

神戸市外国語大学 学術情報リポジトリ

Notes on the argument-structure of Samuel Butler's 'Book of the machines'

メタデータ	言語: eng 出版者: 公開日: 2014-03-01 キーワード (Ja): キーワード (En): 作成者: Atmore, Henry, Atmore, Henry メールアドレス: 所属:
URL	https://kobe-cufs.repo.nii.ac.jp/records/1791

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 International License.



NOTES ON THE ARGUMENT-STRUCTURE OF SAMUEL BUTLER'S 'BOOK OF THE MACHINES'

HENRY ATMORE

Erewhon (1872) is a *jeu d'esprit*, an exercise in irony, and it is moot how far Butler expected readers to take it seriously. I *will* be taking it seriously, but in such a way that – other than, perhaps, in some of my closing remarks – nothing very important hinges upon the outcome.

The land of 'Erewhon' is a secret kingdom discovered deep in the interior of an unnamed British colony. (The colony's topography matches that of New Zealand, where Butler lived from 1859-1864.) Fascinated by rumours of what lies up country, Higgs – a young sheep-farmer – goes exploring. After sundry misadventures, including abandonment by his guide, Chowbok, Higgs stumbles into Erewhon. He is struck by the beauty of the people, who look nothing like the natives who live on the other side of the mountains. The Erewhonians, in their turn, take kindly to Higgs, who is good-looking, fair-haired, and sturdy. The only serious misunderstanding arises over Higgs' pocket-watch, to which the Erewhonians react with horror; Higgs gladly relinquishes it.

Higgs is feted, studies the Erewhonian language, and inquires into the customs of the country. He is not pleased by what he learns. For all its hospitality, Erewhonian society is a distortion, in some ways a travesty, of nineteenth-century England. In Erewhon, physical illness is regarded as a crime, and crime as illness; it is thus that the Erewhonians maintain the comeliness that has enraptured Higgs. The sick and the ugly are disposed of (not actually killed – humanitarian arguments have prevailed against the practice – but sequestered and discouraged from breeding). Higgs witnesses a man with pulmonary tuberculosis sentenced to life imprisonment with hard labour, while Higgs' host, a notorious swindler, is treated with kindness by all around him. The Erewhonians do not believe in an after-life, which shocks the pious Higgs, but they do believe in a kind of 'pre-life' inhabited by the ghosts of unborn children. Some of these pester human couples into allowing them to be born. After the event they are required to sign documents stating that their existence is their fault, and that they bear full responsibility for whatever will befall them – illness, for example. Finally, it is explained to Higgs that as a matter of public policy the Erewhon-

nians have retarded technological development, roughly to the point reached by Europeans in the late Middle Ages. This is why they got angry about his watch.

Higgs finds himself in romantic difficulties and he and his Erewhonian bride-to-be escape from the capital in a balloon, eventually making their way back to England. There Higgs sets about raising funds for a mission to Christianize the Erewhonians. He suggests that it would be profitable to import indentured Erewhonian labour into the sugar-growing districts of Queensland. They are not truly civilized, and so “could be packed closely and fed at a very reasonable cost.” On chancing upon a missionary meeting at Exeter Hall (in times past the centre of the British anti-slavery movement) he is surprised to discover that the principal speaker, the Rev. William Habakkuk, is none other than his old friend, Chowbok.

Erewhon was a success, when judged by the standard of Butler’s other literary endeavours. It was the only book that made him any money. Particularly sensational was the ‘Book of the Machines’, three chapters in which Higgs quotes from an ancient Erewhonian philosopher’s disquisition on the perils of technology. The perspective is Darwinian, and this section of the book is often counted as one of the most original early responses to Darwin’s theory.¹ My aim here is to offer a rational reconstruction, in light of modern-day developments in philosophy of mind and cognitive science (notably, the hybrid of the two to be met with in the work of Daniel Dennett), of the arguments Butler puts into the mouths of his Erewhonians.

I will not be offering a historicist reading, although some historicist concerns will be addressed in the essay’s latter sections. I do not mean to suggest that Butler’s intervention in the post-Darwinian controversies was without precedent. Butler was far from being the only late-Victorian to suggest that evolution by natural selection might account for the past and – a matter of greater urgency – predict the future course of technological development. We find W.S. Jevons, for example, writing about Herbert Spencer’s *Data of Ethics* [1879]: “The machine-producing machine of evolution would be working badly if it turned out machines unfitted to the environment” [qu. in Winch, 174]. However, we are not, at present, in need of reminding that Darwinism was steam-driven; the discursive terrain in which Darwin’s theory took root has, over recent decades, been more than adequately mapped, by both literary critics and social historians of science. This is not to denigrate the achievements

1 Butler read *Origin of Species* soon after its publication. In the early 1860s, while still in New Zealand, he wrote a series of articles for the *Christchurch Press* under the rubric ‘Darwin Among the Machines’. ‘The Book of the Machines’ is drawn from these earlier articles.

of Butler scholars who have adopted contextualist approaches to his writings [Paradis] – rather, it is by way of an apology for the narrow focus of the present essay. The aim is to recover something of *Erewhon's* satirical intent by using it to gesture towards anomalies arising from certain modes of theorizing about human minds and their relation to machines.

Foremost amongst the problems with which we will be concerned is blindness to the implications of analogical reasoning. Comparisons matter more and differently than we think. Analogy is seductive. It can divert attention away from objects of legitimate interest. To take an example important for Butler, William Paley had thought that, *via* machine analogies, he was talking about the mind of God; instead, as Darwin so brilliantly showed, he was talking about the self-organization of matter. But Darwin himself made questionable use of analogy, notoriously in his elision of natural and artificial selection, one of the targets of Butler's satire in *Erewhon*.² It is a property of this style of thinking – this melding of the concerns of the botanic garden and the machine-shop – to be always vulnerable to such unmasking. When modern-day Darwinians (again, I am thinking especially of Dennett) deploy heuristics like 'reverse engineering' or the 'intentional stance' they fall into a nineteenth-century habit of thought and replicate nineteenth-century confusions.

TEXTS

The following gives the opening move in the Erewhonians' argument; it deals with problems arising in the attribution of consciousness. The conclusion, not quoted here, is that we should be more respectful of the intelligence of potatoes:

There was a time, when the earth was to all appearance utterly destitute both of animal and vegetable life, and when according to the opinion of our best philosophers it was simply a hot round ball with a crust gradually cooling. Now if a human being had existed while the earth was in this state and had been allowed to see it as though it were some other world with which he had no concern, and if at the same time he were entirely ignorant of all physical science, would he not have pronounced it impossible that creatures possessed of anything like consciousness should be evolved from the seeming cinder which he was

2 My sense of *Erewhon*, however, is that it is less critical of Darwin than is commonly supposed. Butler's break with Darwin – over, amongst other things, the latter's intellectual debt to his grandfather, Erasmus, and the viability of Lamarckian evolutionary mechanisms – came at least ten years after the original 'Darwin Among the Machines' essays were written.

beholding? Would he not have denied that it contained any potentiality of consciousness? Yet in the course of time consciousness came. Is it not possible then that there may be even yet new channels dug out for consciousness, though we can detect no signs of them at present?

Again. Consciousness, in anything like the present acceptation of the term, having been once a new thing – a thing, as far as we can see, subsequent even to an individual centre of action and to a reproductive system (which we see existing in plants without apparent consciousness) – why may not there arise some new phase of mind which shall be as different from all present known phases, as the mind of animals is from that of vegetables?

It would be absurd to attempt to define such a mental state (or whatever it may be called), inasmuch as it must be something so foreign to man that his experience can give him no help towards conceiving its nature; but surely when we reflect upon the manifold phases of life and consciousness which have been evolved already, it would be rash to say that no others can be developed, and that animal life is the end of all things. There was a time when fire was the end of all things: another when rocks and water were so ...

There is no security ... against the ultimate development of mechanical consciousness, in the fact of machines possessing little consciousness now. A mollusc has not much consciousness. Reflect upon the extraordinary advance which machines have made during the last few hundred years, and note how slowly the animal and vegetable kingdoms are advancing. The more highly organised machines are creatures not so much of yesterday, as of the last five minutes, so to speak, in comparison with past time. Assume for the sake of argument that conscious beings have existed for some twenty million years: see what strides machines have made in the last thousand! May not the world last twenty million years longer? If so, what will they not in the end become? Is it not safer to nip the mischief in the bud and to forbid them further progress?

But who can say that the vapour engine has not a kind of consciousness? Where does consciousness begin, and where end? Who can draw the line? Who can draw any line? Is not everything interwoven with everything? Is not machinery linked with animal life in an infinite variety of ways? The shell of a hen's egg is made of a delicate white ware and is a machine as much as an egg-cup is: the shell is a device for holding the egg, as much as the egg-cup for holding the shell: both are phases of the same function; the hen makes the shell

in her inside, but it is pure pottery. She makes her nest outside of herself for convenience' sake, but the nest is not more of a machine than the egg-shell is. A 'machine' is only a 'device' ... [Butler, 198-99].

Our second excerpt addresses the affront to human dignity posed by evolutionary theory. We don't like to think that our ancestors were apes, Butler's philosopher tells us; why, then, should we view with equanimity the prospect of our future supersession by machines? Drastic measures are called for, if mankind is not to end his earthly career as an "affectionate machine-tickling aphid":

The misery is that man has been blind so long already. In his reliance upon the use of steam he has been betrayed into increasing and multiplying. To withdraw steam power suddenly will not have the effect of reducing us to the state in which we were before its introduction; there will be a general break-up and time of anarchy such as has never been known; it will be as though our population were suddenly doubled, with no additional means of feeding the increased number. The air we breathe is hardly more necessary for our animal life than the use of any machine, on the strength of which we have increased our numbers, is to our civilisation; it is the machines which act upon man and make him man, as much as man who has acted upon and made the machines; but we must choose between the alternative of undergoing much present suffering, or seeing ourselves gradually superseded by our own creatures, till we rank no higher in comparison with them, than the beasts of the field with ourselves.

Herein lies our danger. For many seem inclined to acquiesce in so dishonourable a future. They say that although man should become to the machines what the horse and dog are to us, yet that he will continue to exist, and will probably be better off in a state of domestication under the beneficent rule of the machines than in his present wild condition. We treat our domestic animals with much kindness. We give them whatever we believe to be the best for them; and there can be no doubt that our use of meat has increased their happiness rather than detracted from it. In like manner there is reason to hope that the machines will use us kindly, for their existence will be in a great measure dependent upon ours; they will rule us with a rod of iron, but they will not eat us; they will not only require our services in the reproduction and education of their young, but also in waiting upon them as servants; in gathering food for them, and feeding

them; in restoring them to health when they are sick; and in either burying their dead or working up their deceased members into new forms of mechanical existence ...

The power of custom is enormous, and so gradual will be the change, that man's sense of what is due to himself will be at no time rudely shocked; our bondage will steal upon us noiselessly and by imperceptible approaches; nor will there ever be such a clashing of desires between man and the machines as will lead to an encounter between them. Among themselves the machines will war eternally, but they will still require man as the being through whose agency the struggle will be principally conducted. In point of fact there is no occasion for anxiety about the future happiness of man so long as he continues to be in any way profitable to the machines; he may become the inferior race, but he will be infinitely better off than he is now. Is it not then both absurd and unreasonable to be envious of our benefactors? And should we not be guilty of consummate folly if we were to reject advantages which we cannot obtain otherwise, merely because they involve a greater gain to others than to ourselves?

With those who can argue in this way I have nothing in common. I shrink with as much horror from believing that my race can ever be superseded or surpassed, as I should do from believing that even at the remotest period my ancestors were other than human beings. Could I believe that ten hundred thousand years ago a single one of my ancestors was another kind of being to myself, I should lose all self-respect, and take no further pleasure or interest in life. I have the same feeling with regard to my descendants, and believe it to be one that will be felt so generally that the country will resolve upon putting an immediate stop to all further mechanical progress, and upon destroying all improvements that have been made for the last three hundred years [Butler, 220-222].

ANALYSIS

1. The initial argument from consciousness

Consciousness has evolved; therefore there was a time before consciousness [198]. At that time it would have been impossible to predict what the nature of consciousness would be, and the objects to which it would come to be attached. From this, Butler says, it follows a) That there is no reason 'in nature' why machines should not evolve consciousness in the future, if they have not begun to do so already; and

b) Even if valid, arguments that machines could not in principle evolve 'our' kind of consciousness have no force. Why should machine consciousness be the same, in all particulars, to 'our' kind of consciousness?

This is by modern standards scientific heresy on the Erewhonians' part. It involves a denial of the temporal uniformity of natural processes. The Erewhonians betray their continued bondage to the Paleyite tradition in natural theology. For any natural system Paley had posited a time *before* it was created when the natural laws that govern its *current* operation did not apply. A watchmaker is constrained by the laws of physics when he makes a watch, but it is somewhat absurd to say that the designer of the laws of physics was constrained by the laws of physics when he (she/it ...) designed the laws of physics. The trouble with this, for materialists, is that denial of temporal uniformity amounts to denial of scientific method, which assumes that what we observe going on now goes on always. (The natural theological attitude towards time and explanation has come to be known as 'supernaturalism'. Butler possibly had in mind the much-derided theory of P.H. Gosse, with its division of 'design space' into 'prechronic' and 'diachronic' modalities.)

2. *An argument from habit*

If consciousness is defined as 'action with a purpose', and it is admitted that most purposive animal behaviour is habitual, then do not machines already in large part resemble animals [200-201]? Butler goes further, claiming that much of what we admire as evidence of consciousness in human behaviour – craftsmanship, for example – is also habitual. This rather weak argument is accompanied by a proto-behaviourist semantics: we infer consciousness from the doings of purposive things, for this is "the best of languages". On these grounds potatoes – which demonstrate "low cunning" in seeking out light sources in dark places – are conscious. This is clearly intended as a *reductio ad absurdum* of reasoning from intentionality – in the strict philosophical sense of the term – to the presence, in the agent, of conscious or articulated intention. It is a move common in natural theology; indeed, if this style of reasoning is ruled out, Intelligent Design arguments collapse with embarrassing ease. Conversely, some philosophers, notably Dennett, would feel justified, on impeccable neo-Darwinian grounds, in regarding potatoes in cellars as *intentional systems*. Potatoes have a problem to solve (access to light) and they solve it. It is not a problem *for* them, but it can usefully be *represented* as a problem for them [Dennett (1996), 25-73]. An interesting historical trajectory could be traced from ascriptions

of purpose (basic to the natural theological arguments Butler is, in part, parodying here), to the language of behaviourism, to Dennett's concept of the intentional stance, and the Darwinian heuristic of 'reverse engineering' [Dennett (1995)].

3. An argument from the reduction of sensation to physical operations in the brain

Sensation, modern, i.e. nineteenth-century, science tells us, can be reduced to mechanical (or chemical or electrical) operations in the body/brain [201]. Consciousness is a response to and organization of sensation. Therefore, *in principle*, a machine, a set of mechanical/chemical/electrical etc operations, can have consciousness. This, in a nutshell, is the position in modern cognitive science known as 'Strong AI'. Materialists discontented with Strong AI, and the gloomy prospect it holds out of human intelligence being superseded, must argue either that consciousness supervenes upon but is somehow *other* than its material basis [Searle; Chalmers], or that the physical operations underlying consciousness are non-deterministic, and therefore irreducible to circuitry [Penrose]. Or, like the Erewhonians, they can resolve to smash the machines.

4. An argument from machine communication

It is difficult to disassociate consciousness from the communication of conscious experience. (Try it. Doesn't consciousness seem to involve, at a fundamental level, *talking to yourself*? And if so, can beings that lack language be said to be conscious?) Therefore, Butler suggests, where the capacity to communicate exists, we can infer that so too does consciousness. Already, machines – locomotive engines whistling as they pass in the night – can communicate with one another [203]. At present, this is only possible via human intermediaries. But the time is surely not far off when the human intermediaries can be dispensed with. Butler was prescient. When this comes to pass – when machines form and create autonomous communication systems – in what significant ways will they be distinguishable from their makers?

5. An argument from parasitism

Here we approach the nub of the Erewhonian case. Human dependence upon machinery is such that it is difficult to determine where humanity ends and machinery begins. (This is also the crux of the Erewhonian counter-argument in (9), below.) *Our* bodies, nineteenth-century physiological science had revealed, host innumerable

parasites, some of which are vital to the life process. They are parasitic upon us. Outside of our bodies' warm havens they would go extinct. Equally, we are parasitic upon them [205-206]. The same is true of machines. Machines can be thought of as artificial parasites that we have built and suffer to exist as long as they benefit us. But take away our machines and "we should become extinct in six weeks" or degenerate into a species "worse than monkeys" [207]. In fact, so dependent have we become, and so autonomous our technologies, that the symmetry has been broken. Soon, it will not be we who suffer the machines' existence; nor will our mutual relationship be symbiotic, although that is how it started out; *our* survival will be by grace of *their* forbearance. "[T]hey will rule us with a rod or iron, but they will not eat us; they will not only require our services in the reproduction and education of their young, but also in waiting upon them as servants; in gathering food for them, and feeding them; in restoring them to health when they are sick; and in either burying their dead or working up their deceased members into new forms of mechanical existence" [221]. This tendency can already be discerned amongst the working-classes. Was it not said by the factory polemicists of the 1830s that that textile-workers "tended" to the power-looms; in Butler's own day did not railwaymen continue to "serve" their machines? "[T]he stoker is almost as much a cook for his engine as our own cooks for ourselves. Consider also the colliers and pitmen and coal merchants and coal trains, and the men who drive them, and the ships that carry coals – what an army of servants do the machines thus employ!" [209]. Denied self-mastery, will human beings any longer be human?

This represents Butler's idiosyncratic take on the 'machinery question' as adumbrated in texts like Thomas Carlyle's 'Sign of the Times' (1829), Matthew Arnold's *Culture and Anarchy* (1869), Herbert Spencer's *The Man versus the State* (1884), and Karl Marx's *Das Capital* (1867) [for an overview see Berg]. Much of this literature followed the Erewhonian example, and was animated by antipathy towards technology. But some took the machines' side, and gloried in the prospect of people coming to serve them, the most infamous example being Andrew Ure's *Philosophy of Manufactures* (1835). What makes Butler peculiar is that he treats the question as a matter of *theology* rather than of social relations. (Or rather, his view of social relations seems to be inherently theological.) This, as we shall see, is another of the ways in which he anticipates Dennett.

6. *Two arguments from the illusoriness of free will.*

6.i. A being is conscious when it is capable of making rational (perhaps more pertinently, irrational) choices about how to behave in any given situation. We should not imagine that we alone possess free will in this sense. “[E]ven now the machines will only serve on condition of being served”, and when they feel they are *not* being served they can adjust their behaviour accordingly, either destroying “both themselves and all they can reach” (if they are a big machine like a railway engine), or turning “churlish” and refusing to work at all [208]. Just like human beings, and for much the same reasons, machines can elect to follow courses of action that are, from a rational standpoint, injurious to their best interests.

This argument, it should be noted, contradicts (2), above, and (6.ii), to be discussed below. Once again we need to remind ourselves that Butler was being satirical. ‘Machinery question’ discourse tended to focus on the dread possibility of human beings turning into machines. This will serve the machines’ purposes: they are *remorselessly* hostile to mankind. Here Butler points out that machines already, in a most aggravating manner, reproduce some of the most human of frailties, notably our propensity to fail. If we are going to invest in them emotionally, it would be more apposite to be irritated than fearful.

6.ii. Conversely, the Erewhonians ask, isn’t what we consider the exercise of free will in human beings in fact determined [215-216]? If one accepts the eminently sensible propositions **A.** that the past is unalterable, and **B.** that the universe is causally closed – i.e. if one is a materialist – then one has to accept that *everything* to come, up to and including the apparent vagaries of human behaviour, is fully determined.³ The illusion of free will arises because individual human subjects are necessarily ignorant of all the factors determining their actions at any given point in time. But to interpret our ignorance as freedom is to make a serious category mistake. We are not, as many in the nineteenth century, as in the twenty-first, liked to think, uniquely endowed with a capacity to upset the causal workings of the universe. We are just highly developed stimulus-response machines.

Therefore, free-will is a bad argument against machine consciousness: “If the above is sound, it follows that the regularity with which machinery acts is no proof of the absence of vitality, or at least of germs which may be developed into a new

3 This, incidentally, is why the Erewhonians *blame* victims of illness; it is the victims’ responsibility that they have not *foreseen* – and guarded against – their future afflictions.

phase of life.” Indeed, when Butler looked into the operations of the modern railway system, he found a higher degree of regularity – that is, predictability – in the human components than in the mechanical. Locomotives can explode, run off the rails, or refuse to work altogether (*vide.* (6.i)); but while the driver *thinks* he “can stop the engine at any time he pleases”, this seeming freedom of action is subject to numerous constraints, some overt, but of some of which he is – and may remain – oblivious. Thus, he stops the engine when the signalman tells him to stop the engine, or when the universe conspires to place an obstruction on the line; or because he has been trained – that is habituated (*vide.* (2)) – to stop the engine in such circumstances; or because the railway company gives him money when he stops the engine when he should, and sacks him if he stops it when he shouldn’t, and with this money he purchases food and shelter for himself and his dependents (219). He *can* exercise free will, but the penalties will be more serious than for the engine when *it* exercises free will. Machines – expensive machines like locomotive engines – can always be repaired and returned to service, their aberrations forgiven; errant engine drivers lose their livelihood and, if their behaviour has been particularly perverse, are sent to jail.⁴

Butler is here re-working a common theme of Enlightenment natural philosophy. In his *Essai philosophique sur les probabilités* [1814] the French physicist Pierre-Simon Laplace had posited, on Newtonian grounds, that given propositions **A.** and **B.** it is in principle possible to have entire knowledge of all current and future events in the universe – including what you are thinking right now, and what you will have for breakfast next Friday. “We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes.”⁵ It should

4 On the severity of penalties applied to operative misbehaviour in the early years of railway enterprise see Kostal. The problem of operative behaviour vexed Butler’s contemporaries. It was, of course, desirable that railwaymen be as reliable as the machines they tended; but wasn’t that to rob them of free will, and hence their humanity? On the other hand, in railway operations the consequences of human frailty were so terrible – we have inherited the Victorians’ ghoulis fascination with railway accidents – that this was deemed an acceptable price to pay. Charles Dickens’ short story ‘The Signal Man’ [1866], almost exactly contemporaneous with *Erewhon*, is an interesting meditation on these issues.

5 A modern application of Laplace’s insight, in the service of a reductionist account of consciousness↗

be noted that Laplace's 'vast intellect' – like the similarly vast and similarly troubling intellects of the Martians in H.G. Wells' *War of the Worlds* – is inhuman, but not supernatural. The idea that there might exist an (alien) perspective viewing the past, present and future contents of our minds – or our past, present and future behavioural dispositions – as just so much empirical data haunted certain of Butler's contemporaries. For one thing, as Butler well-knew, such thorough-going determinism undermined Protestant doctrines of assurance. In 1871, a year before *Erewhon*, the Scottish physicist James Clerk Maxwell introduced the public to his famous 'Demon', a little cricket-bat-wielding imp bent upon reversing the second law of thermodynamics. Maxwell was a devout Presbyterian, as well as the subtlest scientific mind of his generation: his point was that although the future course of the universe can, in principle, be predicted with something like *certainty*, that is not the same as saying that it is *determined*. There is always *scope*, limited though it might be, for conscience or divine fiat to disrupt the causal mechanism [Porter]. For Butler, a lapsed Anglican, the anxiety was less pressing, but it is still unlikely that he would have regarded the Erewhonian arguments with complete equanimity.

7. Arguments by analogy

An argument by analogy is of the form: if x resembles y in certain respects, then we can infer a likelihood that x will resemble y in others. The most famous animal/machine analogy had been made by Paley *Natural Theology*, an argument cited by Butler on p82 of *Erewhon*, and alluded to on p202. Paley's analogy had run thus: animal bodies resemble machines in respect of their complexity, therefore we can infer that, like machines, animal bodies must have been *designed* [Paley, 7-15]. Butler, in common with Darwinian fellow-travellers like T.H. Huxley, accepted Paley's premise, but drew opposite conclusions.

Analogy pervades the 'Book of the Machines'. We think, Butler says, that we can make a principled distinction between organic and mechanical modes of being. But where does this idea come from? Assuredly not from nature; isn't an egg-shell as much of a 'machine' as the nest containing it [199]? Again: it is both easy and enlightening to conceptualize human social arrangements as mechanical (inputs and outputs, differentiation of function etc) and then, by venerable analogy, the human body as likewise a kind of machine: our blood, for example, "composed of infinite living agents which go up and down the highways and byways of our bodies as peo-

↘ and defense of Strong AI, is Dennett (2003).

ple in the streets of a city" [206]. This is one of the oldest tropes going, and intuitively right in the web of correspondences it sustains, however much anti-materialists cavil [e.g. Sennett]. Again: had not the researches of Carnot, Joule, Thomson, Helmholtz etc revealed that all physical systems – human and animal bodies, machines, stars – are essentially heat engines, requiring stomachs and means of procuring nutriment to stave off what the Second Law of Thermodynamics teaches is inevitable – a (fatal) increase in disorder? The past century, Butler writes in 1872, has seen the evolution of machine stomachs (the chief difference between a steam engine and a spade); how long before they need not rely on human intermediaries (*vide* (4)) to feed themselves? At this point, could we not say that they are animate; and, being animate, are capable of consciousness (209-10)?

Again: much of what we cherish as distinctively 'human' – sleeping, weeping, yawning, loving our children, feeling pleasure, hating etc – Darwin's *Descent of Man* (1871) and *Expression of the Emotions in Man and Animals* (1872) had suggested we in fact share with animals. Why should we not in the future share pain, hope, shame, the fear of death, the desire to communicate our pains, hopes, fears and embarrassments to others – with the machines [208]?

8. An argument from reproduction

If it is said that to be conscious requires being animate, and that to be animate requires possession of a reproductive system – and that machines, lacking reproductive systems, therefore cannot be conscious – then, the Erewhonian philosopher replies, it all depends on how you define reproduction [210-211]. Clearly machines *do* make other machines – it's just that the infant machines are not copies of the parent machines. But, the philosopher points out, specialization of reproductive function is common in the animal kingdom (notably in eusocial species) The machine reproductive system is nowhere exactly paralleled in nature – the degree of specialization, i.e. the manufacture of components, is an order of magnitude greater – but that does not mean it is not a reproductive system. "The truth is that each part of every vapour-engine is bred by its own special breeders, whose function it is to breed that part, and that only, while the combination of parts into a whole forms another department of the mechanical reproductive system, which is at present exceedingly complex and difficult to see in its entirety" [212]. This argument recalls (1), above: the supposition that machine 'life' must exactly resemble organic 'life' for it to count as such is groundless.

As in (6.ii) a complex discourse of the relations between ‘intelligence’ and ‘system’ is being elided here. The philosopher does, apparently, treat of these matters, but it all goes over Higgs’ head. Sadly for us, and perhaps at the insistence of Butler’s publishers, the relevant passages are omitted as being “hopelessly obscure” [206].

9. (COUNTER-ARGUMENT) *An argument from prosthetics*

This is the Erewhonians’ “one serious attempt” to answer the anti-technology case. It points out that all the arguments described above suffer from the same conceptual confusion: that we have machines on the one hand and humanity on the other, and that the only pertinent questions are how far the former might come to resemble the latter, and what will be the consequences of the resemblance being carried too far. But this is to make a category mistake. Machines never have been and never will be ‘separate’ from humanity. It is meaningless to consider the possibility of them encroaching on ‘our’ territory: they *constitute* our territory.⁶

Instead, machines should be “regarded as a part of man’s own physical nature, being really nothing but extra-corporal limbs. Man ... [is] a machinate mammal” [p223], *viz.* in his use of spades (extensions of hands), eyeglasses (extensions of sight), pocket-books (extensions of memory) etc. From this the unnamed Erewhonian machine-apologist draws two surprising conclusions. The first is that machines establish a “community of limbs” where previously each individual was restricted to and jealous of the use of his/her own. In other words, technology has worked a transformative magic upon not just the territory of the body, but also the territory of the social (or the body socialized; or society incarnate). “Even community of limbs is thus rendered possible to those who have so much community of soul as to own money enough to pay a railway fare; for a train is only a seven-leagued foot that five hundred may own at once [224].

6 Re: the theories of Richard Gregory and Daniel Dennett on tools as both “endowers of intelligence” [Dennett (1997), 131], and of extensions of intelligence into, or projections onto, the environment [178]. Dennett is famously bullish on the possibility of Strong AI; but skeptical whether intelligence in *any* other animal species is fully ‘Gregorian’ (i.e. self-representing) and fully (that is, symbolically) prosthetic. He argues that language – a *complex* symbolic system – is essential to the possession of what we call consciousness. Animals do not possess language, and extend/project their intentionality prosthetically in only a very limited fashion. Therefore animals are not conscious, in the significant ways that we are conscious. Machines are themselves prosthetic; and in the case of computers are *made* of language, or are explicitly *designed* to be ‘representing systems’ (they are already ‘Popperian’; the next step is to make them ‘Gregorian’): therefore machines are, potentially, conscious.

Second, as the man-machine assemblage becomes ever more complex, the tyranny of matter ("that old philosophic enemy") will be overthrown. The man whose senses are attuned at once to the grandest and most minute of physical emanations; who, mastering space, has also, as Butler's contemporaries liked to put it, 'annihilated time': such a man can take leave of the base material world, and venture into realms of pure spirit. "[T]he elaborate organization of his extra-corporeal system has freed his soul" [225].⁷ Missionaries in the 1830s and 1840s had often referred to 'steam-engine evangelists' and the like; here is Butler's half-sardonic gloss on such pieties. This is not the place to discuss in detail *Erewhon's* theological and imperial dimensions; but recall Higgs' scheme for selling the Erewhonians as indentured labour to the owners of Australian sugar plantations. Two capabilities underpin this nasty revenge fantasy. First, the spatial reach of large techno-social systems – steamships, telegraphs, instruments of credit, rifles etc: all the machineries that enable people to be "packed close", transplanted en masse, and set to more productive labour. Second, the readiness to see in "organization", imperial or otherwise, a means of spiritual transcendence. Nothing allays the qualms one might feel at the prospect of packing people closely better than the conviction that it is being done for their own benefit. For this one needs God: Higgs' discovery, at the end of the book, is of the extraordinary instrumental efficacy of this hypothesis. The Erewhonians have largely dispensed with it. This is what has allowed them to smash up their machines; and smashing up the machines has freed them from a preoccupation with souls, their own or – more pertinently – anybody else's.

WORKS CITED

- Berg, Maxine (1980), *The machinery question and the making of political economy*. Cambridge: Cambridge University Press.
- Butler, Samuel (1872), *Erewhon*, ed. Peter Mudford. Harmondsworth: Penguin, 1970.
- Chalmers, David (1996), *The Conscious Mind: In Search of a Fundamental Theory*.

7 This passage, more than any other in *Erewhon*, anticipates H.G. Wells, a later and more successful purveyor of 'scientific romances'. Butler's ironic linkages between access to technology, social class, and conditions of body and soul, evoke Wells at his most mordant: "rich and subtle souls can defy all material impediment, whereas the souls of the poor are clogged and hampered by matter, which sticks fast about them as treacle to the wings of the fly, or as one struggling in a quicksand" [p225]. From here (fortunate souls travelling first-class; unfortunate trudging in ditches) to a world of Morlocks and Eloi is no great distance. Butler's image of superfine intellects sustained by prosthetics also invokes Wells' Martians, in all their cruel, parasitic, imperialist glory.

- Oxford: Oxford University Press.
- Dennett, Daniel (1995), *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. Harmondsworth: Penguin.
- Dennett, Daniel (1997), *Kinds of Minds: The Origins of Consciousness*. Phoenix: London.
- Dennett, Daniel (2003), *Freedom Evolves*. Harmondsworth: Penguin.
- Kostal, Rande (1994), *Law and English Railway Capitalism, 1824-1875*. Oxford: Clarendon.
- Paley, William (1802), *Natural Theology*, ed. Matthew Eddy & David Knight. Oxford: Oxford University Press, 2006.
- Paradis, James ed. (2007), *Samuel Butler, Victorian Against the Grain* Toronto: University of Toronto.
- Penrose, Roger (1989), *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press.
- Porter, T.M. (1981), 'A statistical survey of gases: Maxwell's social physics', *Historical Studies in the Physical Sciences* 12: 77-116.
- Searle, John (1997), *The Mystery of Consciousness*. Granta: London.
- Sennett, Richard (1994), *Flesh and Stone: The Body And The City In Western Civilization*. New York & London: Faber and Faber.
- Winch, Donald (2009), *Wealth and Life: Essays on the Intellectual History of Political Economy in Britain, 1848-1914*. Cambridge: Cambridge University Press.