# Motivations

D. GOLDSTICK

The contested question which a major strand in the argument of the following essay is concerned to take up is that of the *possibility* of computer mentality. Those inclined to answer the question with a blanket negative are reluctant even to credit the contraptions, literally, with purpose or motivation, let alone *creativity*. Those defending the opposite view point out that the brains of even the most creative human beings are every bit as physical as any computer. At that physical level of description, just as with computers—they insist—there is nothing there but matter and energy, governed by all the laws of physics and chemistry. At a 'higher' level of description, analogous to the *design level* of description for a machine such as a computer, a brain's operations—just like the operations of any of the other separate parts of living things—are adequately described *biologically,* in terms of organic functioning, it is argued. At that level of description, so the story goes, the operations described are indeed (in an everyday sense of the word) decidedly 'mechanical' for both a brain and a computer. It is only at a still higher level of description, the *intentional* level, that anything like creativity can be identified. It is only at that level of description that any psychological vocabulary at all can properly be employed. That, at any rate, is the way the argument runs. But how independent of one another can these different 'levels' really be? It is surely implausible to suggest that just *any* physical or organic basis could underlie bona fide mental life.

What is called for will be a 'top-down' inquiry into the very idea of anything such as a *motivation,* for instance. An argument which will be advanced here is that, if computers as we know them could think, they would think computationally. Their ends would be 'lexically ordered', and they would follow a decision procedure in determining what to do to achieve them. But anything worth calling a mind, I shall argue, is motivated ultimately in a different fashion than that, and does not just follow an algorithm when it is thinking. Can what it does instead be characterized in non-circular terms, or must it be pronounced immune to even the limited-depth explication contemplated here? That the reader must be left to decide.



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

* * *

In Ronald Dworkin's famous book on the philosophy of law, *Taking Rights Seriously,* a sharp distinction is drawn between what Dworkin calls 'rules' and 'principles'.

> Only rules dictate results, come what may. When a contrary result has been reached, the rule has been abandoned or changed. Principles do not work that way; they incline a decision one way, though not conclusively, and they survive intact when they do not prevail.[1]

The rational case for the authority or weight of particular principles is generally not demonstrative, Dworkin insists.

> There is no litmus paper for testing the soundness of such a case—it is a matter of judgment....[2]

The distinction between rational considerations which 'dictate' a conclusion and rational considerations which do not is relevant not only to the theory of judicial decision making. It applies to the process of reaching any conclusion, whether the reaching of a conclusion about what to do or a conclusion about what is the case. The mind exercising its 'judgement' attends to different competing considerations and allows itself to be differentially influenced by them. Nor need the considerations influencing this process always be *conscious*. The outcome of more or less long 'experience' contributes to 'good judgement'—or, alternatively, to unreasonable 'prejudice'—in ways that the mind so conditioned is seldom, if ever, able to specify fully. But in any event the various influences favouring different conclusions in their tendency are apt to have unequal 'strengths', as we say, and the resulting conclusion will reflect that.

It is not only in juristic contexts that the 'weighing' metaphor strongly suggests itself in this connection. On the face of it, there exists a radical contrast between arriving at a conclusion in this kind of way, on the one hand, and arriving at a conclusion by following a *decision procedure,* on the other hand. Not that the latter, *algorithmic* way of thinking cannot also be found in use in both the great and small aspects of human life. One uses algorithms for simple mental arithmetic every day. And it is the characteristic of human *bureaucracies,* whether in the private or the governmental sector, to save time and trouble, and minimize 'inconsistency' by means of laying

[1] Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA, 1977), p. 35.

[2] Ibid., p. 36. The bearing of a legal precedent, in common-law jurisdictions, is explained by Dworkin on page 111 in the evocative words, 'the earlier decision exerts a gravitational force on later decisions'.



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

down decision procedures for resolving most of the issues by which their front-line officers are ever confronted in the normal course of business. It certainly is very often possible to serve good ends through the following of explicit or implicit decision procedures.

By 'following' a decision procedure here, I mean acting *on purpose* as the decision procedure prescribes. Since one is not necessarily, in whole or part, conscious of each of one's purposes, it is possible to *follow a decision procedure* without doing so *consciously*. One no more needs to be conscious of one's habitual thought procedures than of one's 'automatic' shoelace-tying procedures. On the other hand, to *act in accordance* with a particular decision procedure it is not necessary to be *following* it, in the present sense, as the *purpose,* whether conscious or not, of acting in accordance with that decision procedure may be completely missing. A conscious mind does not need—any more than something inanimate—to be doing *on purpose* whatever it does.

These last distinctions will enable us to answer the objection that thinking really is entirely algorithmic after all, given that it is the deterministic result of finitely complex initial conditions in conjunction with finitely complex psychological and other laws. The idea behind this objection is that, when different considerations and combinations of considerations (whether consciously entertained or not) influence the more or less receptive mind, as they do, to differing degrees, these differences in their influentiality must be due to some common factor—call it 'Apparent Weightiness' or 'A.W.'—and the decision procedure in accordance with which the deliberating mind is proceeding, in effect, is one prescribing *the maximization of deference to A. W.* The objector here is not arguing for any readily apparent commensurability among the different considerations entertained by the deliberating mind, but only insisting that it does succeed, in practice, in rendering them mutually commensurable, at least by its lights, in allowing them severally to influence it to various differing degrees. However, in so far as the mind in question lacks any purpose, conscious or otherwise, of *maximizing deference to A. W.,* it is not *following* any such decision procedure.

And so, the objection that thinking really is all algorithmic fails. But the point the objection is making is well taken, though, in its insistence that, to the extent the deliberating mind is *rational*, it will think in accordance with a rule—akin to the rule of 'universalizability' in ethics—to the effect that sufficiently similar considerations ought to influence it in the same direction and to the same degree, so that the various operative considerations' influentiality will indeed be due to, and covary with, some (possibly very complex) common factor, such as we have called 'A.W.' here. (Possibly very complex, but not unlimitedly complex; for, to the extent that



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

the deliberating mind is *rational,* it is predisposed already as regards how any conceivable considerations would influence it; and mental predispositions of *unlimited* complexity are beyond any finite being's capacity.)

We are left, then, with a distinction between two genuinely different ways of thinking. When people call the *algorithmic* way of thinking 'unimaginative', the implication surely is that *imagination* especially creative *imagination,* somehow partakes of, or is akin to, the other, non-algorithmic way of thinking. Let me advance the suggestion that what it is for a mind to *imagine* something is for that mind to pretend to itself that it believes it, i.e., it is for that mind to *make believe* that such a fact exists, to *make as if* it so believed—in other words, in its own intrinsic state to *map* the actual or imaginary fact in what would be the belief-mode but for the lack of any real causal hook-up to potential acting-upon behaviour[3]—and to allow itself then to be influenced as a result in such a way as to give rise to pretend-beliefs of this sort on other, related points. (E.g., 'If she had used the gun, it would have her fingerprints on it now'.) In *creative imagination* the mind willingly allows itself to be affected in this sort of way by influences of which it hasn't any conscious notion until they manifest themselves in the imaginative result. This, of course, will normally make that result *unforeseeable* by it.

* * *

At least since Plato's *Gorgias,* the difference has been appreciated between two logically distinct sorts of *goal*. It will be recalled that Plato there compared a pleasure-seeker to someone perpetually

[3] As far as non-normative opinions go, I argue for the thesis of inseparability from behavioural dispositions in my 'Belief', *American Philosophical Quarterly,* Vol. 26, No. 3 (July, 1989), pp. 231–38. However, isn't any hook-up at least to the outward manifestation of behaviour cut off during sleep most of the time, despite the fact that sleepers are apt to *believe* what they dream? Certainly one sort of hook-up to outward behaviour is mostly cut off during sleep, but another sort is not. Usually under sleep conditions no outward result is, in fact, attainable just at will. But it remains true that, if such a behavioural result *were* to be so attainable (and was not at all doubted to be), and if that behaviour were on balance desired because it was *believed* to be in the circumstances a sufficient precondition for some further result or other, in such a case the behaviour in question would then ensue (because *then* the believer concerned would in fact be awake). The logical belief-desire-behaviour relationship is further discussed below.



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

engaged in trying to fill a leaky jar.[4] No matter how much pleasure is secured, for no matter how long, the result achieved must inevitably fall short of a *conceivable* result more pleasurable still. In this regard, the objective of securing pleasure distinctly differs from that of shutting the door, or that of coming to know whether or not p (for any proposition-expressing substitution instance of the 'p'). Let us say that a goal is 'infinitistic' if and only if, for any conceivable state-of-affairs, a state-of-affairs is conceivable wherein that goal is realized more fully. A goal can be said to be a 'finitistic' one otherwise.[5] And corresponding to this distinction between two different sorts of goal is a distinction between two different sorts of motivation, *infinitistic* and *finitistic,* depending on whether the (simple or complex) goal that its possessor is motivated to pursue is an infinitistic or finitistic goal.

Could the motivations actuating any mind fail to be infinitistic *as an ensemble?* Might they, for example, fail to be infinitistic as an ensemble by failing at all to constitute jointly any one ensemble? But

[4] See Plato's dialogue, the *Gorgias,* 493b3–494a3. In the dialogue, Socrates attributes the conception to an unnamed (Pythagorean?) 'wise man'.

[5] The expression 'conceivable' is used here rather than 'logically possible' because some goals logically impossible to realize to any degree whatsoever are still quite capable of being pursued in good faith: e.g., that of 'squaring the circle'; or of amassing a fortune of $n, where n is the highest prime number. A state-of-affairs is 'conceivable' if and only if it is logically possible to *believe* it to be realized. (For the distinction between logical impossibilities which are and which are not logically impossible to *believe,* see my 'Cognitive Synonymy', *Dialectica,* Vol. 34, No. 3 (1980), pp. 183–203, and my 'When Inconsistency is Logically Impossible', *Logique et Analyse,* Vol. 32, Nos 125–26 (March–June 1989), pp. 139–42.)

In the present context, a conceivable *state-of-affairs* should be understood as an intentional object specifiable in value-neutral non-modal terms entirely. That will make it possible, e.g., to disqualify *perfection* (in contrast, say, to *a perfect darts score)* from counting as any sort of finitistic goal, since, no matter how fine any state-of-affairs is, another, superior one will always be *conceivable*—if only through *conceiving* the characteristics that make for perfection to be different from what they actually are (a sort of ruling which would not be possible if *perfection* itself could be considered a 'conceivable state-of-affairs' for these purposes). By the same token, *making as much money as possible in the circumstances,* for example, can be disqualified as any sort of finitistic goal, on the grounds that, no matter how much money is made, a state-of-affairs is always *conceivable* in which still more money can be made and is *(possession of the maximum wealth possible in the circumstances* not being allowed to count as a 'conceivable state-of-affairs' for these purposes).



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

just how could those motivations, however much in competition and even in conflict with one another, nevertheless count as being all the motivations of *one mind* if there were not, at any time, some aggregation of them into a *single will,* however complex, and however vacillating over time? Could *such* an ensemble of motivation fail to be, overall, infinitistic? Only if the mind so motivated had *no* infinitistic goals; and there would then have to be something determining which of its finitistic goals it would pursue in cases of an apparent clash between any two or more of them—something such as, for example, a 'lexical ordering' among all the mind's finitistic goals, so that it would only be motivated to realize any one of them on condition that this would not in the slightest reduce the chances of realizing any others of higher priority. It is easy to see why no biologically evolved mind is like that. For that would be incompatible with the characteristic flexibility of response which gives mentality its survival value in comparison to the mindless forms of life with their 'purely mechanical' reactions to environmental conditions. It seems not unreasonable taxonomically to specify this flexibility of response, and hence the infinitistic character of the life form in question's ensemble of motivation, as an *essential feature* of 'mentality'. And so, wherever there is a *mind,* its motivation is infinitistic overall, and if it is ever motivated to pursue a finitistic goal, this is only because of that goal's contribution, or supposed contribution, to the infinitistic goal(s) which it is motivated to pursue.

A bureaucrat, for example, may be motivated by the finitistic goal of reducing rule-infractions to zero. But why? It may well be the prospect of reward (including the reward of avoiding penalization); or it may be the particular external goals of the organization served which supply the bureaucrat's deepest motivations; or it may instead merely be loyalty to that bureaucratic organization; or then again, lastly, the bureaucrat may just be personally *pained* by any infractions of the rules. In all of these cases but the last, the finitistic goal specified is not what motivates the bureaucrat ultimately. And in the last case the motivation, even if non-derivative, is inherently overridable. Even in the case of an 'ideal bureaucrat' who is not pained by anything else except rule-infractions, and has no other motivation whatsoever, *can* all rule-infractions, and all combinations of rule-infractions, be exactly equally painful, so that there is no motivation to prefer the avoidance of some rule-infractions or combinations of rule-infractions to the avoidance of others? It is possible, indeed, to imagine a super-rule laying down an order of preference here, but could all infractions, and all combinations of infractions, of *this* rule be exactly equally painful to our imaginary bureaucrat? We of course could imagine a super-super-rule to deal



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

with this, but the same question would arise again in relation to it. And so on. At any rate, it is obvious that no real, flesh-and-blood bureaucrat is anything like that.

The extentialists used to make a point like this in their own way, perhaps, when they insisted that, whatever thought-and-action-dictating *persona* one adopts, there is always a motivationally deeper level at which its adoption by one is (ever anew) an undictated *free choice* on one's part.

Not only is any *mind's* motivation necessarily infinitistic overall, it cannot have any finitistic goals at all to which its attachment is strictly absolute and non-overridable. Such goals would have to hold absolute priority for it, under any conditions, over all infinitistic motivations, and the flexibility of response characteristic of *mentality* as such would thus be lost. And so, it seems reasonable to require this as well of any *mind* worthy of the name.[6]

On any psychological theory of motivations, they are all either inborn or else acquired, on the basis of inborn motivations, through processes of 'canalization' or conditioning. Accordingly, the infinitistic character of all organically inborn motivations ought to determine at any rate the potential overridability of all the motivations thus derived from them, as well as of themselves.

* * *

Here is a viewpoint which appears to go at least some way in a contrary direction.

> Human needs arrange themselves in hierarchies of pre-potency. That is to say, the appearance of one need usually rests on the prior satisfaction of another, more pre-potent need....
>
> If all the needs are unsatisfied, and the organism is then dominated by the physiological needs, all other needs may become simply non-existent or be pushed into the background. It is then fair to characterize the whole organism by saying simply that it is hungry, for consciousness is almost completely preempted by hunger. All capacities are put into the service of hunger-satisfaction, and the organization of these capacities is almost entirely determined by the one purpose of satisfying hunger.... For the

[6] Everybody has their price, as the proverb says. *Could* there be a rule-follower so mindless as to be absolutely undeflectable by all conceivable appeals to greed, compassion or any other infinitistic motivation whatsoever? No, for to be absolutely mindless is no longer even to be a *rule-follower.* Such is one consequence of our present understanding of *mentality.*



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

> man who is extremely and dangerously hungry, no other interests exist but food. He dreams food, he remembers food, he thinks about food, he emotes only about food, he perceives only food and he wants only food....
>
> For our chronically and extremely hungry man, Utopia can be defined very simply as a place where there is plenty of food. He tends to think that, if only he is guaranteed food for the rest of his life, he will be perfectly happy and will never want anything more. Life itself tends to be defined in terms of eating. Anything else will be defined as unimportant....[7]

In these vivid words, first published in 1943, the famous psychologist Abraham Maslow does seem to be saying that certain finitistic motivations, such as that of securing at least a certain minimum of food to eat, take absolute precedence over all other human motivations; which would make a goal like the satisfaction of elemental hunger absolutely non-overridable psychologically. And, indeed, instances of people who have deliberately starved themselves to death for any reason are known to be rare. Does this mean that most people, by contrast, are in fact absolutely under the dominion of such a finitistic motivation as that of securing at least a certain minimum of food to eat?[8] Maslow, at any rate, expressly qualifies what he has said on this score:

> So far, our theoretical discussion may have given the impression that these five sets of needs are somehow in a step-wise, all-or-none relationships [sic] to each other. We have spoken in such terms as the following: 'If one need is satisfied, then another emerges.' This statement might give the false impression that a need must be satisfied 100 per cent before the next need emerges.

[7] A. H. Maslow, 'A Theory of Human Motivation', *The Psychological Review,* Vol. 50, No. 4 (July, 1943), pp. 370, 373–74.

[8] Maslow is surely right to call attention to circumstances that on certain occasions in effect operate to render humans or other animals blind to any possible facts that might deflect them from 'single-mindedly' pursuing some particular limited goal. But this need not stop that motivation from nonetheless being overridable, in the sense that if certain conceivable facts were known, or at any rate believed, to be actual, that goal *would* then be given up. Even less can the intellectual impairment of a mind preoccupied with hunger or whatever prevent its motivation from remaining in principle *overridable* in the still weaker sense—the pertinent one for these purposes—that if it did *care enough* about something else *and* knew or believed certain conceivable facts to be actual, then the goal it now pursues would be immediately given up. (I am indebted to Ernest Benz for pointing out the need for this clarification of the meaning of 'overridable' here.)



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

> In actual fact, most members of our society who are normal, are partially satisfied in all their basic needs and partially unsatisfied in all their basic needs at the same time. A more realistic description of the hierarchy would be in terms of decreasing percentages of satisfaction as we go up the hierarchy of prepotency.[9]

* * *

The discussion of motivations so far has relied on a *weighing* analogy and an analogy with opposed physical *tugs* tending to pull an object in different directions. How much better than that can we do? Surely what motivational tugs are, in the first place, is causes—causes potentially militating against each other's effects, the eventual results depending upon their different relative magnitudes. With words like 'militating' and possibly even 'magnitudes' we are, of course, still immersed in metaphor. What more can we say?

To characterize a motivational cause of behaviour as infinitistic is to assert that the result which it tends to produce is not any definite single state-of-affairs but rather *movement in a certain direction* towards states-of-affairs of ever higher preferential priority. For each such motivation, that is, there is an asymmetrical, transitive relation ranging certain conceivable states-of-affairs in their reverse order of preference (from the standpoint of that motivation).[10]

We cannot say that each such motivation, in so far as it were unopposed, would actually result in *any* of those states-of-affairs ensuing—if only because the being thus motivated might not be *able* to secure just any result it desired. Under conditions of normal functioning there are certain results, such as simple body movements, which are, indeed, attainable at will, and by means of which such a being endeavours to attain everything else which it seeks. But the results sought may in fact not be attainable by those means (or by any means). However, any time that something is done as a means to some result, it is the case that the probability of that result occurring *would* be increased thereby if certain propositions were all true, namely those assigning to each point on which the being in question has any non-normative opinion a probability corresponding to whatever degree of confidence the being has on that point. These probability estimates need not be completely precise. But

[9] Op. cit., p. 388.

[10] The suggestion is not being made here that overall preference is (or is not) asymmetrical and transitive. What is being suggested is that any mind's overall motivation is factorable into particular infinitistic *desires,* whose conceivable fulfilments-and-nonfulfilments do (in each case) line up in that way.



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

what *is* definite is that, if (maybe *per impossible)* they *were* all to be true, then the being's actual behaviour *would* increase the likelihood of the outcome sought.[11]

We cannot necessarily say that this behaviour occurs because of its aptness to produce a certain result (here, the increasing of an aimed-at outcome's probability).[12] What we can say is that the behaviour occurs because of the fact that, if certain (probability-estimating) propositions were all true, that behaviour then would increase a certain eventuality's likelihood of ensuing.[13]

An infinitistic desire which some being has is a cause of a certain specific kind: it *tends* to produce certain occurrences as a result of (whatever it is which accounts for) the fact that, if certain (probability-estimating) propositions were all true, those occurrences would then raise the probabilities of certain further outcomes ensuing, those probability increases being ordered by an asymmetrical, transitive relation of such a sort that which of them actually would result if all those propositions were true would depend upon what other such causes were currently present in the being in question, the greater, in any instance, the *magnitude* of the particular cause in question relative to that of each other such cause present in the being, the further along in the order generated by that relation the (probability-increase) results which would be caused to ensue.

Two problems remain outstanding. The first is the possibly metaphorical language whereby a motivational cause of behaviour is said to have a particular 'magnitude' relative to other such causes present in the being that is motivated. The second problem is that, while for any mind there certainly are occurrences—'behaviour'—which under conditions of normal functioning it could produce just

[11] See my paper 'Belief', op. cit.

[12] Cf. Charles Taylor, *The Explanation of Behaviour* (London, 1964), p. 12: 'The first term of a teleological correlation ... identifies the antecedent condition of the event to be accounted for, *B,* as a state of affairs in which *B* will lead to *G*. Thus the antecedent is identified in terms of its law-like connexions with two other events, *B* and *G*, i.e., as that state of affairs in which, when *B* occurs, *G* will follow.'

[13] On occasion something is done because of the known fact that doing it will increase the probability of some *undesired* result: for instance, I accidentally drop a fragile plate as a result of nervousness due to the realization that dropping it will lead to intense embarrassment or worse. (Cf. Roderick M. Chisholm, 'Freedom and Action', in *Freedom and Determinism,* edited by Keith Lehrer (New York, 1966), pp. 29–30.) But outcomes like that do not result *in a systematic way* from my motivations. Indeed, even where I do drop the plate, I *would* take means to prevent that happening if I could. And so, *systematically* functioning causally in this way is not just necessary but also sufficient for being a *desire.*



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

at will, the conditions actually obtaining do not have to be at all normal. Quite apart from total paralysis, etc., it might even be inexact not to count nightly *sleep* as 'normal functioning'.

To deal with the second problem first, we can say that motivations are not functioning normally *as causes* unless there actually is some behaviour producible just at will (and known to be, of course; or at least believed to be; but, unless this is known or believed, that behaviour is not, in fact, producible just at will). As with a spark plug, whatever constitutes a *desire* does so by being of such a character as to function causally in a certain way under certain conditions. The way in question is the way of functioning causally which we have been discussing. The 'conditions of normal functioning' here are just those conditions in which that sort of functioning does occur.

There remains the problem of the different 'magnitudes' of desires, thanks to which they are comparable with one another as 'stronger' or 'weaker', more or less causative—non-motivational factors permitting—of the sort of results which they are desires *for*. The idea is that there is something intrinsic about any desire which accounts for its motivational causativeness and which admits of variations rankable in a series with no limit at one end—that is, with no limit in the direction of increasing magnitude—and which is of such a character that, whenever some desires which a being has fail to produce, in the normal direct way, the behaviour which they would produce, in the normal direct way, but for other desires the being has which in the normal direct way produce that behaviour's *non*occurrence, the latter desires jointly count as 'stronger' than the former desires jointly with respect to the occurrence of that behaviour then. By 'the normal direct way' here is meant the normal way in which desires produce behaviour *otherwise than through the taking of means to its production*. In a clash of desires such as is here envisaged, one set of desires does not come out as stronger than another set 'absolutely', but only *with respect to the occurrence of specific behaviour* of this sort *on a given occasion,* because how important that behaviour's occurrence or nonoccurrence is believed to be for securing the different ends favoured by the two sets of desires will necessarily play its part in determining which set 'wins out' on the given occasion. If you desire X and Y equally strongly and believe, with complete confidence, that behaviour B is producible or preventable at will and will, if produced, increase X's likelihood of realization from 5% to 95% but decrease Y's likelihood of realization by only 5%, from 95% to 90%, and if no other considerations have any influence on your decision, then you surely will resolve on behaviour B. (It is worth taking note of the point that while finitistic



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

desires like the desire for X and the desire for Y in this example *can* indeed have 'absolute' strengths in the psyche of the being possessing them, infinitistic desires cannot, because each one of those, with whatever magnitude it possesses, will involve a motivation to pursue any number of different finitistic goals with various different degrees of keenness.)

It remains true that an infinitistic desire cannot express itself in action except by way of giving rise to finitistic desires. However, this does not mean, on the other hand, that it is impossible to have an 'unimplemented' infinitistic desire, which has not (yet) actually given rise to any finitistic desires. How separate are the *existences* of an infinitistic desire and the (oftentimes multiply motivated) finitistic desires which do 'implement' it? It would no doubt be desirable to have a developed theory interrelating the two—their existences, their causal roles, and their magnitudes and degrees of strength—but this is as far as it seems possible or necessary to go in elaborating all that here.

* * *

Can computers think? Not—evidently—unless they can *follow algorithms*. They would use them to determine the optimal behavioural response to different existing situations, in pursuit of their ends. The ends of any computer mind would have to be lexically ordered if there were ever to be more than one of them motivating it. But our preceding considerations have led to the conclusion that no *mind* worthy of the name ever follows an algorithm (or does anything on purpose) unless it is motivated to do it by motivation that is infinitistic. In other words, whatever ultimately motivates the following of a decision procedure on any occasion does so by making that course *more attractive* than any alternative way of proceeding in the circumstances. It does so, that is, by *outweighing* any contrary motivation. And computers, as we have made them and know them, do not think like that. Undoubtedly they do operate *in accordance with* algorithms. But we cannot say that they do this—or do anything—*on purpose,* since ultimately all purposes are infinitistically motivated: and so, being without any purposes, computers are mindless and, accordingly, do not think at all, even algorithmically.

All the same, however, because the present paper throughout is so determinedly 'top down' in its approach, readers may well want to ask whether this paper does really rule out anything specific in the way of hardware. And more than one reader of an earlier version of the paper has brought up the question whether a Turing machine programmed to do something again and again indefinitely



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

could not qualify as 'infinitistically motivated' by the criteria developed here. Indeed, such a description does *remind* us of Plato's tireless leaky-vessel filler. The question is, though: could such a Turing machine qualify as *motivated* to do what it does indefinitely?

Instead of answering off the bat that no *deterministic* mechanism could possibly possess even a lower-animal-type 'point of view', it will surely be uncontroversial, in the first place, to take note that at any rate being programmed to do the same thing over and over again indefinitely cannot by itself count as a *sufficient* condition for being motivated in any manner, let alone infinitistically. A machine programmed to travel ten metres in a westerly direction, and then repeat, and so on, obviously *could* be mindless. If such a machine were to be seriously credited with infinitistic *motivation*, this would involve postulating that the completion of each leg of its journeying is somehow a thing the machine does *because* this is a step towards further travel (or, rather, because it would be if certain conditions obtained). Can the fact that the bare definition of a Turing machine does not actually rule that out open the door to the possibility of genuine thinking computers in the future?

It is not that science-fiction conceptions are out of order in this discussion. Of course if artificial thinking machines ever were to be fashioned in the future they might well incorporate computer 'logic'.[14] Would that then make them computers? The range of what can be described as a 'computer' is indeed very wide. Standard textbook definitions of the term are broad enough to encompass (for instance) even an abacus. And there is no need here to deny (or affirm) that at some level of description the real brain processes involved in thinking may in fact be all describable 'computationally'. But, on the other hand, we do not call animals 'plants' even though they do indeed possess cell structure and other features characteristic of plants.

In the abstract it is easy enough to *imagine* machines, just as deterministic in their operation as any computer, which incorporated, for example, electromagnets attracting pieces of metal unequally in different directions, and which arguably thus produced results in a way analogous to that in which the outcome of a mental 'exercise of judgement' is produced. But literal thinking machines like that, were they ever to be constructed, would not much resemble the machines which today we call *computers*. (Certainly it is possible at least to conceive of a computer programmed to 'calculate' the resultant of a set of force vectors and then react differentially in accor-

[14] I am grateful to Martin Hilgerdenaar for insisting on this point in debating the possibilities for 'computer thought'.



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press

dance with whatever that resultant is. But the resultant will have been *literally* (mentally) calculated only if the calculator was ultimately motivated to perform that algorithmic computation by infinitistic—and hence overridable—motivations.[15])

Two-and-a-half centuries ago, some scientific-minded thinkers were bold enough to dream of artificially constructed machines which one day would be capable of replicating even the intelligent behaviour of animals, including humans. As the most sophisticated machinery then known operated by means of clockwork, it tended to be along clockwork lines that these machines were envisaged as being constructed. If literal thinking machines are ever constructed in the future, will the mechanisms by which they operate be as remote from those of our computers as *they* are, in turn, from the clockwork mechanisms that so thrilled our illustrious predecessors?

*University of Toronto*

[15] The sense in which 'connectionist' models of actual brain functioning are *non-computational* is this, that, even if a properly programmed computing device *could,* in principle, 'calculate' such a brain's outputs from its inputs (and its workings), the connectionist approach does not need to view a functioning brain as itself a mechanism which does that with respect to its own sensory inputs and motor outputs.



https://doi.org/10.1017/S0031819100000486 Published online by Cambridge University Press