# A VARIANT OF A RECURSIVELY UNSOLVABLE PROBLEM

EMIL L. POST

By a string on $a$, $b$ we mean a row of $a$'s and $b$'s such as *baabbbab*. It may involve only $a$, or $b$, or be null. If, for example, $g_1$, $g_2$, $g_3$ represent strings *bab*, *aa*, *b* respectively, string $g_2g_1g_1g_3g_2$ on $g_1$, $g_2$, $g_3$ will represent, in obvious fashion, the string *aababbabbaa* on $a$, $b$. By the *correspondence decision problem* we mean the problem of determining for an arbitrary finite set $(g_1, g_1')$, $(g_2, g_2')$, $\cdots$, $(g_\mu, g_\mu')$ of pairs of corresponding non-null strings on $a$, $b$ whether there is a solution in $n$, $i_1$, $i_2$, $\cdots$, $i_n$ of equation

$$(1) \qquad g_{i_1}g_{i_2}\cdots g_{i_n} = g'_{i_1}g'_{i_2}\cdots g'_{i_n}, \qquad n \geqq 1,\ i_j = 1, 2, \cdots, \mu.$$

That is, whether some non-null string on $g_1$, $g_2$, $\cdots$, $g_\mu$, and corresponding string on $g_1'$, $g_2'$, $\cdots$, $g_\mu'$, represent identical strings on $a$, $b$.

In special cases, of course, the question posed by (1) may be answerable. Thus, if, with $\mu=3$, $(g_1, g_1')$, $(g_2, g_2')$, $(g_3, g_3')$ are $(bb, b)$, $(ab, ba)$, $(b, bb)$ respectively, $g_1g_2g_2g_3 = bbababb = g_1'g_2'g_2'g_3'$, and (1) has a solution. Again, if each $g_i$ is of greater length than the corresponding $g_i'$, or if each $g_i$ starts with a different letter than the corresponding $g_i'$, (1) has no solution. We proceed to prove, on the other hand, that in its full generality *the correspondence decision problem is recursively unsolvable*,[1] and hence, no doubt, unsolvable in the intuitive sense.

We start with the known recursive unsolvability of the decision problem for the class of normal systems on $a$, $b$.[2] A normal system $S$ on

Received by the editors October 20, 1945.

[1] It suffices here to consider "recursively unsolvable" to mean unsolvable in the sense of Church [1]. Of course the general problem remains recursively unsolvable if we allow null $g$'s and $g''$s. Numbers in brackets refer to the references cited at the end of the paper.

[2] See [4, §2] for an informal proof. As far as the printed literature is concerned, we must refer to [2] for a formal proof, though there then remains the actual verification, via Gödel representations, that the reduction effected is indeed recursive. This verification, at least for the reduction of $S'$ to $S'''$ [2, p. 51], is immediate if we use the following simpler method of reducing $S'$ to a system $S''$ in canonical form than that there given by Church. The primitive symbols of our $S''$ are those of $S'$ and one additional primitive symbol $\alpha$. The basis of $S''$ in part consists of the two primitive assertions $\alpha I$, $\alpha J$, and the operation $\alpha P$, $\alpha Q$ *produce* $\alpha(PQ)$. It will follow that $\alpha P$ is asserted in $S''$ when and only when $P$ is a combination without free variables. The remainder of the basis of $S''$ consists of the primitive assertion of $S'$ as primitive assertion, and the thirty-eight operations of $S'$ each modified as follows. For each operational variable $P$ occurring in the operation, $\alpha P$ is introduced as additional premise.

$a$, $b$ is given by a basis consisting of an initial non-null string $A$ on $a$, $b$, and a finite set of operations $\alpha_i P$ *produces* $P\alpha_i'$, $i=1, 2, \cdots, \nu$, where the $\alpha$'s and $\alpha'$'s are given strings on $a$, $b$, while the operational variable $P$ represents an arbitrary string on $a$, $b$, possibly null. The assertions of $S$ consist of $A$ and all non-null strings obtainable from $A$ by repeated use of the $\nu$ operations. The known recursively unsolvable problem is then the problem of determining for arbitrary $S$, as given by a basis therefore, and arbitrary non-null string $B$ on $a$, $b$, whether $B$ is an assertion of $S$. This unsolvability is undisturbed if the $\alpha$'s and $\alpha'$'s are all non-null,[3] a condition which automatically excludes the possibility of null assertions, and will henceforth be assumed.

Referring to operation $\alpha_i P$ *produces* $P\alpha_i'$ by the subscript $i$, string $B$ on $a$, $b$ will be an assertion of $S$ when and only when some finite sequence of operations $i_1, i_2, \cdots, i_n$ leads from $A$ to $B$. Now operation $\alpha_i P$ *produces* $P\alpha_i'$ can be applied to string $C$ to yield string $D$ when and only when for some string $P$, possibly null, $C=\alpha_i P$, $P\alpha_i'=D$. Hence $B$ is an assertion of $S$ when and only when the following set of equations has a solution in $n, i_1, i_2, \cdots, i_n$, and the $P$'s.

$$(2) \quad A = \alpha_{i_1}P_1,\ P_1\alpha'_{i_1} = \alpha_{i_2}P_2, \cdots, P_{n-1}\alpha'_{i_{n-1}} = \alpha_{i_n}P_r,\ P_n\alpha'_{i_n} = B.$$

Here $n$ may be 0, (2) then becoming $A=B$. We proceed to show that (2) is equivalent to a single equation somewhat like (1) subject, however, to certain length conditions.

Given (2), we can eliminate the $P$'s by forming $A\alpha'_{i_1}\alpha'_{i_2}\cdots\alpha'_{i_n}$ and successively substituting for the left members of (2) the right to obtain

$$(3) \qquad A\alpha'_{i_1}\alpha'_{i_2}\cdots\alpha'_{i_n} = \alpha_{i_1}\alpha_{i_2}\cdots\alpha_{i_n}B.$$

Likewise, starting with $A\alpha'_{i_1}\alpha'_{i_2}\cdots\alpha'_{i_{m-1}}$, we obtain

$$(4) \qquad A\alpha'_{i_1}\cdots\alpha'_{i_{m-1}} = \alpha_{i_1}\cdots\alpha_{i_m}P_m,$$

whence,

$$(5) \quad \textit{length}\ (A\alpha'_{i_1}\cdots\alpha'_{m-1}) \geqq \textit{length}\ (\alpha_{i_1}\cdots\alpha_{i_m}), \quad m = 1, 2, \cdots, n,$$

the length of a string being the total number of occurrences of letters therein, here $a$'s and $b$'s. Conversely, let (3) be given, with (5) satisfied. With the length of $\alpha_{i_1}\cdots\alpha_{i_m}$ less than or equal to that of $A\alpha'_{i_1}\cdots\alpha'_{i_{m-1}}$, (3) shows that the former must be identical with an

[3] It suffices to modify the production starting on page 214 of [3] in accordance with footnote 3 thereof to insure that the final normal system has no $g$ or $g'$ null.

"initial segment" of the latter. Hence, $P_m$ can be determined so that (4) is satisfied, and for $m=1, 2, \cdots, n$. For $m=1$, (4) yields the first equation of (2). By substituting the right side of (4), with $m=j$, for the left, (4) for $m=j+1$ becomes $\alpha_{i_1} \cdots \alpha_{i_j} P_j \alpha_{i_j} = \alpha'_{i_1} \cdots \alpha_{i_j} \alpha_{i_{j+1}} P_{j+1}$, whence $P_j \alpha'_j = \alpha_{i_{j+1}} P_{j+1}$, $j=1, 2, \cdots, n-1$. Likewise, the last equation of (2) is obtained from (3) via (4) for $m=n$. Hence, (2) has a solution when and only when (3) has a solution subject to (5). That is, $B$ is in normal system $S$ when and only when (3) has a solution in $n$, $i_1$, $i_2$, $\cdots$, $i_n$ subject to (5). Comparing (3) with (1), we see that to reduce the decision problem for the class of normal systems on $a$, $b$ to the correspondence decision problem, and thus have the unsolvability of the former lead to the unsolvability of the latter, we must on the one hand eliminate the length condition (5), on the other, the $A$ and $B$ of (3).

We achieve the first aim by reducing normal system $S$ in three stages to a normal system in which (3) implies (5). If $C=x_1x_2 \cdots x_n$, the $x$'s $a$'s or $b$'s, let $\overline{C}=x_n \cdots x_2x_1$. For a letter with subscript, superscript, we shall only bar the letter. Now for the normal system $S$ on $a$, $b$ with initial string $A$ and operations $\alpha_i P$ *produces* $P\alpha'_i$, $i=1, 2, \cdots, \nu$, form the system $S'$, not normal, with initial string $\overline{A}$, and operations $P\bar{\alpha}_i$ *produces* $\bar{\alpha}'_i P$. Clearly, string $B$ on $a$, $b$ will be an assertion of $S$ when and only when $\overline{B}$ is an assertion of $S'$. Next form $S''$ with initial string $\overline{A}h$, and operations $P\bar{\alpha}_i h$ *produces* $\bar{\alpha}'_i Ph$. String $\overline{B}$ is then in $S'$ when and only when $\overline{B}h$ is in $S''$. We finally form a normal system $S'''$, though on the three letters $a$, $b$, $h$, whose assertions are the assertions of $S''$ and all cyclic permutations thereof, a cyclic permutation of string $x_1 \cdots x_j x_{j+1} \cdots x_n$ here meaning any string $x_{j+1} \cdots x_n x_1 \cdots x_j$. The initial string of $S'''$ is again $\overline{A}h$. For its first $\nu$ operations we take $\bar{\alpha}_i hP$ *produces* $Ph\bar{\alpha}'_i$, premise and conclusion being a cyclic permutation of the premise and conclusion of the corresponding operation of $S''$. We finally add the operations $aP$ *produces* $Pa$, $bP$ *produces* $Pb$, $hP$ *produces* $Ph$ which serve to transform a string on $a$, $b$, $h$ into any of its cyclic permutations. System $S'''$ is therefore normal, and, by induction, is easily seen to have the stated property.[4] It follows that $B$ is an assertion of normal system $S$ when and only when $\overline{B}h$ is an assertion of normal system $S'''$.

Let the operations of $S'''$ be resymbolized $\beta_i P$ *produces* $P\beta'_i$, $i=1, 2, \cdots, \nu+3$. Though $S'''$ is a normal system on three letters, the discussion of equations (2)–(5) is equally applicable to it. Hence $B$ is an assertion of $S$ when and only when the following equation (6) has a solution subject to (7).

[4] Cf. [3], final reduction.

$$\overline{A}h\beta'_{i_1}\beta'_{i_2}\cdots\beta'_{i_n} = \beta_{i_1}\beta_{i_2}\cdots\beta_{i_n}\overline{B}h. \tag{6}$$

$$\textit{length}\,(\overline{A}h\beta'_{i_1}\cdots\beta'_{i_{m-1}}) \geqq \textit{length}\,(\beta_{i_1}\cdots\beta_{i_m}),\ m = 1, 2, \cdots, n. \tag{7}$$

Suppose (6) had a solution with (7) not satisfied for a certain $m$. For that $m$, the length of $\beta_{i_1}\cdots\beta_{i_m}$ would exceed the length of $\overline{A}h\beta'_{i_1}\cdots\beta'_{i_{m-1}}$, and hence, in virtue of (6), we would have

$$\beta_{i_1}\cdots\beta_{i_m} = \overline{A}h\beta'_{i_1}\cdots\beta'_{i_{m-1}}Q \tag{8}$$

with non-null $Q$. Recall that $(\beta_i, \beta'_i)$ is $(\bar{\alpha}_i h, h\bar{\alpha}'_i)$ for $i \leqq \nu$, $(a, a)$, $(b, b)$, $(h, h)$ for the three remaining $i$'s. With $\alpha$'s and $\alpha'$'s on $a$, $b$ only, $\beta_i$ and $\beta'_i$ are then either both free from $h$, or have exactly one $h$ apiece. Were the $\beta_{i_m}$ of (8) $a$ or $b$, the right side of (8) would have at least one more occurrence of $h$ than the left, which is impossible. In any other case, $\beta_{i_m}$ ends with $h$. Non-null $Q$ therefore ends with $h$, and again the right side of (8) would have at least one more $h$ than the left. Hence, every solution of (6) must satisfy (7). That is, $B$ is an assertion of $S$ when and only when (6) has a solution.

The elimination of $\overline{A}h$ and $\overline{B}h$ from (6) is more easily effected. Corresponding to the $\nu+3$ couples $(\beta_i, \beta'_i)$, $i=1, 2, \cdots, \nu+3$, and $\overline{A}h$, $\overline{B}h$, we introduce $\nu+5$ couples $(\gamma_i, \gamma'_i)$ as follows. With $x$'s and $y$'s representing letters, in this case $a$, $b$, or $h$, if $\beta_i$ and $\beta'_i$ are $x_1x_2\cdots x_\kappa$ and $y_1y_2\cdots y_\lambda$ respectively, $\gamma_i$ and $\gamma'_i$ are to be $x_1kx_2k\cdots x_\kappa k$ and $ky_1ky_2\cdots ky_\lambda$ respectively. If $\overline{A}h$ and $\overline{B}h$ are $y_1y_2\cdots y_\lambda$ and $x_1x_2\cdots x_\kappa$ respectively, $\gamma_{\nu+4}$ and $\gamma'_{\nu+4}$ are to be $kk$ and $kky_1ky_2\cdots ky_\lambda$, $\gamma_{\nu+5}$ and $\gamma'_{\nu+5}$, $x_1kx_2k\cdots x_\kappa kk$ and $kk$, respectively. It then follows that (6) has a solution in $n$, $i_1$, $i_2$, $\cdots$, $i_n$, $n\geqq 0$, $i_p=1, 2, \cdots, \nu+3$, when and only when the equation

$$\gamma'_{j_1}\gamma'_{j_2}\cdots\gamma'_{j_m} = \gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_m} \tag{9}$$

has a solution in $m$, $j_1$, $j_2$, $\cdots$, $j_m$, $m\geqq 1$, $j_q=1, 2, \cdots, \nu+5$. In fact, if $i_1, i_2, \cdots, i_n$ makes both sides of (6) equal to $z_1z_2\cdots z_l$, $(j_1, j_2, \cdots, j_m)=(\nu+4, i_1, i_2, \cdots, i_n, \nu+5)$ makes both sides of (9) equal to $kkz_1kz_2\cdots kz_lkk$. On the other hand, suppose (9) has a solution. Then since $\gamma'_{j_1}$ and $\gamma_{j_1}$ must start with the same letter, $j_1=\nu+4$. For in every other case $\gamma'_{j_1}$ starts with $k$, $\gamma_{j_1}$ with $a$, $b$, or $h$. Similarly, $j_m=\nu+5$, (9) forcing $\gamma'_{j_m}$ and $\gamma_{j_m}$ to end with the same letter. If, now, the intermediate $j$'s are all different from $\nu+4$ and $\nu+5$, they directly give a sequence of $i$'s satisfying (6). Otherwise, let $j_\mu$ be the first $j$ beyond $j_1$ that is $\nu+4$ or $\nu+5$. Were $j_\mu=\nu+4$, $\gamma'_{j_1}\gamma'_{j_2}\cdots\gamma'_{j_\mu}$ and $\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_\mu}$ would take the forms $kkx_1kx_2\cdots kx_pkkx_{p+1}k\cdots x_q$ and $kky_1ky_2k\cdots y_rkkk$ respectively, with $x$'s and $y$'s $a$, $b$, or $h$. But then the second occurrence of $kk$ in the left side of (9) would be immediately followed by $a$, $b$, or $h$, in the right side, by $k$, contradicting (9). Hence $j_\mu=\nu+5$, and $\gamma'_{j_1}\gamma'_{j_2}\cdots\gamma'_{j_\mu}$ and $\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_\mu}$ consequently take the form $kkx_1kx_2\cdots kx_pkk$ and $kky_1ky_2\cdots ky_qkk$. But the left side of (9) through the second occurrence of $kk$ must equal the right side of (9) through its second occurrence of $kk$. Hence $\gamma'_{j_1}\gamma'_{j_2}\cdots\gamma'_{j_\mu}=\gamma_{j_1}\gamma_{j_2}\cdots\gamma_{j_\mu}$, and we have a solution of (9) of the type previously seen to lead to a solution of (6). It follows that $B$ is an assertion of $S$ when and only when (9) has a solution.

In the reduction thus effected we have introduced the new letters $h$ and $k$. But now in $(\gamma_i, \gamma'_i)$ replace the letters $a$, $b$, $h$, $k$ by *bab*, *baab*, *baaab*, *baaaab* respectively,[5] and call the resulting pair of strings on $a$, $b$, $(\delta_i, \delta'_i)$. Then (9) is seen to be equivalent to

$$(10) \qquad \delta'_{j_1}\delta'_{j_2}\cdots\delta'_{j_m}=\delta_{j_1}\delta_{j_2}\cdots\delta_{j_m},$$

immediately so in passing from (9) to (10), and conversely. For if, for example, $\delta'_{j_1}$ starts with *baab*, $\delta_{j_1}$ must also start with *baab*, and likewise for the next group of letters, and so on till (10) is seen to be a translation of (9).

Given normal system $S$ on $a$, $b$ with basis $A$, $\alpha_iP$ *producesP*$\alpha'_i$, $i=1, 2, \cdots, \nu$, and string $B$ on $a$, $b$, the above gives an effective method for forming the pairs of strings on $a$, $b$, $(\delta_i, \delta'_i)$, $i=1, 2, \cdots, \nu+5$, such that $B$ is an assertion of $S$ when and only when (10) has a solution. But (10), with left and right hand members interchanged, is a case of (1). We have therefore effectively reduced the decision problem for the class of normal systems on $a$, $b$ to our correspondence decision problem. If, then, the former is unsolvable in the intuitive sense, so must be the latter. Actually, by introducing Gödel representations, we readily verify that the above effective reduction is indeed recursive, the recursive unsolvability of the former problem then leading to the recursive unsolvability of the correspondence decision problem.

## References

**1.** Alonzo Church, *An unsolvable problem of elementary number theory*, Amer. J. Math. vol. 58 (1936) pp. 345–363.

**2.** ———, Review of [3], Journal of Symbolic Logic vol. 8 (1943) pp. 50–52.

**3.** Emil L. Post, *Formal reductions of the general combinatorial decision problem*, Amer. J. Math. vol. 65 (1943) pp. 197–215.

**4.** ———, *Recursively enumerable sets of positive integers and their decision problems*, Bull. Amer. Math. Soc. vol. 50 (1944) pp. 284–316.

The City College,
New York City

[5] Cf. [4], footnote 5.