S.G. Shanker

THE NATURE OF PHILOSOPHY

'When you are philosophizing you have to descend into primeval chaos and feel at home there.' (*Culture and Value*, 65)

In his Introduction to *The Legacy of Wittgenstein* Anthony Kenny sadly intones that 'the experience of the last decade has been chastening for those who think that wittgenstein's work is important and should be more widely appreciated' for wittgenstein's influence, he reports, 'seems to be declining rather than increasing.'[1] Of course it is difficult to know precisely how one judges such matters. As early as 1930 Moritz Schlick was arguing that Wittgenstein's work heralded 'the decisive turning-point in the history of philosophy'. But does the fact that comparatively few have echoed this accolade over the past fifty years actually serve to falsify Schlick's optimism ? The trouble here is that a revolution in philosophical thought is far from being an obvious or spontaneous event. Wittgenstein himself cautioned : 'Philosophy hasn't made any progress—If somebody scratches the spot where he has an itch, do we have to see some progress ? Isn't it genuine scratching otherwise, or genuine itching ? And can't this reaction to an irritation continue in the same way for a long time before a cure for the itching is discovered ?' (CV 86,7) But philosophers long for immediate results, not a gradual remission; which perhaps accounts for the fact that philosophers are so drawn—as the previous introductory essays of this collection more than amply demonstrate—to political and activist metaphors. Perhaps this is merely a consequence of the cloistered thinker's secret admiration for the man of action, or perhaps it reflects the frustration which he feels at the tortuous pace of philosophical progress. But

whatever the cause, it is clear that it can be gravely misleading to compare a philosophical to a political revolution. For whereas in politics the success of a revolution is marked by, as Bierce described it,'an abrupt change in the form of misgovernment', the influence of a revolutionary figure in a conceptual area as abstract as philosophy is something which only really manifests itself over generations, and in ways not immediately obvious to the protagonists of these epic struggles.

The problem as Kenny sees it, however, is that while 'we may say that Wittgensteinian scholarship has blossomed', it is unfortunately the case that 'Wittgensteinian philosophy, as opposed to wittgensteinian scholarship, has not made progress, and some of the gains we owe to Wittgenstein seem in danger as being lost.'[2] But the obvious difficulty with this argument is that until the latter activity has been at least partially successful in disclosing the purport of Wittgenstein's remarks, there can be little hope of the former prospering. Indeed, it is not at all clear that the two can be so sharply separated. For as we can see particularly clearly from the articles gathered in this final volume, Wittgensteinian scholarship is far from being a simple exegetical matter. The very manner in which Wittgenstein developed and presented his arguments militates against any orthodox critical approach to the writings of a master. For the basic fact of Wittgensteinian scholarship is simply that we are not presented with a body of philosophical theses, waiting to be catalogued and perhaps reordered or expounded. On the contrary, what we are confronted with are first and foremost philosophical *problems*, which Wittgenstein clearly expected us to resolve for ourselves, albeit on the basis of the guidelines laid down in his brief discussions.

To be sure, Wittgenstein chose these problems—and the manner in which he sought to re-orient our approach to the confusions on which they rest—in such a way that they would be seen to interlock with one another, and thence both arise from and contribute to a

new understanding of the nature of philosophy. Thus one of the primary tasks of Wittgensteinian scholarship has been to uncover the careful design underlying Wittgenstein's presentation of his major themes. Nevertheless, at an even deeper level we cannot hope to come to terms with the full dimensions of Wittgenstein's thought without immersing ourselves in the problems themselves. For Wittgensteinian scholarship can only genuinely proceed via the rigorous application of wittgenstein's arguments to specific philosophical problems; on the basis, that is, of the vigorous growth of Wittgensteinian philosophy. As Baker and Hacker have shown in the first volume of their commentary of *Philosophical Investigations*, it is only with the synthesis of textual exegesis and the simultaneous resolution of major philosophical issues on the basis of the major insights afforded by this exegesis, that Wittgensteinian scholarship will be able to come to terms with the profound implications of Wittgenstein's thought. By no means, however, should we underestimate the complexity or labour involved in such a task. For we must be clear about one thing : Wittgenstein was not involved in the piecemeal construction of abstrue philosophical theories, Rather, he was engaged in the clarification of problems which have troubled philosophers for thousands of years. The very obduracy of these issues must itself constantly remind us of the fact that the confusions which have inspired these questions must be exceptionally deep-rooted, and for that reason, exceptionally difficult to expose and extirpate. Indeed, had they been facile or superficial matters, they would not have become *philosophical* problemes. But by the same token, we connot then expect the essence and the consequences of Wittgenstien's resolution of these issues to be immediately forthcoming.

One can certainly sympathize, however with kenny's feelings of frustration and impatience. Yet Kenny worries that there is a larger issue at stake here. The problem is not simply that Wittgenstein's insights have not been fully absorbed : it is that the very spirit of

modern philosophy is proceeding in a direction which will make it increasingly difficult for future philosophers to grasp the nature and the point of Wittgenstein's thought. Thus kenny warns that Wittgenstein's 'contribution has been neglected because more and more philosophers have attempted to model their studies on the pattern of a rigorously scientific discipline'. 'This kind of scientism in philosophy', he quite rightly points out, 'was something which Wittgenstein abominated, and in such a climate the seeds he planted have a poor chance of flourishing growth.'[3] This is without question a crucial theme, and one which Wittgenstein himself took great pains to emphasize. Thus in *Philosophical Remarks* he forthrightly declared his intention to undermine the scientistic conception of philosophy, insisting that 'in philosophy it's always a matter of the application of a series of utterly simple basic principles that any child knows, and the—enormous—difficulty is only one of applying these in the confusion our language creates. It's never a question of the latest result of experiments with exotic fish or the most recent developments in mathematics. But the difficulty in applying the simple basic principles shakes our confidence in the principles themselves.' (PR 133)

Wittgenstein was not only aware of, but extremely eager to, draw attention to the fundamental confrontation between this and the Russellian conception of philosophy. Russell, of course, had argued that 'Philosophy, like all other studies, aims primarily at knowledge'; that philosophy should 'seek to base itself upon science', to 'study the methods of science, and seek to apply these methods, with the necessary adaptations, to its own peculiar province.'[4] Moreover, according to Russell, 'to a great extent, the uncertainty of philosophy is more apparent than real: those questions which are already capable of definite answers are placed in the sciences, while those only to which, at present, no definite answer can be given, remain to form the residue which is called philosophy.'[5] The upshot of t he argument formulated in *Philosophical Remarks*, however, is that

the point of a philosophical investigation is not—*pace* Russell—to seek to acquire a greater intellectual respectability for the 'science of pure reason' by locating its accomplishments within the mainstream of contemporary scientific progress ; nor is it, as Russell hoped, to place philosophy on an inexorable path of eventual self-extinction, so that there would ultimately be no such thing as 'philosophical' questions left to excite man's latent metaphysical instincts, thus leaving us free to pursue genuine scientific matters. Rather, the goal of philosophy is purely and solely to clear up the confusions created by the misuses of logical grammar: 'What we find out in philosophy is trivial ; it does not teach us new facts, only science does that. But the proper synopsis of these trivialities is enormously difficult, and has immense importance. Philosophy is in fact the synopsis of trivialities.' (LWL 26)

Needless to say, Russell did not take at all kindly to the challenge which Wittgenstein had thrown down. If Wittgenstein was right, Russell complained in *My Philosophical Development*, then 'philosophy is, at best, a slight help to lexicographers, and at worst, an idle tea-table amusement.'[6] Wittgenstein's stress on the 'enormous difficulty' involved in arriving at the banalities which are the hallmark of philosophy was little consolation to Russell. What most concerned him was Wittgenstein's renunciation of the principle that 'the point of philosophy is to start with something so simple as not to seem worth stating it, and to end with something so paradoxical that no one will believe it.'[7] Where the 'scientific philosopher' stands in awe before the hidden mysteries of thought—which he feels we are only now just beginning to penetrate—Wittgenstein had perversely insisted that 'there is no mystery here'- Thus, although it was perfectly clear that wittgenstein did not intend this theme to be seen as in any way detracting from the difficulty or the importance of philosophy, all that really mattered was Wittgenstein's belief that, as he described it in *Zettel*, 'The philosopher is not a citizen of any

community of ideas. That is what makes him into a philosopher.' (Z 455)

For all the attention which Wittgenstein devoted to elucidating the significance of this point, it remains a theme which it is all too easy to misconstrue and thus dismiss. For the 'scientific philosopher' may well feel that this is nothing more than an unwarranted *ad hominem*. After all, if he can actually be said to be engaged in the pursuit of truth, and if he can be seen to be making progress—however limited—in the construction of theories which are at least *in posse* capable of yielding new insights into the nature of thought, then by what right is he to be denied his claim to an honoured position amongst the community of scientific burgers? This theme comes out particularly clearly in the current exchanges between 'theorists of meaning' and Wittgensteinian philosophers. The theorist of meaning wants to argue that philosophy carries the rational behaviour which mankind as a species is continually developing to an even greater extent. The use of language, he maintains, is to be seen as the principle manifestation of our rationality : the vehicle of our thought. But it is only in so far as we can acquire *explicit* knowledge of the rules governing the construction of significant expressions that we will have an adequate conception of our mastery of a language, and hence not succumb to the slavery imposed by linguistic confusion or ignorance. His apologia thus rests on the point that there are many terms within natural language which express this mastery and which are used by language-speakers to elucidate the nature of language. Like any other scientist, therefore, the philosopher *qua* theorist of meaning is pursuing and sharpening the original elementary terms in his struggle to develop and refine these comparatively primitive tools. These 'second-order' expressions will become incorporated in the philosopher's meta-theory about language, but if we are to appreciate the proper status of philosophy it is essential that we recognize the

origin of these expressions within natural language itself. Philosophy is thus seen as the rational continuation of the man on the Clapham Omnibus' attempt to understand the nature of language, But whereas the ordinary language-user has only an intuitive conception of the mechanics of the language which he speaks, philosophy attempts to transform this inchoate knowledge into a concise and perspicuous theory about language.

The theorist of meaning bases his enterprise, therefore, on the premise that the point of philosophy arises from our innate and rational drive to reflect upon ourselves as language-users. But this in itself is certainly not a theme which Wittgenstein wished to repudiate. The source of the conflict lies in the theorist of meaning's further assumption that the philosopher can be legitimately engaged in the empirical construction of linguistic theories ; for philosophical problems, Wittgenstein insisted, are conceptual, not empirical, and the theorist of meaning's enterprise rests squarely on a violation of precisely this grammatical demarcation. But then, perhaps this too will strike the theorist of meaning as an *ad hominem* objection, given the prior assumption that the guiding principle behind philosophical progress just is to displace what he regards as idle *a priori* speculation with verifiable hypotheses about the structure and use of language. However, there is a further problem which the 'theorist of meaning' simply passes over in his haste to acquire the badge of scientific respectability ; for, unlike the sciences, philosophy might be said to be an ongoing process in which the earliest problems and writings on the subject remain timelessly relevant and illuminating. In 1931 wittgenstein wrote :

> You always hear people say that philosophy makes no progress and that the same philosophical problems which were already preoccupying the Greeks are still troubling us today. But people who say that do not understand the reason why it has to be so.

The reason is that our language has remained the same and always introduces us to the same questions.

I read : "... philosophers are no nearer to the meaning of 'Reality' than Plato got,...". What a strange situation. How extraordinary that Plato could have even as far as he did ! Or that we could not get any further ! Was it because Plato was so ***extremely*** clever ? (CV 15)

From the point of view of a science there is clearly an anomaly here—and certainly a complication—which needs to be acknowledged and explained.

Conversely, we must read this theme against Wittgenstein's notorious(and frequently repeated) insistence that 'The real discovery is the one that makes me capable of stopping doing philosophy when I want to.—The one that gives philosophy peace, so that it is no longer tormented by questions which bring *itself* in question.' (PI 133) But here too, far from wishing to in any way denigrate philosophy, Wittgenstein's main concern was to attack a basic premise which underlies 'scientific philosophy' : the idea that philosophy is hierarchical—in the sense that there are certain fundamental areas of philosophical inquiry which provide the material for solving problems in the secondary branches of the subject—and hence progressive, in the sense that the philosopher slowly moves his way up the ladder of philosophical enlightenment. In Frege : *Philosophy of Language* Michael Dummett suggests that in the *Tractatus* Wittgenstein established 'philosophical logic as the foundation of philosophy, and relegated epistemology to a peripheral position.'[8] Whether this applies to the *Tractatus* is, peshaps, a moot issue (for Wittgenstein did not so much relegate epistemology to the fringes as consign it to oblivion). What is clear, however, is that this argument definitely cannot be applied to Wittgenstein's later conception of philosophy.

The point of Dummett's centre/periphery picture is that we must resolve certain problems in philosophical logic in order to furnish the logical tools with which to answer the qnestions plaguing the 'peripheral' philosophical disciplines (such as epistemology, ethics, political philosophy etc). The overriding danger in this picture is that it encourages one to suppose that there are *answers* to philosophical *questions* : i.e. there is a body of significant logical truths (which successive generations of philosophers inherit and augment). This, of course, is a point which Wittgenstein vociferously and continually rejected, always insisting that philosophy is an acitivity as opposed to a supra-scientific body of knowledge. Nor was Wittgenstein's point that we must engage in certain forms of activity – i.e. master certain techniques – before we can address secondary philosophical topics ; for this would be but a covert attempt to reintroduce theses into the realm of philosophy. Certainly there is no such theme present in the articles reprinted in this volume. On the contrary, what we find here are a series of attempts to apply the techniques and themes which Wittgenstein developed to resolve the philosophical problems which occur in the wide range of subjects represented by these papers. Moreover, there is not the slightest indication that the authors view their subjects as in any way 'peripheral'; and there is never any suggestion that they view the philosophical solutions which they have propounded as the consequence of applying a body of basic logical truths to their fields.

Thus, the obvious trouble with the 'periphery' picture is simply that it encourages us to look at philosophy as a pseudo-scientific discipline ; or rather, a group of closely related quasi-scientific subjects which together constitute the 'Queen of the Sciences'. Apart from Wittgenstein's repudiation of the scientific conception of philosophy, it is also important that we recognize that Wittgenstein was trying to break down the artificial barriers departmentalizing philosophy into separate autonomous subjects. His license for

talking about mathematics, he told his students in 1939, resided simply in the fact that 'I can as a philosopher talk about mathematics because I will only deal with puzzles which arise from the words of our ordinary everyday language, such as "proof", "number" "series", "order", etc.' (LFM 14) The basic point here is that there is only one discipline—activity—that goes by the name of philosophy: the removal of philosophical problems, which are caused by the confusions engendered by violations of logical grammar. If there is a meaningful distinction to be drawn within the body of philosophical activity, therefore, it is only in so far as the philosopher is concerned, in the various orthodox subjects of philosophy. with some identifiable category of concepts (e.g. epistemology is concerned with the removal of philosophical problems which occur from the confused application of epistemic concepts ; the philosophy of mathematics removes the confusions which are created when mathematical concepts are used illicitly, etc.) But the various topics in philosophy, far from being different species of *a priori* sciences, are merely different areas of conceptual clarification, where in each case we employ very much the same techniques to remove the philosophical problem involved : viz, we clarify the logical grammar of the terms concerned by achieving an *Ubersicht* of the segment of grammar which has generated our perplexity.

As it stands, however, relatively few 'scientific philosophers' have been persuaded to abandon their putative theories on the basis of having asked themselves : 'Why is philosophy so complicated ? It ought, after all, to be *completely* simple.--Philosophy unties the knots in our thinking, which we have tangled up in an absurd way ; but to do that; it must make movements which are just as complicated as the knots. Although the *result* of philosophy is simple, its methods for arriving there cannot be so. The complexity of philosophy is not in its matter,

but in our tangled understanding' (PR 2) yet the explanation for their intransigence is not simply to be accounted for by the refusal of 'scientific philosophers' to consider the basis and thence the implications of Wittgenstein's thought. To be sure, Wittgenstein presented a conception of the nature of philosophy which offends very deeply against the scientistic mood which currently dominates analytic philosophy. But the aspiring Wittgensteinian philosopher must be extremely careful that he does not place too much of the responsibility for his failure to win over a significant number of converts to his conception of the nature of philosophy on the 'spirit of the times'. For this is not just a matter of the state of receptivity to wittgenstein's anti-scientistic conception of philosophy in an age which is profoundly governed by scientific attitudes and aspirations. As we saw above, Wittgenstein himself emphasized that the source of philosophical problems lies in the very essence of language : a point which transcends any particular cultural epoch and applies to the human condition *simpliciter*. Hence, the source af our difficulties cannot simply be dismissed on the grounds that the seeds which Wittgenstein planted have little chance of flourishing in a climate such as the persent. For if Wittgenstein was right, philosopher must not allow this to, in any way, deter him from his task. Rather, he must accept that, if he is to have any chance of reaping a bountiful harvest, he cannot simply scatter his seeds to the wind and hope that some will eventually germinate. Rather, he must cultivate his fields assiduously, and diligently tend the growing plants. For it is not the case that 'scientific philosophers' have been wholly unwilling to tackle Wittgenstein's arguments, nor for that matter, that they have been uninterested in what Wittgenstein had to say. It is really that they remain unconvinced by Wittgenstein's objections in the face of what they regard as the overwhelming counter evidence provided by modern science. Indeed, a great many philosophers and scientists

alike interpret the Wittgensteinian attack on scientism in philosophy as a manifestation of an intellectual primitivism which is fundamentally incapable of coming to terms with the great advances that have been made by modern technology. To respond to their objections that a computer will prove no more advantageous than an abacus in the solution of philosophical problems will, one suspects, strike then as the folly born from the failure to realize the vast difference between the early digital computing devices and the enormous reasoning potential of the 'fifth-generation computer' (or, as we must all learn to call it, the KIPS, or 'knowledge information processing system').

It is without question crucial that the Wittgensteinian philosopher has fully grasped the nature of this great divide separating the scientistic from Wittgenstein's conception of philosophy before he attempts to implement any of the particular themes raised in Wittgenstein's writings Yet he might well worry that, stated in these sweeping terms, the debate is unlikely to win many converts on either side of the issue. It is understandable—and no doubt prudent—that the Wittgensteinian philosopher should find himself despairing that Wittgenstein's conception of philosophy was written only 'for such men as are in sympathy with its spirit.' ! for the worry must remain that *Weltanschauungen* are ultimately far more powerful than understanding. As stated in these stark terms, however, the argument fails to do justice to the full depth of Wittgenstein's achievement. The only way in which we can truly appreciate the thrust and the significance of Wittegenstein's conception of the nature of philosophy is by applying his arguments to the actual resolution of the major philosophical issues which predominate in philosophical circles. For there is all too great a danger that, if the argument is left at this level of generality, it will str ke the 'scientific' philosopher as nothing more than an empty rhetorical gesture. Indeed, it might even strike him as irrelevant. And perhaps it is just this pervasive sentiment which accounts for the continuing growth of scientism in philosophy. For only one thing could be worse for the modern

spirit than being seen to have been wrong in an argument, and that is to be deemed irrelevant. Yet we shall gain the attention—and perhaps thereby earn the respect—of our scientific colleagues by seeking, not to mimic their own activities, but rather, to clarify why philosophical problems are important—why the consequences of conceptual confusion may not just be pernicious, but might even be dangerous—and why the solution to philosophical problems is categorially different from that which occurs in the sciences. Yet to accomplish this, we must leave the lofty plain of philosophical generalization and dirty our hands in the actual details of the reat issues which preoccupy our times.

It would be doing a great disservice to Wittgenstein, however, to suppose that he himself was guilty of such a shortcoming. There is a tendency today to quote Wittgenstein's remarks on the nature of philosophy *vis a vis* science and leave the matter at that, Yet the truth of the matter is that the basis for Wittgenstein's comments on this issue is to be found in the bowels of Wittgenstein's arguments on any number of specific technical issues. Quite clearly, therefore, the cause of Wittgensteinian scholarship rests squarely on our preparedness to follow Wittgenstein into these murky areas where chaos reigns. Wittgenstein's curt remarks on the incoherence of the mechanist thesis provide an excellent example of the types of problems which the Wittgensteinian philosopher who is prepared to accept this challenge must fully expect to encounter. Wittgenstein's sparse arguments have noticeably failed to win many converts among those actively engaged in extending the frontiers of Artifical Intelligence. Indeed, among practising exponents of this 'knowledge-revolution' Wittgenstein's comments must appear to be positively bizarre. Yet it is no good our simply retorting that the mechanist thesis is guilty of a version of the 'homuncu us fallacy'—e.g. the mechanist thesis rests on the misapplication[1] of normative concepts to mechanical operations—and leave he matter at that. If Wittgensteinian philosophy is to be at all effective, it must - as was argued above—enter into the actual details of such problems as they appear to the main disputtants

engaged in these areas, and try to reveal to them how at each of the key stages of their argument they have committed some variation of this basic grammatical confusion. Otherwise, there is a distinct danger that the Wittgensteinian philosopher will strike his opponents as a dilletante engaged in the superficial critique of their fields without truly understanding the complex nature of the arguments involved. Even worse, he is apt to strike the 'scientific philosopher' as a Luddite, intent not so much on clarifying as on thoroughly repudiating the problems thrown up by modern science.

Nothing could have been further from Wittgenstein's own intentions. We must be extremely careful, therefore, that we do not confuse Wittggenstein's attack on scientism in philosophy for an attack on science *per se*. On the contrary, one of our most pressing tasks *qua* philosophers, according to Wittgenstein, is to clarify the significance of the discoveries which modern science is constantly providing. Bouwsma recalls in his unpublished 'Notes' how Wittgenstein had argued that :

> This is the age of popular science, and so this cannot be the age of philosophy. He was not objecting to this. In fact he recommended Faraday's *The Burning of the Candle* as an illusration of fine popular science. He objected to the sentationalism, and what he called the cheating. Eddington and Jeans cheat. A fine work in this order would have to be very careful, analogies would be well chosen and nicely worked out. In fact the consummation of philosophy might very well be just such fine popular science, work which does not cheat and where the confusions have been cleared up. He was especially resentful of philosophy on the radio – more sensationalism.

The problem is, however, that philosophy must clear up the muddles thrown up by science without itself encroaching on the domain of

the scientist. Thus, if we argue that it is unintelligible to debate whether machines can think, the emphasis must be placed firmly on the *unintelligibility* of the mechanist thesis. In other words, the confusions which undermine the mechanist thesis are entirely conceptual, not empirical. In *The Blue Book* Wittgenstein argued that :

> the problem here arises which could be expressed by the question : "Is it possible for a machine to think ?" (whether the action of this machine can be described and predicted by the laws of physics or, possibly, only by laws of a different kind applying to the behaviour of organism). And the trouble which is expressed in this question is not really that we don't yet know a machine which could do the job. The question is not analogous to that which someone might have asked a hundred years ago : "Can a machine liquefy a gas ?" The trouble is rather that the sentence, "A machine thinks (perceives, wishes)" : seems somehow nonsensical. It is as though we had asked "Has the number 3 a colour ?" (BIB 47).

Our task in this issue is to unfold the implications of this passage, and use these insights as our guideline as we descend through the layers of conceptual confusion which underpin the mechanist thesis, spelling out at each stage precisely why the mechanist thesis has transcended the bounds of sense. We are not concerned, therefore, with the question whether Artificial Intellingence-researchers will one day be able to construct the sophisticated prosthetic apparatuses which they envisage. It is whether it is coherent to describe these devices as 'expert'—i.e. 'knowledge-based'—systems. Thus, our concern is not with the technical claims that have been made by knowledge engineers : it is whether it is meaningful to describe the type of activity in which they are engaged as 'knowledge engineering'.

The basic issue we are concerned with here turns on the normativity of rule-following behaviour. To say to someone that his notions are normative—that he is following a rule—entails that he is consciously guiding himself by the application of that rule. In other words, we only say of someone that he is following a rule if he can satisfy a host of normative criteria ; e.g. if he can explain or justify his actions in reference to the rule, if he can teach, evaluate, or criticize others in the application of the rule, etc. Thus, mere regularity does not suffice to constitute rule-following behaviour ; it is only against the background of this cluster of normative criteria that we attribute rule-following behaviour.

This argument is particularly important when we come to consider the logical grammar of the key concepts which lie at the heart of the mechanist thesis, such as 'calculation' or 'inference'. In *Remarks on the Foundations of Mathematics* Wittgenstein asked :

Does a calculating machine *calculate* ?

Imagine that a calculating machine had come into existence by accident ; now someone accidentally presses its knobs (or an animal walks over it) and it calculates the product 25×20.

I want to say : it is essential to mathematics that its signs are also employed in *mufti*.

It is the use outside mathematics. and so the *meaning* of the signs, that makes the sign-game into mathematics.

Just as it is not logical inference either, for me to make a change from one formation to another (say from one arrangement of chairs to another) if these arragements have not a linguistic funcion apart from this tranformation. (REM V § 2).

This passage will no doubt strike AI-researchers as hopelessly perverse ; for surely, they will argue. it just is the case that there are

machine-calculable algorithms, and information theory has unquestionably become one of the most important of the new fields in knowledge-engineering. How could Wittgenstein possibly deny the existence of calculating machines when all around us there is so much evidence to the contrary ? Was Wittgenstein simply unaware of the computer revolution that lay just on the horizon : a revolution that would take us far beyond the primitive state of calculating devices which existed at the time that Wittgenstein was writing, into the realms of the "executive' robot ? Or did Wittgenstein perhaps harbour some obscure reactionary instinct in the area of mathematical innovation ?

The point Wittgenstein was raising, was, however, solely concerned with the intelligibility of speaking of a *mechanical calculation*: with the question of whether it makes sense to describe the operations of such sophisticated machines as 'calculations', let alone as thinking, understanding, knowing, inferring etc. The tension he was concerned with lies solely in the fact that *calculation* as we understand the term, is a normative concept. A monkey can be trained to perform various tasks with pebbles, but can it be trained to calculate with the pebbles ? The problem here is not simply that the concept of calculation involves a considerable amount of background knowledge of the rules governing the basic operations of arithmetic ; more importantly, it is that the concept of calculation, *qua* normative concept, demands the ability to follow a rule. The main theme of our investigation into this issue, therefore, must be to consider how we should properly describe the operations which a machine performs. In short, the questions we are concerned with are purely a matter of the logical grammar of *calculation*, *inference* etc. *vis-a-vis* the logical grammar of mechanical operations ; we are not in the least concerned with the empirical question of the extent to which machines can be developed which will simulate complex aspects of human behaviour.

The basic grammatical confusion we are concerned with is one which is profoundly entrenched in the history of the subject. No better example of this could be found than Babbage's ill-fated vision of coustructing an 'Analytical Engine'. Given the central role that the Jacquard loom played in Babbage's invention, it is proof of the strength of this confusion that Babbage believed that the machine which he hoped to construct using Jacquard-like card-controlled programme could be meaningfully described as an *analytical* engine. Eady Ada Lovelace remarked that Babbage's 'Analytical Engine *weaves algebraic patterns* just as the Jacquard-loom weaves flowers and leaves'. It is a comment worthy of serious consideration. Unfortunately, what has actually happened is the creation of a tangled skein of grammatical confusions masquerading as conceptual analyses. Thus, far from being merely concerned with the mechanical manipulation of symbols—as opposed to yarn—Babbage's idea, according to one modern philosopher, successfuly 'mechanized the more mentalistic idea for a plan or design for a calculation.[9]

The confusion here is clearly one which is deeply embedded in the modern psyche Confronted with an obvious paradox, the mechanist refuses to examine the grammatical basis for the expressions he employs which have created his puzzlement, and instead searches for an escape in the realms of technical innovation. Douglas Hofstadter tells us at the beginning of *Godel, Escher, Bach* that, when first approaching the subject.

> One runs up against a seeming paradox. Computers by their very nature are the most inflexible, desireless, rule following beasts. Fast though they may be, they are nonetheless the epitome of unconsciousness. How, then, can intelligent behaviour be programmed ? Isn't this the most blatant of contra-

dictions in terms ? One of the major theses of this book is that it is not a contradiction at all.[10]

But before Hofstadter even embarks on his long exposition of some of the more recondite 'heuristic' and 'executive' computer programmes that can be devised, one wants to confront him with the very normative premises which are contained in this outline, and which vitiate his subsequent efforts to vindicate his thesis with a lesson in the subtlety of computer programming. For the solution to a contradiction lies in grammatical, not empirical clarification.

From the beginning Hofstadter assumes that it make sense to describe computers as 'rule-following beasts' ; but this is the very issue we want to examine, and not whether complicated heuristic programmes can be constructed which lessen the 'inflexibility' of computers. Indeed, once this former issue has been properly clarified, we will see why it is so misleading to employ all of the subsequent terms which he presents. What does it mean to say that a computer is 'inflexible', or that it is 'desireless' ; these are hardly contingent features of computer 'behaviour'. Indeed, what does it mean to say that a computer is 'the epitome of unconsciousness' ? Surely a computer does not display the type of 'behaviour' that psychologists have long described as 'unconscious' it is that the concept of consciousness simply cannot be applied to a machine.

At the end of his book Hofstadter produces a long fictionalized dialogue with Babbage. It would be a valuable exercise if one were to go through this end-piece carefully, with an eye to spotting all of the instances when Hofstadter unreflectingly introduces normative and international concepts to characterize what he describes as 'computer behaviour'. In the interests of space, we shall limit ourselves to one such example, since it is one of the standard arguments in the AI battery. What is offered as a completely

innocuous example which is supposed to secure the basic minimum of support for the proof of machine intelligence betrays a profound intensitivity to the grammatical confusions which are concealed at even the most fundamental levels of the argument :

Crab : What about imbuing a machine with human mental faculties, such as intelligence ?

Babbage : I have given such matters some thought from time to time, but I have never had the proper hardware to take up the challenge. This is indeed a felicitous suggestion, Mr. Crab and I would enjoy nothing more than working with your excellent Theme. Tell me—did you have any specific kind of intelligence in mind ?

Crab : An idle thought which had crossed my mind was to instruct it in such a manner as to play a reasonable game of chess.

Babbage : What an original suggestion ! And chess happens to be my favourite pastime. I can tell that you have a broad acquaintance with computing machinery, and are no more amateur...I can make two of the subprograms play against each other, in the manner of two persons who play chess together in select chess club. Meanwhile, the third subprogram will play Mr. Crab. That way, all three internal chess players will be occupied.

Crab : That's an amusing suggestion—an internal mental game, while it combats an external opponent. Very good !

Tortoise . What else could this be called, but a three-part chess-fugue ?

Crab : Oh, how recherche ![11]

Recherche indeed. To be sure, sophisticated computer chess programmes have been devised, and it is entertaining to play chess in this fashion. But now, although I might be playing chess against the computer, is the computer playing chess against me ? Could, as Hofstadter suggests, the computer plays chess against itself, or against another computer ? How should we clarify such matters : by looking at the nature of the programmes, or by examining the logical grammar of the concept *playing a game* ?

In 'Machines with Minds' Anthony Quinton argues :

> We talk of compters playing chess but they do not do so in the standard human way, they do not move pieces about the board. They do so more in the manner of one who plays chess by post. They say what move is to be made, but do not themselves make it...They can play chess in the fashion I have just described. They can play backgammon well enough to beat the human world champion as happened at Monte Carlo in 1979. As Dr. Hans Berliner very creditably admits 'the dice rolled well for it', but it remains a memorable feat. They can solve problems in calculus and in algerba of a not too sophisticated sort. They can find proofs of theorems in mathematical logic, sometimes neater and more elegant than those appearing in the logical classic *Principia Mathematica*, which were, so to speak, hand made by A.N. Whitehead and Bertrand Russell.[12]

The basic confusion we are concerned with here—the the ascription of normative and international behaviour to the mechanical manipulation of symbols—runs throughout this passage. It is hardly the case that the objection we want to put against the suggestion th t computers play chess is that they cannot move the pieces themselves; nor, at a deeper level that the actual manner in which computer programmes function may be completely unlike the methods

whereby humans decide which moves to make. What really concerns us are such suggestions as that the computer 'beat' the backgammon champion ; was the compter trying to win ? And when Quinton says that computers can 'find proofs of theorems in mathematical logic', does that mean that they were looking for such proofs ? Clearly something has gone away in such arguments. What we are faced with is not a case of conceptual relignment, but rather, of grammatical confusion : for it just is unintelligible to apply a normative/intentional concept to a mechanical manipulation of a symbol.

In order to understand the questions whether machines can 'think', 'play games', 'calculate' etc., we must clarify the grammar of the concepts *think*, *play games*, *calculate* *vis a vis* the logical nature of *machines* with an eye io clarifying the distinction between *normativity* and *mechanical correlation*. Perhaps the greatest difficulty that we face in discussing this question is keeping firmly in mind that the problem we are concerned with is logical, not empirical. In many ways the problem is similar to asking whether animals can think. No amount of neurophysiology would resolve this latter question, just as no amount of physical or mathematical analysis of the structure of a computer programme could resolve the former. The criteria for saying whether animals think, understand, infer etc. rest on the nature and the complexity of the behaviour which they manifest. We will only explain their behaviour as normative if it satisfies the various criteria that we have for saying of a creature that it is following a rule. In other words, in order to ascribe normative behaviour, there must be not simply regularity, but also a fair amount of external criteria whose complexity is such that we call it rule-following behaviour. Yet the argument is by no means a simple behaviourist claim : a point which is particularly borne out, in fact, by considering the mechanist thesis.

There is, of course, a well-known argument which was designed precisely in order to deflate this objection : the so-called 'Turing Test'. Turing's hypothesis was that if in a blind game situation we were unable to tell whether the hidden opponent we were playing against was a man or a machine, we would then have to conclude in the cases of machine opposition that such a machine could think. As it stands, however, the test accomplishes nothing of the sort, and if anything would indicate the very opposite from what Turing had intended : viz. that the criteria for ascribing normative behaviour can be extremely complex, and in a situation such as Turing's 'blind' game context, we are simply not afforded enough evidence on which to form the judgment whether we are dealing with normative behaviour. But of course, that was not at all the conclusion that Turing drew. For him the 'imitation game' sufficed to establish that, if we were unable to differentiate between the behaviour of a machine and a human being, then we would have no choice but to say that that machine thinks.[13]

Whatever the actual drawbacks of Turing's version of the 'imitation game', it is clear that the point of the Turing Test is supposed to be that if we could not distinguish between machine operations and human behaviour in identical circumstances, then we would have substantiated the claim that machines think. And since this test is simply an empirical problem, and there are no *a priori* reasons to deny the feasibility of such technological accomplishments, there is no conceptual barrier to the central thesis of AI. We can see from the cluster of normative concepts in the ascription of rule-following behaviour, however, that Turing failed to appreciate the full logical implications of his argument. For it involves supposing that the machine could perform the full range of actions which licenses our ascription of thinking : explaining, asking questions, correcting mistakes, instructing another etc. Yet this

argument does not in itself suffice, however, to undermine the Turing Test. For this is just the point which Turing believed he had accounted for, and which so many following him have stoutly defended Their claim is that the suggestion of a contradiction which the mechanist thesis seems to present owing to the fact that we dismiss the idea that human behaviour *tout court* could be simulated by a machine, but that when we treat the various criteria involved case by case we can demonstrate the viability of machine parallels, thus dispersing the air of paradoxicality which surrounds this issue. Thus, Turing's answer to this objection would simply be : make the feedback as complex as you like, there is no *a priori* reason why this should be deemed mechanically impossible. But what we have to see here is that the problem with the Turing Test is not that it is inadequate for Turing's purposes : it is that it is *au fond* mosconceived. For Turing confused two separate issues : the *empirical* question whether we could tell if something was human or a machine, and the *logical* question whether this 'test' would license us in the conclusion that machines can think. As far as the former—empirical—issue is concerned, the crux of the matter is to see exactly that : that it is an empirical issue.

There is certainly no *a priori* reason to deny the possibility of constructing a machine which was so sophisticated that it fooled everyone who came into contact with it into concluding that it was human. (For that matter, we should not exaggerate how sophiscated such a machine would have to be ; *vide* the reactions to Weizenbaum's mock psycho-analytic programme 'Eliza'.) But the basic question that we are confronted with here is, what exactly should we say if we subsequently discovered that we had been dealing with a machine : that the machine should be described as *thinking*, or simply that we had been fooled into assuming that such was the case ? Unfortunately, this is precisely the question which the Turing Test skips over ; or rather, it tries to break the argument

off just before this question. Yet this contains the vital heart of the matter : for the crucial answer is that my judgment was wrong, not that machines must *per impossible* be able to think. As soon as I learn that it was a machine I realize that someone had created a marvellously clever programme for correlating my input with symbols in such a manner as to convince me that I was dealing with intentional behaviour. (One might say, the responses which would have to so convince us have grown increasingly sophisticated through the course of human history, but still remain extraordinarily primitive, as the reaction to Eliza demonstrated.)

Where Turing went wrong, therefore, was in his belief that my judgement must in some sense be indefeasible. If something appears to satisfy my criteria for ascribing normative behaviour it cannot, he felt, subsequently turn out that I was wrong. Thus, Turing sought to tailor the concept of thought to these indefeasibility conditions. And it is precisely this step which was wrong. For the demarcation between *normative behaviour* and *mechanical symbol manipulation* cannot be obliterated by empirical regularities, however sophisticated these might conceivably become. The latter operation can never break into the closed circle of normative concepts : for the very reason that that is how we distinguish between *normative behaviour* and *mechanical operations*. Hence it is not our concept of *thought* which must change : it is our understanding of the defeasibility of any particular judgment that something is actually thinking.

Quinton argues in 'Machines with Minds' that 'before we can say that machines think in the sense that really interests us we have to equip them with feelings, desire and emotion. But the fact that we are still a long way from doing so does not imply that it cannot be done.[14] Once again it is clear that what is an *a priori* question has been reintegrated as an empirical problem. But Quinton has certainly placed his finger on the important point : to say of a machine that it thinks entails that a host of other 'human'

of "conscious' characteristics can be attributed to it. The difficulty in accepting Quinton's conclusion, however, lies wholly in the unintelligibility of ascribing thought or pain to a prosthetic apparatus :

> If one thinks of thought as something specifically human and organic, one is inclined to ask "could there be a prosthetic apparatus for thinking, an inorganic substitute for thought ?"
>
> But if thinking consists only in writing or speaking, why shouldn't a machine do it ? "Yes, but the machine doesn't know anything." Certainly it is senseless to talk of a prosthetic substitute for seeing and hearing. We do talk of artificial feet, but not of artificial pains in the foot.
>
> "But could a machine think ?"—Could it be in pain ?—Here the important thing is what one means by something *being in* pain. (PG 105).

The point is that, by trying to apply the very concept of *behaviour* (and *a fortiori, normative behaviour*) to the mechanical manipulation of symbols, the mechanist thesis is continually forced to transgress the bounds of sense. Advocates of the mechanist thesis gladly accept that in order to say of a machine that it thinks we would also have to be able to say of it that it understands : that it graps the meaning of the symbols with which it is operating. But the direction that this pulls the mechanist in is diametrically opposed to the fundamental premise of his thesis, which just is that the machine performs its operations mechanically, on meaningless symbols. It is manifest, however, that the logical point which is operating here has not been grasped, and it is not surprising to find the claim that the very strength of Turing's thesis is that it manages to 'circumvent meanings'.[15] Clearly, this is a point which cannot be passed over lightly.

The themes that have so far been touched upon have been handled in far too cursory a manner to be regarded as satisfactorily disposing of the various problems involved. Rather, they should be seen as offering a survey of the areas in which the discussion should proceed. However, it might be felt that, no matter how far you carried any of these individual investigations, the argument still operates at far too general a level to be seriously damaging to the mechanist thesis. Accordingly, we must descend still a level lower, to what might justly be described as the bowls of the mechanist thesis. For where the basic conceptual trouble underlying the mechanist thesis lies is, perhaps, in the initial assumption that, as Donald Knuth describes it in 'Algorithms' : 'an algorithm is a set of rules or directions for getting a specific output from a specific input. The distinguishing feature of an algorithm is that all vagueness must be eliminated : the rules must describe operations that are so simple and well defined that they can be executed by a machine.'[16] Overlooking the problem of what it means to say that rules 'describe' anything, we might feel that, in so far as it goes, this conception is entirely in accord with the way in which we interpret algorithms as rules which *we* calculate. What Knuth is discussing, however, are the nature of the rules which a machine calculates. Before we can investigate the nature of these 'rules', however, we should be asking ourselves what it means to speak of rules in the first place in the context of the mechanical manipulation of symbols. And this is precisely the—philosophical - problem that has been entirely overlooked.

This attitude has led to the pervasive confusion that algorithms can be described as complex systems of meaningless sub-rules, each of which can be applied purely mechanically. To begin with, we should consider what it means to describe a rule (or a 'sub-rule') as 'meaningless' ; tautologies and contradictions are meaningless, but what sort of rule would a 'meaningless rule' be ? (Presumably, one which told us nothing). Yet the point of this manoeuvre is clear :

to reduce rules to a body of seemingly mechanical components so that it will appear plausible to say that, by performing the totality of these operations the machine is in effect following a rule. Thus Gregory J. Chaitin argues in 'Randomness and mathematical proof' that :

> Instructions given to the computer must be complete and explicit, and they must enable it to proceed step by step without requiring that it comprehend the result of any part of the operations it performs. Such a program of instructions is an algorithm. It can demand any finite number of mechanical manipulations of numbers, but it cannot ask for judgments about their meanings.[17]

There are important confusions buried in this argument. How, first of all, could you apply a rule without comprehending that you had ? To be sure, we can map certain behaviour patterns onto rules, but that does not mean that the original agents were *following those rules.* Yet that is precisely the confusion which has been made here, for while the step in a programme might well be a rule that we follow, that does not license the conclusion that the computer is *following* that rule. By transgressing this point an intentional concept has, so to speak, been slipped in through the back door. Certainly there is a distinction to be drawn between *mechanical symbol manipulation* and (humanly effective) *calculation* ; the philosophical problem we encounter is how to describe these machine operations, and how to elucidate the distinction between them and the correct application of a rule/algortithm.

One source of the confusion operating here can be traced back tot he belief that Turing had really analyzed what we understand by *calculation* and *computation.* Indeed, the 'analysis' of *calculation* which Turing introduced in 1936 has been celebrated as, in the

words of Joseph Weizenbaum, 'one of the greatest triumphs of the human intellect.[18] And Martin Davis insists that Turing provided us with 'a cogent and complete logical analysis of the notion of "computation".[19] When he says this he certainly does not mean that Turing simply introduced a new concept which he called "machine computability' ; on the contrary, Turing is supposed to have revealed to us what the concept of computation as ordinarily understood *really* means. Davis goes so far as to claim : 'Thus it was that although people have been computing for centuries, it has only been since 1936 that we have possessed a satisfactory answer to the question : "What is a computation?"[20]

There is a fundamental confusion running throughout this argument. Turing's 'analysis' did not take us from a vague to a precise concept of *calculation*, but rather, constructed an entirely new concept which Turing misleadingly called by the same name. For what Turing had done was to introduce a non-normative concept which applies to the mechanical manipulation of symbols. Of course, there is nothing to stop us from creating such a new concept which is to be called 'calculation' ; where the danger lies is in supposing that this new version is in some sense more true to the facts than our previous (normative) concept of calculation. Our task here, therefore, is to distinguish between these two different concepts, whose only similarity is homophonic. That is, we must elucidate the rules governing the application of Turing's new concept of 'calculation', and elucidate how these differ from the logical grammar of *calculation* as properly understood.

This is very much the opposite, however, from the manner in which Turing's 'analysis' was received. Wang, for example, insisted that the problem we are concerned with here is 'How can we formalize, i.e. bring into a sharper form, the notion of a systematic procedure ? If we begin with a vague intuitive concept, how can we

find a sharper concept to correspond to it faithfuly ?[20] This is a dangerous source of confusion-; for by assuming that the sense of the notion of a 'systematic procedure' was indeterminate, Wang could then prepare the way for the conclusion that the Turing 'analysis' offered a more precise version of the same concept. But the notion of a 'systematic procedure' was neither vague nor intuitive : it simply did not perform the task which Turing sought for the programming of his machines, and consequently Turing was forced to introduce a new concept which would be suited to the mechanical purposes that he had in mind. In fact, Wang himself was uneasy with Turing's claim that the definition of 'computability' was equivalent to Church's concept of 'effective calculability',[22] remonstrating that 'A closer look reveals that the sharp notion, often referred to as recursiveness or Turing computability, is actually not as sharp as it appears at first sight.[23] Wittgenstein was also profoundly unhappy with the Turing thesis, but for reasons totally divorced from Wang's constructivist qualms, and indeed, reasons which would have rendered him equally hostile to Wang's own position.

In *Philosophical Remarks* and *Philosophical Grammar* Wittgenstein moved sharply away from the conception of logical analysis adumbrated by Davis and Wang towards a method which, as we saw above, strives to achieve an *Ubersicht* of the use of a concept. In the *Tractatus* Wittgenstein had formally pledged his allegiance to the cause of logical analysis, conforming with the view that philosophy seeks to define concept-words in such a way that the necessary and sufficient conditions for the correct application of a concept are specified. Definition is thus carried out by a method of 'logical decomposition' which breaks a concept down into its constituent logical elements (what Frege called its *Merkmale*, and Wang, echoing this, called its 'basic or essential characteristics'), bringing us ultimately to the logically simple (i.e. indefinable) constituents of

the concept. But Wittgenstein dramatically turned away from this conception of analysis in the early 1930s, abandoning the *Tractatus* conception of formal concepts (many of which were henceforward described as 'family resemblance' concepts). The meaning of a concept is strictly determined by the totality of rules governing its use within a *Satzsystem*, but certain concepts are closely related to one another by overlapping *Satzsysteme* parallels in their uses. The crrucial idea here was that there is nothing which is fundamentally common to every member of the family. but rather, the features which create the parallels between the uses of the concepts in the network of *Satzsystem*-family, and hence logical analysis in the Frege/Russell/early-Wittgenstein mould is a fruitless task because it rests on an illusory picture of the type of quarry that philosophical elucidation is seeking.

The later Wittgenstein was thus fundamentally opposed to the sort of conclusion offered by Davis that 'Turing based his precise definition of computation on an analysis of what a human being actually does when he computes'.[24] ; or as Wang described it : 'What Turing did was to analyze the human calculating act and arrive at a number of simple operations which are obviously mechanical in nature and yet can be shown to be capable of being combined to perform arbitrarily complex mechanical operations'.[25] Moreover, Wang followed this with a qualification which was the very point which Wittgenstein rejected : 'It is often not necessary [in logical analysis] to be entirely faithful. So long as a hard core is preserved, trimming on the margin is quite acceptable. Rigidity of formalized concept leads to decisions in cases where mere use of the intuitive notion was insufficient'.[26] Wittgenstein not only rejected the notion that there is any such 'hard core' –the fundamental property described above –, but he also rejected the assumption that such a formalized concept would enable us to decide cases of the application of the concept for which the 'intuitive' concept

proves 'insufficient'. He denied the demarcation between 'formal' and 'intuitive' concepts precisely because he denied the suggestions that there can be 'gaps' in a concept.

Furthermore, he denied the basic Godelian idea that a formal system is simply a mechanical procedure for producing theorems. Godel's basic idea was that :

> The concept of formal system requires that reasoning be completely replaced by 'mechanical operations' on formulas in just the sense made clear by Turing machines. More exactly, a formal system is nothing but a many valued Turing machine which permits a predetermined range of choices at certain steps. The one who works the Turing machine can, by his choice, set a lever at certain stages. This is precisely what one does in proving theorems within a formal system.[27]

It is fascinating to see the crucial shift which occurs in this passage. To be sure, the operator can programme the machine to correlate certain symbols according to what correspond to his choices at various stages, but that does not mean that the machine itself is making the choices !

Significantly, Wittgenstein would have agreed that the concept of *reasoning* is completely inappropriate here ; but then, the force of his point renders Godel's intended use of the expression 'proving theorems' entirely spurious. Wittgenstein vigorously opposed the suggestion that meaningless formal systems exist, and hence he rejected the very assumption that the 'derivations' of such systems (which are actually supposed to constitute the system) are purely mechanical. The problem here is that it is unintelligible to speak of a 'mechanical deductive structure'. For the two notions operating here—*inferring* and *meaningless symbols*—cannot be fused together You can speak of comparing the orthography, size etc. of

meaningless marks, but you cannot speak of deducing meaningless marks in a *purely formal system*. Hence it makes no sense to speak of a computer's 'making inferences'. We can programme inferences into a computer, but all that the computer is doing is correlating the various symbols which we regard as inferential patterns. Thus, to return to the argument from *Remarks on the foundations of Mathematics* V §2 (quoted above), if someone accidentally presses the knobs of a calculating machine and obtains the result $25 \times 20 = 500$ he has not thereby *calculated* the product of 25×20. For 'it is essential to mathematics that its signs are also employed in *mufti*' That is, without grasping the mathematical application, you are not actually calculating. Likewise, unless you grasp that q follows from p—where to understand the nature of conceptual relation between q and p *just is* to know that q follows from p—it is equally unintelligible to speak of inference.

Wittgenstein certainly did not deny that we can construct or imagine mechanical procedures for manipulating the symbols of a proof, but where we must be especially careful is in how we describe such a mechanical operation. In what sense would it constitute 'working over' the proof ? (RFM V §3). The danger here is in that of saying that such a machine was 'testing the correctness of the inferences'. It is the meaning of this latter phrase that we are concerned with, not with the 'accuracy' of the machine's operations. And the point that Wittgenstein wanted to make is that concept of *testing the correctness of an inference* rests on the presupposition that these inferences have been understood as such : that the proof has been grasped as a proof. A machine can perform mechanical operations which yield symbols that seemingly correspond to our own checks of validity, but in so doing it is not testing the correctness of the inferences, for the concept of *checking inferences* only has meaning when applied to the significant concepts of proof and inference.

For all of these reasons we must see the sort of exercise which Turing undertook, not as the clarification of the concept of computation, but rather, as the constructlon of a new concept, operating under different (mechanical) rules of logical grammar. The whole problem with this issue lies in the very framework in which we approach these questions, and not in e.g. the limitations imposed by a finitist position. Wang was particularly concerned to elucidate the nature of Godel's distinction between a *mechanically effective procedure* (i.e. a well-defined procedure which need not always terminate) and a *humanly effective procedure* (which must terminate), in order to bolster the Godel/Kresel objection that the Turing Thesis only establishes that 'every *mechanically* effective computation can be simulated by a Turing Machine'. But the philosophical point that we are really concerned with here is the intelligibility of the expression 'mechanical computation', and this is just the point which Wang overlooked when he described :

> a distinction which is often overlooked. What is adequately explicated is the intuitive concept of mechanical procedures or algorithms or computation procedures of finite combinatorial procedures. The related concept of effective procedures or constructive procedures, meaning procedures which can in the most general sense be carried out, suggests somewhat different elements which are related to the difference between mental and mechanical procedures and the question as to the method by which a Turing machine or a set of equations is seen to be one which defines a Turing computable or general recursive function.[28]

The problem with this argument, from Wittgenstein's point of view, does not concern the issue of classical versus constructive mathematics ; rather, the problem is to consider the logical grammar of the concepts of *mechanical operation* and *calculation*, in order to clarify the logical barriers which separate the two.

To illustrate the distinction between *mechanically effective* and *humanly effective* calculation, Wang gave as an example the case of a pupil learning how to apply a Euclidean algorithm correctly without knowing why it gives the correct results : Giving an algorithm for solving a class K of problems and a problem belonging to K, anybody can solve the problem provided he is able to perform the operations required by the algorithm and to follow exactly the rules as given. For example, a shoolboy can learn the Euclidean algorithm correctly without knowing why it gives the desired answers'.[29] The example does indeed draw our attention to a very important aspect of the problem ; for, before we can even consider the issues raised by the assumption that an algorithm can be calculated mechanically when the agent/machine does not understand why it gives the correct results (or what it means to understand 'why' in this sense ; i.e. whether machine can in some sense be said to understand 'why' the algorithm yields the correct results), we must first clarify what it means to say that the shoolboy has learned how to apply the Euclidean algorithm. Thus the immediate philosophical problem here concerns the conditions under which we would say that 'the pupil has grasped the rule', 'the pupil has applied the rule correctly'. Some are prepared to concede that a machine can be programmed to apply a rule correctly, but it cannot be programmed to learn or understand a rule. But this seemingly innocuous schoolboy example indicates precisely where the confusion in this latter argument lies.

There is obviously something profoundly out of focus in the final sentence of the above passage ; the thoughts which it contains are pulling in opposite directions, crediting the schoolboy with the very piece of knowledge which is immediately denied him. For to learn how to apply an algorithm correctly involves more than merely producing the 'right' results. In order to say of the schoolboy that

he has learnt—grasped—the rule, we will demand more than simply the set of his results to justify such a judgment, The criteria for crediting someone with the mastery of a rule are far more complex than this ; we place it against the background of his explaining, justifying, correcting, answering certain questions, etc. We would no doubt be extremely puzzled by the case of a schoolboy who could variably give the 'right' results for an algorithm and yet could provide us with absolutely no information about how or why he had derived that result. The fact that in such a case we would be extremely hard-pressed to accept that he had learned the algorithm is evidence of the complexity of the criteria underpinning our use of the concept of *mastery of a rule.*

Presuambly what Wang meant in the message quoted above, however, was that a schoolboy can perform each step of the algorithm correctly, but cannot see the overall pattern connecting each of these individual tasks ; and it was in this sense that he said that the schoolboy has learned the algoritm without knowing why it gives the desired results. But this is not at all what *learning the algorithm* is taken to mean. All that we could rightly say in such a situation is that the schoolboy has learned a series of (for him) independent rules ; but to learn how to apply each of these sub-rules does not amount to learning the algorithm. (Someone who learns the individual rules for all the chess-pieces without grasping that the point of the game is to mate his opponent's king has not learn how to play chess)

This in turn points to yet another problem involved in Wang's account, for we must also clarify what is involved in saying that the schoolboy has learned how to apply the individual sub-rules of the algorithm correctly. It is noteworthy that Wang himself used the word 'learned', and did not suggest that the schoolboy has merely applied the rules mechanically. For the fundamental point here is

that there could be no method of moving from such a mechanical performance to the conclusion that the schoolboy has learned the algorithm. And even more to the point is the fact that there would not even be any method of moving from such a mechanical operation to the conclusion that the schoolboy has learned how to apply the sub-rules of the algorithm correctly. To have accomplished the latter, the schoolboy would, of course, have had to have grasped the sub-rule.

The problem here, therefore, is that we cannot reconcile the two concepts—*mechanical operation* and *learning the correct application of a rule*—with one another. And to try to fuse the two ideas together into a single concept—a *mechanical rule*—is a straightforward contradiction, postulating a non-normative normative concept. If an operation is mechanical, then there is an empirical probability (which in the case of powerful computers can be exceptionally high) that the operation will yield the symbol which we describe *a priori* as the correct calculation of the algorithm. The point is that to say of an agent/machine that he/it has learned how to apply the algorithm correctly presupposes that he/it has grasped the rule it is applying. The result of applying an algorithm correctly is the consequence of understanding and following a rule, whereas the result of performing a mechanical operation successfuly in the consequence of a successful experiment.

Of course, there are many who are quite happy with this objection, for they cheerfully respond that in that case we shall simply say that the machine has understood the algorithm it has calculated. The underlying feeling here is that machines can in some sense be said to think if, as Wang put it, they can be programmed to perform operations which 'imitate' the operations of the human mind.[30] And in order to recognize the presence of such 'imitation' we do not turn to the neurophysiologist to aid our inquiry, but rather,

refine (in the manner of Turing's analysis) the nature of the algorithms which the human mind (brain ?) computes. Hence, we are supposedly led to see that :

> The intuitive notion of an algorithm is rather vague, For example, what is a rule ? We would like the rules to be mechanically interpretable, i.e. such that a machine can understand the rule (instruction) and carry it out. In other words, we need to specify a language for describing algorithms which is general enough to describe all mechanical procedures and yet simple enough to be interpreted by a machine...What Turing did was to analyze the human calculating act and arrive at a number of simple operations which are obviously mechanical in nature and yet can be shown to be capable of being combined to perform arbitrarily complex mechanical operations.[31]

Perhaps the most notable feature of the extensive discussion of this topic is that despite the emphasis on the mechanical nature of the procedures that are being devised, the really important issue that underlies this—the relation of the machine to this programme—is charaterized in terms which clearly beg or confuse the issue. Passing over Wittgenstein's objections to the concept of constructing such a suitable (i.e. formal) 'language', what really concerns us in the above passage is the unreflecting assumption that the machine interprets the programme which it follows ; i.e. that in some special context it makes sense to say that a machine understands or follows a rule, even though it cannot grasp the meaning of a concept The important problem here is to clarify the nature of this *cannot*. For it is not the nature of the programme that should be worrying us here : it is the very manner in which we describe the interaction between the machine and this programme, and here we find Wang slipping into an idiom which presupposes the very issue which we are ultimately supposed to be establishing. The reason for this

biased form of description would seem to be that, since Wang continued to think (tacitly) of the sub-rules of the algorithm within a basically normative conceptual framework, he implicitly assumed that it makes sense to describe the programme input as a series of instructions which the machine must naturally interpret if it is to follow the programme correctly Thus, for all the preoccupation with defining 'complex mechanical operations', the language in which the problem is couched remains firmly committed to the fundamentally normative concept of an algorithm.

As we can see, therefore, the ultimate problem with the mechanist thesis is that we have a system of interlocking concepts—all internally related—which rest on the fundamental concepts of understanding, meaning, and of course, following a rule. The mechanical manipulation of symbols cannot break into this conceptual circle at any point. A computer programme is no more an algorithm (which must be understood normatively) that is the roll in a pianola, and no more a proof than Jacquard's cards were a proof of the patterns that were woven. Obviously the mechanist thesis is not going to succumb without a spirited struggle. Let us simply assume for the sake of argument, however, that the Wittgensteinian philosopher has effectively managed to work his way through all the various levels of conceptual confusion involved in the mechanist thesis, not simply to his own, but more importantly, to the satisfaction of the AI-philosopher. The ultimate question which then arises is : can the Wittgensteinian philosopher expect his efforts to be crowned with success in the form of the mechanist thesis ? Unfortunately, he must be prepared to accept that the matter is by no means so straightforward. For in speaking of the 'spirit of the times', we are not just referring to an intellectual inability to grasp the nature of Wittgenstein's conception of philosophy. At what is perhaps an even more fundamental level, we are

speaking of a basic disinterest in the fruits yielded by Wittgenstein's approach.

All of the emphasis in this paper has been devoted to the issue of how our understanding of Wittgenstein's conception of the nature of philosophy should govern our actual practice of the resolution of a philosophical problem. But no reason has yet been given why anyone should actually welcome the insights thus afforded. And it is not at all clear that there is any compelling *philosophical* answer to this question. Certainly the Wittgensteinian philosopher cannot *assume* that, as soon as a 'scientific' philosopher has grasped and accepted the outlines of the Wittgensteinian solution to a philosophical problem he will automatically renounce his earlier pledge to the scientistic cause. Perhaps there are elements here of what is essentially an ethical issue, with overtones of the Grand Inquisitor's speech in *The Brothers Karamazov*. For one of the hallmarks of the spirit of the modern age is that teconological progress is regarded as an overriding goal : so much so that it little matters how it is obtained. Indeed, perhaps the real allusion that should be made here is to *Faust*.

In *The Fifth Generation* Feigenbaum and McCorduck warn that Japan enjoys a distinct advantage over the West in AI-research because 'they've spent no time at all in those arid little debates so beloved by Western intellectuals, debates centred on the question whether a machine really can be said to think. They regard our obsession with that topic the same way we regard their eating raw fish—an odd, puzzling, but harmless cultural quirk. Instead, their debates are about the best way to design an intelligent machine'.[2] The Wittgensteinian philosopher must tread warily here. Our most natural instinct is to try to respond to the mechanist thesis by seeking to illuminate what type of society will evolve if these confusions are allowed to persist unattended. The obvious worry is

that if you institute a conceptual revolution in the concept of *thought* so that it henceforward becomes intelligible to describe mechanical operations as *thinking*, then conversely there seems little reason why the argument should not proceed in the opposite direction, thereby denying human beings the notions of autonomy and consciousness which underpin our conception of man as a *rule-following creature*. And if this conceptual revolution is allowed to proceed unhindered, the social and political consequences can surely not be far behind. Yet AI-philosophers invariably respond that this argument is hollow, because our experience has hitherto been exactly the opposite : viz. computers have only served to enrich our daily life and human potential. But, of course, our complaint here has nothing whatsoever to do with the benefits of computers *per se*—which without question are one of the most exciting tools that mankind has devised—but rather, with the conceptual confusions that surround this marvellous technical discovery. AI-researchers like Feigenbaum and McCorduck can brusquely dismiss philosophers' anxieties as illusory (if not hysterical), and yet in the very next breath describe the wonders of the 'geriatic robot' : 'The very best thing about the geriatic robot', Feigenbaum and McCorduck write, 'is that it *listens*. "Tell me again," it says, "about how wonderful/dreadful your children are to you. Tell me again that fascinating tale of the coup of '63. Tell again....." And it means it. It never gets tired of hearing those stories, just as you never get tired of telling them. It is difficult to know how to respond to someone who can accept such a prospect with no qualms whatsoever.

Like all great ethical issues, every individual must ultimately choose for himself in which direction he wants society to proceed, and what shall constitute his values of social harmony and personal fulfilment. But our task as philosophers is to try to enable the individual to make his decision on the basis of clarity rather than confusion. And yet even here, we cannot expect that understanding

itself will guarantee the nature of his decision. In *Philosophical Grammar* Wittgenstein argued that 'Philosophical clarity will have the same effect on the growth of mathematics as sunlight has on the growth of potato shoots. (In a dark cellar they grow yards long'). (PG 381) But the 'scientific philosopher' might well argue that he actually *prefers* the potato shoots : that the greatest scientific discoveries come from the completely unexpected. After all, what if Wittgenstein had phrased his argument : philosophical clarity will have the same effect on mathematics as refrigeration has on the development of bread mould ? So the Wittgensteinian philosopher must only resort to the *practical* benefit of conceptual clarification with extreme caution ; for it could well be that it is the very confusions enshrined in the mechanist thesis which have provided the driving force behind the astonishing development of AI. Perhaps without the picture of a thinking machine to guide them, scientists would not have been nearly so eager to develop 'expert systems' nor have achieved such astonishing strides in developing prosthetic apparatuses which simulate complex human abilities. And perhaps it was in response to just such a Faustian theme—which runs throughout Western thought—that Wittgenstein was led to define the nature and role of philosophy in deeply religious terms :

> Tolstoy : 'The significance of an object lies in its universal intelligibility'. That is partly true, partly false. When an object is significant and important what makes it difficult to understand is not the lack of some special instruction in abstruse matters necessary for its understanding but the conflict between the right understanding of the object and what most men *want* to see. This can make the most obvious things the most difficult to understand. What has to be overcome is not a difficulty of the understanding, but of the will. (CV 17).

Here, finally, we can see the intrinsically ethical foundation of Wittgenstein's conception of the nature of philosophy which from his earliest efforts provided the mainstay of his thought. It was for precisely this reason that I argued above that Wittgenstein's conception of the nature of philosophy transcends the cultural values or attitudes of any particular epoch and applies to the 'human condition' in general. Indeed, as he takes up his burden the Wittgensteinian philosopher might do well to reflect on Milton's description of 'false Philosophie' in Book II of *Paradise Lost*, for it highlights archetypal dimensions of the struggle to vanquish 'false philosophy':

> Others apart sat on a Hill retir'd,
> In thoughts more elevate, and reason'd high
> Of Providence, Foreknowledge, Will, and Fate,
> Fixt Fote, Free Will, Foreknowledge absolute,
> And found no end, in wand'ring mazes lost.
> Of good and evil much they argu'd then,
> of happiness and final misery,
> Passion and Apathy, and glory and same,
> Vain wisdom all, and false Philosophie :
> Yet with a pleasing sorcery could charm
> Pain for a while or anguish, and excite
> Fallacious hope, or arm th' obdred breast
> With stubborn patience as with triple steel.

So too, Wittgenstein saw genuine philosophy as a never-ceasing struggle, not simply against the confusions enshrined in any particular society or culture, but more importantly, against one's inner self : against the constant temptation to abandon the arduous quest for conceptual clarity and accept instead a 'pleasing sorcery' which 'charms Pain for a while or anguish, and excite Fallacious hope' :

To say, when they are at work, "Let's have done with it now", is a *physical* need for human beings ; it is the constant necessity when you are philosophizing to go on thinkig in the face of this need that makes this such strenuous work. (CV 75,6)

FOOTNOTES

1 Anthoney Kenny, *The Legacy of Wittgenstein*, Oxford, Basil Blackwell, 1984, p. vii.

2. *Ibid.*

3. *Ibid.*, p. vii.

4. cf. 'On Scientific Method in Philosophy', in ***Mysticism and Logic***, London, Longmans, Green & Company, 1918, and 'The Value of Philosophy', in ***The Problems of Philosophy***, Oxford, Oxford University Press, 1959.

5. ***The Problems of Philosophy***, p. 70.

6. B. Russel, ***My Philosophical Development***, London, George Allen & Unwin, 1959, p. 161.

7. 'The Philosophy of Logical Atomism', in *Logic and knowledge*, ed. R.C. Marsh, London, Ge ° rge Allen & Unwin, 1977, p. 193.

8. M. Dummet, *Frege : Philosophy of Language*, London, Duckworth, 1981, p. xxxiii.

9. J. C. Webb, ***Mechanism, Mentalism, and Metamathematics***, Dordrecht, D. Reidel Publishing Co., 1980, p. 24.

10 Hofstadter, *Godel, Escher, Bach*, Harmondsworth, Penguin, 1980, p. 26.

11. *Ibid* , pp 729,30.

12. A. Quinton, 'Machines with Minds', 3 London, BBC Publications, March 1983, p. 19.

13. A. Turing, 'Computing Machinery and Intelligence', in *Minds and Machines*, ed. A.R. Ross, Englewood Cliffs, Prentice-Hall Inc., 1964, p. 6.

14. 'Machines with Minds', p. 21.

15. *Mechanism, Mentalism, and Metamathematics*, p. 225.
16. Dr. Knuth, 'Algorithms', *Scientific American* 236 (April 1977),
17. G. J. Chaitin, 'Randomness and Mathematical Proof', *Scientific American* 232 (May 1975) p. 47.
18. J Weizenbaum, *Computer Power and Human Reason*, San Francisco, W.H. Freeman and Company, 1976, p. 58.
19. Martin Davis, 'What is a computation', in *Mathematics Today*, ed. L.A. Steen, New York, Springer-Verlag, 1978, p. 241.
22. *Ibid.*
21. Hao Wang, *From Mathematics to Philosophy*, London, Routledge & Kegan Paul, 1976, p. 81.
22. 'On Computable Numbers', in *The Undecidable*, ed. M. Davis, Hewlett, Raven Press, 1965, 117.
23. *From Mathematics to Philosophy*, p. 81.
24. 'What is a Computation', p. 243.
25. *From Mathematics to Philosophy*, p. 91.
26. *Ibid.*, p. 81.
27. *Ibid.*, p. 84.
28. *Ibid.*, pp. 89, 90.
29. *Ibid.*, p. 90.
30. *Ibid.*, p. 83.
31. *Ibid* , p. 91.
32. E. A. Feigenbaum and P. McCorduck, *The Fifth Generation*, London, Pan Books, 1984, p. 31.
32. *Ibid.*, p. 124.

Department of Philosophy Punjab University
LIBRARY
8880