RICHARD LAING

ANOMALIES OF SELF-DESCRIPTION*

1. INTRODUCTION

Theoretical psychology, broadly conceived, is the study of all mind-like behavior wherever it may arise. In this sense, the subject matter of theoretical psychology includes *artificial* mental mechanisms as well as those mental mechanisms which have arisen naturally (or may yet so arise). Automaton theory is the abstract and precise study of all discrete deterministic devices and processes. In so far as mental mechanisms (natural or artificial) can be explained in discrete deterministic terms, automaton theory is a convenient and valuable source of precise concepts for a broadly conceived theoretical psychology.

A putative important function of mental mechanisms is the control, guidance, and regulation of the activities of larger physical systems of which the mental mechanisms are a part. For a mechanism to be a good guide, controller and regulator of a system, the mechanism must possess a good model of the system to be controlled, guided or regulated; for a system to be a good guide, controller and regulator of itself, it must have access to a good model of itself. Thus, organisms and artificial systems alike must have access to information about such properties as their temperature and pressure tolerances, their size, strength, endurance, speed, reasoning abilities, etc. in order to survive, to move about safely, or to plan wisely. Presumably, the more complete and accurate the knowledge of itself which a system can make available to itself, the better it can perform, and the more successful it will be.

In this paper, we explore some ways in which a discrete deterministic system can make available to itself information about itself. We concentrate on the problem of a system coming to possess logical descriptions of itself (e.g. in the sense of Putnam, 1960) of sufficient completeness and accuracy that it can simulate its own possible alternative future courses of action. In order to speak with some specificity, we model the self-description process in a particular Turing machine format. We will begin by examining the strategy of machine complete self-description enunciated by Lee (1963) and given a

Synthese 38 (1978) 373–387. *All Rights Reserved.*

definite design by Thatcher (1963). We then identify the sources of some possible anomalies and pathologies of the self-description process. Explanations of empirical psychological or neurological malfunction (by locating their possible origins in specific flaws of the underlying automaton system) can thus be suggested. We conclude with some brief comment on the relevance of the self-description capacity in the development of human intelligence and inter-human relations.

2. COMPLETE SELF-DESCRIPTION

We begin by considering a Turing machine (a deterministic finite-state automaton which has access to an indefinitely expandible tape, a tape which the automaton can move to right or left, and upon which it can print symbols, and from which it can read). In explicit self-description, the Turing automaton is to make available to itself, upon its tape, information on its own structural organization (ideally in a form such that the automaton can employ this information to simulate its own alternative behaviors).

Although our discussion will not oblige us to employ any very extensive mathematical expositional apparatus, we will, for concreteness, fix upon a particular characterization of Turing machines and their descriptions. We will employ the 'computer program' characterization proposed by Wang (1957) in which Turing machines can be expressed as a fixed finite length program composed from the following repertoire of instruction types: *P*0 (*print* a *zero* in the tape square under scan), *P*1 (*print* a *one* in the tape square under scan), *R* (move *right* one square of the tape), *L* (move *left* one square of the tape), *T*0, *k*. (if scanning a *zero* on the tape go to line *k*. of the program; otherwise, go to the next instruction), *T*1, *k*. (if scanning a *one* on the tape, go to line *k*. of the program; otherwise, go to the next instruction), *H* (*halt* action). Wang showed that any Turing machine computation can be carried out by machine characterized in this 'program' fashion.

We will wish to *describe* such Turing machine programs. To do this we assign a unique fixed length code word of zeros and ones to each of the basic instruction types, and, in addition, to each of the transfer types (*T*0, *k*., *T*1, *k*), we assign its own 'count' code word which (in order to specify the location *k*.) will be repeated a *k* number of times. A description of a Turing machine program will then consist of the sequence of 0,1 code words of each of the instructions in order, while each transfer instruction will also be accompanied by a number of count code words giving the transfer locations. Fixing upon a specific such code and its details will not be necessary to the issues we will herein consider.

Turing (1936) showed that there is one of his machines which can be presented with a coded description of the structure of any of his machines and can read and interpret the description and carry out the actions of the machine described. Such a *universal* machine can read its *own* description, interpret it, and simulate what it would do in various alternative circumstances.

Can a Turing machine, its tape initially blank, provide itself with its own complete functional description, suitable for such perusal and exploratory self-simulation? Lee (1963) showed that this self-description was possible, and Thatcher (1963) provided a simple design of such a machine. This self-describing machine can be viewed as a model of certain sorts of psychological introspection, and it is the properties and possible anomalies of this model we shall herein explore in some detail.

We first explain the workings of the self-describing machine. The self-describing machine will consist of three distinguishable sub-programs: an arbitrary finitely long *general processor* (which is not vital to the self-description strategy but which is necessary if our machine is to make any use of its description: the general processor may possess whatever known effective procedures we may wish to ascribe to it); an *inference routine* (the properties of which we shall describe below); and a 'simple-minded' printing routine or *emitter.* The emitter will print out upon the initially blank tape a description, in the agreed upon 0,1 code, of all of the general processor (if one is included) and the inference routine. The structure of the emitter will be extremely simple, being composed of alternating *R* instructions (move right on the tape), and either *P*0 (print zero) or *P*1 (print one) instructions. (If the sequence 010 is to be printed upon the tape, the printer would consist of the instructions *R-P0-R-P1-R-P0*.) Thus, whatever 0,1 sequence describes the general processor and the inference routine, an emitter can easily be designed which will print out exactly that sequence.

The self-description process proceeds as follows. Our tri-partite self describing machine begins in its emitter routine which prints out upon the initially blank tape, a sequence of zeros and ones which is the coded description of the successive instructions of the general processor and the inference routine. At this point, all of the Turing machine program, save that of the emitter itself, will have been described upon the tape. It is the task of the inference routine to supply the deficiency by producing a description of the emitter routine.

The inference routine examines the 0,1 sequence presently inscribed upon the tape. If the first symbol is a 0, the inference routine moves to the end of the already printed sequence and prints out the 0,1 code words for an *R*, and for a *P*0 instruction; if the first symbol is a 1, the inference routine moves to the end of the already printed sequence and prints out the 0,1 code words for an *R*, and for a *P*1 instruction. The inference routine then returns to read the second symbol of the originally printed string, etc. Thus, the inference routine is able to supply the (heretofore absent) description of the emitter routine, thus completing the description of the machine program, and making available to the machine a description of its own program sufficient for simulation. (It should be clear that it is the especially simple structure of the emitter that makes this inference possible and that other suitably simple structures could be employed.)

3. GENERAL PROPERTIES OF THE PRINCIPAL ORGANS

We now develop more fully some behavioral implications of the three principal organs (general processor, emitter, inferrer) of the self-describing system.

A *general processor* not equipped with the special sorts of emitter and inferrer we have described can yet by itself carry out a wide range of complicated mind-like processes, some of these processes making repeated (though rudimentary and piecemeal) use of some self-description. This is so because every instance of printing of a 0 or a 1 (in our programmed Turing machine system, these are produced by *P*0 and *P*1 instructions, respectively) makes available internally stored information which can then be read by a 'test and transfer' instruction (in our system *T*0,*k*. and *T*1,*k*. instructions). Such printing and testing is ubiquitous in any very complex Turing machine program (and in any program for a general purpose digital computer) and *can* be a highly special and truncated instance of self-description (if the zeros and ones which are printed code for properties of the very system of which they are a part). However, such stored and printed values are usually not present in systems in a form sufficiently explicit and complete to permit simulation by the system of the possible future actions of the system as a whole.

The *emitter* is essentially a distinguishable 'memory' organ, of a particular form. If stimulated into action it supplies a sequence which, we have seen, could be an explicit description (in an agreed upon code) of a large part of the system itself. (In von Neumann's explication of machine self-reproduction (von Neumann, 1966), such units were called *pulsers* and their role was to store and produce on demand whole memory 'chunks', or complete instruction sequences, to implement stereotyped actions.) Emitters could also contain sets of goals or values or standards which the system as a whole contrives to achieve or sustain. (Present system status, or external input information, would be compared with the internally stored goals or values or standards, and the system, as a consequence of the detection of those comparison differences would work upon the actual state of affairs to reduce the discrepancy between the actual state and the stored 'ideals'.)

If the emitting units of a system can be altered, so that the emitted contents (and thus values, goals, etc.) can be changed, the basis for an *adaptive* system is provided. When the system alters its emitter-stored values in such a way that the system performs better (according to some criterion of success) then the system, in altering its stored values in this way, is *adapting*. The precise mechanism (presumably including natural selection as one of them) by which such memory organs have been created and loaded (in the species and in the individual) and by what means, and upon what occasions their contents are revised, is one of the great unsolved problems of biology (as well as of psychology and philosophy). (A way in which a system can inspect itself, obtain its own structural description, and construct an emitter containing the description is given in Laing, 1976.)

The *inference routine* takes descriptions in one coding system and converts them to descriptions in another coding system. Thus, the inferrer is more generally a transducer (and is related to the *decoders* and *detectors* of von Neumann's self-reproducing machines). It is composed of a series of 'sensors' (test and transfer instructions) which detect individual symbols or sequences of them, 'decode' them, and produce new code word equivalents, or specific activities which can stand for the sequences detected. One task of such transducing organs is to detect through its sensors, symbols originating external to the system and re-code them for internal system use, in the form of descriptions of things.

4. ANOMALIES AND ABNORMALITIES OF THE SELF-DESCRIPTION PROCESS

We have been explicating the ways in which our principal machine organs can be combined and used in modelling various complex system processes with special attention being given to the problem of complete structural self-description. We now wish to point out how the goal of self-description might, through various machine disabilities, completely or partially fail of attainment.

A. In the first set of ways in which the system could fail of our intended self-description purpose for it, we concentrate on disabilities of the mechanisms underlying the operation of the principal components. In living systems, these disabilities would have their counterparts in basic biochemical and neurological incapacities. In machines, the flaws would arise at the level of the implementation of the individual program instruction types (or the functioning of equivalent sub-automata) or in the sequencing or routing of program activation (or, at the 'electronic' level, of mis-wiring). We assume that the emitter does indeed produce a sequence which under some uniquely decipherable coding convention is indeed a description of all or a usably substantial part of the remaining system, but there is a failure in the system's basic capacities to access and make use of this information.

(i) The inference routine or the general processor is not equipped to 'read' or distinguish the significant symbols produced by the emitter. The emitter and the other organs do not 'understand' each other at the level of the smallest meaningful units. The description employs one or more symbols S_i and neither the processor nor the inferrer possess any instruction types of the form TS_i,k. which can detect and act upon the presence of any S_i. The description information is available, but the would-be cognizant organs are, at a sensory level, blind to it.

(ii) The emitter produces a description and the process or inferrer can detect the symbols of the description, but the result of detecting a symbol is a mis-routing of activation to an inappropriate location in the machine or program. For example, a 0 symbol is present, and is (properly) detected by a *T0,k.*; but the resulting routing of activation is incorrect. This is equivalent to an neurological 'crossed-connection'.

(iii) The emitter produces a description but the rest of the system, though it can detect properly the individual symbols of the description, does not properly interpret the resulting code words: where, for example, 001 might actually code for a *P*1, the system acts upon it as if it were another sort of instruction entirely.

(iv) The emitter produces a description but the inference routine may not be designed to infer the form that the emitter as a matter of fact takes; for example, the inference routine may be designed to assume that the emitter is always of the 'standard' R-P_i-R-P_i ... form when, in fact, it is of the functionally equivalent but redundant form R-L-R-P_i-R-L-R-P_i... In general, different emitters or different parts of the same emitter system may employ different structural conventions, and a routine which must infer emitter structure from emitter output must be designed to be aware of each emitter's structure conventions.

(v) The emitter produces a description but the system can not make use of it because the reading of the code words begins at the wrong place (a 'frameshift' error) or, for example, the description sequence is read in reverse.

B. Another set of ways in which the system could fail of our intended purpose for it of complete accurate self-description would arise owing to the ultimate finiteness of every machine physical property. That is, departing from the ideal features of a Turing machine, any actual machine will have an upper bound on the amount of tape available, on the length of persistence of inscribed symbols, on the rate or accuracy at which symbols can be read and interpreted.

(i) The emitter produces a sequence which is indeed an accurate and complete description of all of the rest of the machine, and the machine is equipped to read the individual symbols and to interpret the code words, and infer the emitter structure, but the symbols of the emitted sequence do not persist long enough to be read and interpreted completely. As a consequence the self-knowledge attainable by the system is piecemeal, intermittent, and disorganized.

(ii) As with (i), the emitted description is accurate and complete, and the system is properly equipped to read and interpret individual symbols and code words, but the 'tape' space within the system is not large enough to accommodate the whole emitted description, or large enough to hold both the emitted description and the inferred description of the emitter. In effect, the system can introspect clearly and methodically but owing to limited tape 'work space' available to the system, it must, for example, constantly transfer its successful results to an 'outside' medium. If original tape space is exceeded, and the emitter output rate exceeds the rate of transfer to an auxiliary storage medium or if emitter produced information continues to overflow and be lost while the system transfers earlier information to another medium, or if (as is likely in the analogous psychological situation) there is no adequate coding scheme for faithfully recording internal events in a more 'public' external form, then there will be a loss in either completeness or accuracy of self-description.

(iii) The description emitted is accurate and complete, and the system is equipped to employ it, but, in addition to the information provided by the description emitter, the output of other symbol producing organs is also 'printed' in the tape space, so that, in the absence of appropriate discriminatory circuitry, 'cross-talk' confuses and corrupts the self-description capacity.

(iv) The description emitted is accurate and complete, but the system suffers from a 'punctuation' incapacity; that is, in reading and interpreting the machine might lose track of the boundaries between the examined and unexamined portions of the sequence. After an emitter has produced a sequence which is a description of some or all of the remaining part of the system, the inference routine must read and interpret *all and only* the description sequence in inferring the composition of the organ or organs which produced the sequence. To accomplish this the inference routine must be guided, or arrange to guide itself, by means of various fixed and shifting markers or punctuation marks. For example, the inference routine must be able to detect where the emitted sequence begins and where it ends; it must also be able to distinguish what it has already read from the as yet unread. Such 'punctuation' incapacity might lead the inference routine 'neurotically' to read, and interpret, re-read and re-interpret, the same symbol or symbol sequence indefinitely.

(If we can ascribe to the general processor some capacity for observing and evaluating its own actions, such simple 'neurosis' might be detected and brought to a halt. Consider however the following slightly more subtle punctuation disability, where the *end marker* of the emitted sequence is lost or undetectable. In this case, after the inference routine has successfully read and interpreted the complete originally emitted string, it will fail to halt and return control to the general processor, but may begin to read and interpret the very sequence produced by itself. Since the interpretation of each symbol is several symbols long the (pseudo) description will grow longer and longer, the inference routine reading and interpreting its own output, but never re-reading the same location. In this 'neurotic' behavior, the futility of the repeated inference is much harder to detect, and again, as with the simpler example, an attempt to use the resulting supposed description would be disastrous relative to an intended self-simulation.)

5. OMISSIONS AND ORGANIZATIONAL INADEQUACIES OF SELF-DESCRIBING SYSTEMS

In the last section we pointed out many of the ways in which a machine system which was to provide itself with a description of itself could fail of that goal. We discussed fundamental disabilities of 'sensory' detection and of 'wiring', as well as limitations of system capacity and difficulties arising out of 'punctuation' accidents. There are other ways in which a system can fail to succeed in the self-description process, among them, inadequacies or omissions in the content or overall organization of the would-be self-describing system. For example, we have heretofore in this paper considered only cases where all three basic organs were present and where the emitter has been capable of producing a description of all of the rest of the system; the failings of the system were owing to local incapacity to make appropriate use of this information. In this present section we consider what might occur if the emitter could *not* always provide a description of all of the remaining parts of the system or if one or another of our three basic organs was inoperative or missing entirely.

Suppose we have a system composed of a general processor, an inference routine, and an emitter which can produce the description of the general processor (but *not* of the inference routine). The emitter will produce a description of the processor, and the inference routine can produce a description of the emitter. Can the system come to know it also contains an inference routine and obtain its description?

The actions required of our general processor may be very difficult to achieve. The processor must examine the description before it and be able to ascertain that it is a description of itself and of an emitter of the description of itself, that existence of the described emitter can account for the description of the general processor but that nothing in the general processor description accounts for the description of the emitter. It must deduce this from the existence of an heretofore undescribed part of the system. From the sequence which must be accounted for (the description of the emitter) it must deduce the structure of the inference routine which could, from the description of the general processor, produce this. In effect, the general processor must contain effective procedures which from particular behavioral consequences of machines deduce the structure which produced the sequence, a problem which in general is not recursively solvable (but which may be solvable, with suitable constraints, in the particular case at hand).

For our next example, let us suppose that the system consists of a general processor and an inference routine, and an emitter of the description of the inference routine (but no description of the processor). Then the emitter could produce a description of the inference routine, and the inference routine could produce the description of the emitter, but the system would not have a description of the general processor. The general processor would have to try to deduce the existence of itself and its description with the help only of the inference routine. It could take the emitter inference behavior as a model, but this might be very misleading, for then it would only be able to infer that within itself there were routines which produced symbols upon the tape, and it would in general be incorrect to assume that print routines originating in the processor had the same simple form as those of a standard emitter. It would seem that the processor must largely remain at a loss in its attempts to determine its own nature. (In the general case it could not even always decide correctly, given a sample of some of its behavioral consequences, whether a proffered description was indeed a description of itself.)

It is especially interesting to consider the sorts of systems which result when a general processor is combined with an emitter of the general processor description (but there is no inference capacity) and when a general processor is combined with a description inference routine (but there is no special emitter of the description of the processor and inference routine).

Considering first the case where the system consists of a general processor and an emitter of the description of the processor *and* of an inference routine (which however is *not* actually present), the emitter can produce the description of the processor and the inference routine, and the general processor can *interpret* the description of the inference routine and so deduce the description of the emitter which produced the description. Thus, this omission of the inference routine need not be a bar to the production of a complete accurate description of the system.

In the case of a processor combined with an emitter of its description (but no inferrer), the processor will have access to and can use the description of itself, but it will be 'blind' to the existence of the emitter which produced the description. In many respects the processor can get along well without a detailed description of the emitter part of the system, or even the knowleddge that it exists.[1] (But as with consciousness though, the processor, could it think about it, may feel a bit uneasy making use of a very personal phenomenon the origin of which forever eludes it.)

If a general processor is combined with a description inference routine (but the system lacks any emitter of a complete accurate description of the processor and inferrer) the ability to produce a self-description useable in system self-simulation is (as we have indicated earlier in the section) very unlikely. The general processor might in the course of its computational actions produce sequences of symbols upon the tape, but the source of these symbols could not necessarily be inferred and described by the inference routine, and there would be no guarantee that vast regions of the general processor did not exist which, since they cast no symbols on the tape-screen available to the inference routine, could never be described or their existence even disclosed

6. SOME FURTHER CONSIDERATIONS

We have presented first, an ideal state of affairs by which a system can make available to itself a complete structural description of itself. We then considered some ways in which such an ideal state of affairs might be perturbed. We now wish briefly to consider what seems to be the actual state of affairs for human, externally unaided, self-description, framing our remarks, where appropriate, in the machine terminology we have introduced.

Humans seem to possess general processors by which as largely unconscious machines much of life's activities can successfully be carried forward. There is also probably some rudimentary capacity, by means of emitter-like storage devices (devies which can be stimulated to emit whole sequences) to bring forth more or less elaborate images of the self, and of the self in the context of past, present, and proposed future actions. In attempts to introspect as to the nature and organization of our minds, we can have little assurance that what can be emitted reflects, in any very useful form, all or even much of the rest of our mental mechanism. We can be (and are) thus constantly surprised at the arousal of memories we did not know we possessed, and by the sudden activation of disused or forgotten or unsuspected parts of the general processor (although by middle-age few new faculties or capabilities announce themselves). Our inferential apparatus, in so far as emitter and processor activity is apparent to it, is reluctant to act; its timidity is undoubtedly justified since though it can deduce (rather tentatively) that the actions it is aware of must have sources, there is little certainty about the logical form these sources might take, and it assuredly does not know for certain that these sources must fall into a 'standard' structural pattern readily deducible from their behavioral properties. Thus complete and accurate self-description of the sort we have shown possible for machines is not presently possible for humans (and perhaps never was and never will be).

There is however considerable pressure upon humans (and perhaps even upon natural organisms very generally) to improve and to enlarge their capacity for useful self-description. We have long been advised to know ourselves, and in contemporary jargon we are urged to 'find out who we are' and to 'get in touch with ourselves'. Solicitous elders constantly harangue the young to develop foresight and to learn to plan their lives wisely; professional moralists urge us to empathize with others by seeing ourselves in others; engineers of society and self insist we constantly submit ourselves to our own intense scrutiny and criticism; and poets wish themselves and us to see ourselves as others see us. Clearly, some considerable available capacity to produce useable accurate models of the self as well as the means to improve that capacity is presupposed.

Quite apart from such conscious humanely intended exhortations, the indifferent natural selection pressures of evolution may have promoted an elaboration of the self-description capacity. Hamilton (1964) has pointed out that a trait for 'altruism' if it arises, would presumably be disadvantageous to the bearer; if however the altruistic trait was coupled with a capacity to recognize those individuals genetically related to the bearer, and to confine aid to them, then the trait would probably spread in the population (those who persistently helped their kind would enhance the production of more of this genetically-selfish helping kind). There is considerable evidence that 'altruism' (of this genetically-selfish sort) exists, parental care and nepotistic concern being only the most obvious and pervasive forms of it, and this trait can persist only in so far as its bearers possess a means to detect in others genuine (i.e. genetic) similarity to self.

Since genes themselves are not readily open to direct inspection, organisms possessing only superficial likeness can take advantage of the altruistic traits of others. It is thus evolutionarily advantageous that individuals of the altruist persuasion constantly improve their discriminatory powers and in addition impose, if they can, increasingly severe sanctions upon would-be parasitic members of the larger population deceitfully posing as related fellow altruists.

Alexander (1975) has pointed out that this situation is undoubtedly further complicated by two additional factors. First, it is likely that most humans simultaneously possess both altruistic and parasitic traits, and second, parasitism will likely be most successful if its possessors are largely blind to this 'immoral' aspect of themselves.

In our presentations of various machine self-description strategies, we have usually assumed that any descriptions of the rest of the system produced by the emitter are indeed correct descriptions and that the 'proper' function of the inference mechanism is to implement a *complete* as well as correct description. For an organism (or a machine) existing in and reacting with a complex environment of other organisms or machines and being 'naturally' selected upon according to continued reproductive success, a complete and correct description of the self may *not* necessarily be the most advantageous. Of course, if the machine or organism is designed to operate under the assumption that the emitter apparatus makes available to it a complete and correct description of itself, and the description is grossly and constantly incorrect, then attempts to produce prudent appropriate behavior on the basis of the distorted notion of self would generally be disastrous to the bearer. Alexander's analysis suggests however that the 'normal' state of affairs is one in which there is a compromise between accuracy and completeness on the one hand and distortion and omission of self-description on the other.

The more accurate and complete the description then ideally more wisely we can behave; but our present and past experience of survival from moment to moment and generation to generation has perhaps produced a genetic policy of partial ignorance and distortion. An incorrect or partial self-description (which can exist within us in multiple differing forms) does not usually diverge from reality to the extent that gross performance and viability are constantly and directly threatened. Indeed we can assume that the blind spots and distortions of our notion of ourselves must have had survival value and may still have. The produced self-description is usually faulty, but often despite this, quite comforting. Relying upon it may sometimes produce bafflement, but rarely disaster. Calculation of future courses of action, based on our beliefs about our capabilities and nature often prove inaccurate and disappointing; assumptions about one's strength, size, comeliness, and mental acumen fall wide of the mark, when tested in the world. Although the psychic consequences of such failures *can* lead to desperate attempts to reconcile this strange state of affairs, and even to despair, we usually find ourselves too busy, just getting along, to dwell long upon it.

The University of Michigan, Ann Arbor

NOTES

*This research was supported in part by the National Science Foundation.

[1] 'Physiologically' the absence of an inference routine would be disastrous. Let us suppose the machine *reproduced* itself 'consciously' on the basis of the (partial) description provided by the emitter. Then the processor would be reproduced, but not the 'hidden' emitter of the description of the processor. The offspring processor would thus not make available to itself a description of itself for use in simulation and guidance, and in turn could not itself reproduce at all.

BIBLIOGRAPHY

Alexander, R. D., 1975, 'The Search for a General Theory of Behavior', *Behavioral Science* 20 (2), 77–100 (esp. p. 97).

Hamilton, W. D., 1964, 'The Genetical Theory of Social Behavior (I and II), *J. Theor. Biol.* 7, 1–16, 17–32.

Laing, R. A., 1976, 'Automaton Introspection', *J. Comp. and Systems Sci.* 31 (2), 172–183.

Lee, C. Y., 1963, 'A Turing Machine Which Prints its Own Code Script', *Mathematical Theory of Automata,* Polytechnic Press, Brooklyn, New York, pp. 155–164.

Putnam, H., 1960, 'Minds and Machines', *Dimensions of Mind*, ed. by Sidney Hook, New York University Press, New York, pp. 148–179.

Thatcher, James, 1963, 'The Construction of a Self-Describing Turing Machine', *Mathematical Theory of Automata*, Polytechnic Press, Brooklyn, New York, pp. 165–171.

Turing, A. M., 1936 'On Computable Numbers, with an Application to the Entscheidungsproblem', *Proc. of the London Mathematical Society*, Series 2, **42**, 230–265.

von Neumann, J., 1966, *Theory of Self-Reproducing Automata*, edited and completed by A. W. Burks, University of Illinois Press.

Wang, H., 1957, 'A Variant to Turing's Theory of Computing Machines', *J. Assoc. Comput. Mach.* **4**, 63–93.