the terminal occurrence of C in $CZ_i'!CZ_i!C$, whereas we saw that no two occurrences of C could overlap. Hence every occurrence of C in D lies wholly within $CZ_i'!CZ_i!C$ for some i. But $CZ_i'!CZ_i!C$ was seen to contain only the three indicated occurrences of C. Therefore D contains only the $3n$ occurrences of C indicated in (14).

Where $C \sim in\ GH$, $CGCHC$ contains, as was seen, only the three indicated occurrences of C, and is hence describable as beginning and ending with C and containing just three occurrences of C and no two adjacent. But, since D contains none but its $3n$ indicated occurrences of C, inspection of (14) shows that the only segments of D fulfilling this description of $CGCHC$ are the segments $CZ_i'\,!CZ_i!C$ for the various i; any other segment of D beginning and ending with C and containing three occurrences of C would contain two adjacent. Hence, if $C \sim in\ GH$ and $CGCHC\ in\ D$, $CGCHC$ must be $CZ_i'\,!CZ_i!C$ for some i; then G and H must be $Z_i'\,!$ and $Z_i!$. Thus, in view of D4,

(15) $(G)(H)\colon \Theta(G, H, C, D)\ .\supset .\ (\exists i)\cdot G = Z_i'\,!\ .\ H = Z_i!.$

Conversely, by (14), (13), and D4,

(16) $(i).\ \Theta(Z_i'\,!, Z_i!, C, D).$

By (15) and (16),

(17)* $(G)(H)\colon \Theta(G, H, C, D)\ .\equiv .\ (\exists i).\ G = Z_i'\,!\ .\ H = Z_i!.$

If $sub1(U, X, Y, V)$, then, by §3, Y occurs in V terminally and only so and U is the result of putting X for Y in V. But, if V contains Y just thus and if further $Z_i!V\ init\ Z$, then V must be Z_{i+1}, so that U becomes Z_{i+1}'; and where U and V are Z_{i+1}' and Z_{i+1} it follows from (10) and (16) that $\Theta(Z_i'\,!U, Z_i!V, C, D)$. Thus

* This exemplifies a general technique, within a concatenation system, for eliminating reference to a finite class or relation of elements in favor of reference to two properly selected elements C and D. Where α is the m-adic relation (or class, if $m=1$) exhibited by just the elements Q_{i1}, Q_{i2}, $\cdots$, and Q_{im} in that order, for the various i from 1 to n, and k is the length of the element

$$Q_{11}Q_{12}\cdots Q_{1m}Q_{21}Q_{22}\cdots Q_{2m}\cdots Q_{n1}Q_{n2}\cdots Q_{nm},$$

and C is $ABB\cdots B$ to k occurrences of B, and D is

$$CQ_{11}CQ_{12}\cdots CQ_{1m}CCQ_{21}CQ_{22}\cdots CQ_{2m}C\cdots CQ_{n1}CQ_{n2}\cdots CQ_{nm}C,$$

it can be shown that

$$CG_1CG_2\cdots CG_mC \text{ in } D\cdot C\sim in\ G_1G_2\cdots G_m$$

if and only if G_1, G_2, $\cdots$, and G_m exhibit in that order the relation α. In (17) this equivalence is proved for the special case where $m=2$; G, H, $Z_i'\,!$, and $Z_i!$ answer to G_1, G_2, Q_{i1}, and Q_{i2}.

$$(U)(V)(i)\colon sub1(U, X, Y, V) . Z_i ! V \, init \, Z . \supset . \Theta(Z_i{}' ! U, Z_i ! V, C, D),$$

and hence

$$(U)(V) \colon . \; sub1(U, X, Y, V) \; . \supset \colon$$
$$(S)(T)\colon (\exists i) . \; S = Z_i' ! . \; T = Z_i ! . \; TV \; init Z \; . \supset . \; \Theta(SU, TV, C, D);$$

that is, by (17),

$$(18)\quad \begin{array}{l} (U)(V) \colon . \; sub1(U, X, Y, V) \; . \supset \colon \\ \quad (S)(T)\colon \Theta(S, T, C, D) \; . \; TV \; init Z \; . \supset . \; \Theta(SU, TV, C, D). \end{array}$$

Again, if Y occurs in V terminally and only so, and V *init* Z, then V must be Z_1; thus, where $sub1(U, X, Y, V)$ and V *init* Z, U and V will be Z_1' and Z_1. But, by (9) and (16), $\Theta(Z_1', Z_1, C, D)$. Thus

$$(U)(V)\colon sub1(U, X, Y, V) \; . \; V \; init Z \; . \supset . \; \Theta(U, V, C, D).$$

From this and (18) it follows, by Dt1, that $\Phi(C, D)$. By Dt2, then,

$$(G)(H) \colon\colon \Psi(W, G, H) \; . \supset \colon . \; \Theta(G, H, C, D) \colon .$$
$$W = G \; . \; Z = H \; . \vee \colon (\exists K) . \; Y \sim in \; K \; . \; W = GK \; . \; Z = HK,$$

that is, by (17),

$$(G)(H) \colon\colon \Psi(W, G, H) \; . \supset \colon . \; (\exists i) . \; G = Z_i' ! \; . \; H = Z_i ! \colon .$$
$$W = G \; . \; Z = H \; . \vee \colon (\exists K) . \; Y \sim in \; K \; . \; W = GK \; . \; Z = HK,$$

and hence

$$(\exists G)(\exists H) . \; \Psi(W, G, H) \; . \supset \colon (\exists i)\colon W = Z_i' ! \; . \; Z = Z_i ! \; . \vee \colon$$
$$(\exists K) . \; Y \sim in \; K \; . \; W = Z_i' ! K \; . \; Z = Z_i ! K.$$

But, whether $Z = Z_i!$ or $(\exists K) . \; Y \sim in \; K . \; Z = Z_i!K$, i must be n: for in neither case does Z contain any occurrence of Y after $Z_i!$. Thus

$$(19)\quad \begin{array}{l} (\exists G)(\exists H) . \; \Psi(W, G, H) \; . \supset . \; W = Z_n' ! \; . \; Z = Z_n ! \; . \vee \colon \\ \qquad\qquad (\exists K) . \; W = Z_n' ! K \; . \; Z = Z_n ! K. \end{array}$$

Of the occurrences of Y in Z, the one in Z_1 is, by construction, the first, the one in Z_2 is the first which begins after the end of that first, and so on. Therefore if each Z_i is supplanted in Z by

Z_i', so that $Z_n!$ is supplanted by $Z_n'!$, the result will be Z' as described in §2. Z' is thus $Z_n'!$ or $Z_n'!K$ according as Z is $Z_n!$ or $Z_n!K$. By (19), then,

$$(20) \qquad (\exists G)(\exists H) \,.\, \Psi(W, G, H) \,.\supset.\, W = Z'.$$

If $\Phi(M, N)$, then, by Dt1,

$$sub1(Z_1', X, Y, Z_1) \,.\supset:\, Z_1 \; init \; Z \,.\supset.\, \Theta(Z_1', Z_1, M, N)$$

and

$$(i):.\, sub1(Z_{i+1}', X, Y, Z_{i+1}) \,.\supset:\, \Theta(Z_i'!, Z_i!, M, N) \,. \\ Z_i!Z_{i+1} \; init \; Z \,.\supset.\, \Theta(Z_i'!Z_{i+1}', Z_i!Z_{i+1}, M, N).$$

But these two results reduce, in view of (8)–(11), to

$$\Theta(Z_1'!, Z_1!, M, N)$$

and

$$(i):\, \Theta(Z_i'!, Z_i!, M, N) \,.\supset.\, \Theta(Z_{i+1}'!, Z_{i+1}!, M, N),$$

and from these it follows that $\Theta(Z_n'!, Z_n!, M, N)$. Thus

$$(21) \qquad (M)(N):\, \Phi(M, N) \,.\supset.\, \Theta(Z_n'!, Z_n!, M, N).$$

By (11) and D1, Z is $Z_n!$ or else $Z_n!K$ for some K; and, as seen above, Z' is in these respective cases $Z_n'!$ and $Z_n'!K$. Moreover, in view of (12), $Y \sim in \; K$. Thus

$$Z' = Z_n'! \,.\, Z = Z_n! \,.\vee:\, (\exists K) \,.\, Y \sim inK \,.\, Z' = Z_n'!K \,.\, Z = Z_n!K.$$

From this and (21) it follows, by Dt2, that $\Psi(Z', Z_n'!, Z_n!)$. Hence

$$W = Z' \,.\supset.\, (\exists G)(\exists H) \,.\, \Psi(W, G, H),$$

and consequently, by (20),

$$(\exists G)(\exists H) \,.\, \Psi(W, G, H) \,.\equiv.\, W = Z',$$

which was to be proved.

HARVARD UNIVERSITY