

FINAL ASSIGNMENT

Papers are officially due on Dec 10th at 4PM. In cases of extension, no paper will be accepted after December 16th at 4 PM. Papers should consist of approximately 8 typed pages (figure 320 words/page). As before, the following questions and topics are meant to be suggestive. If you wish to modify them or invent a topic of your own, you may do so, but the object of the discussion should be one (or more) of the texts read and discussed so far this term and should deal centrally with one of the texts read since we undertook reading and discussion of Butler's *Erewhon*. Once again, please remember that you are writing an essay, not a book-report. We have read the book and do not require a rehearsal of its contents. What an essay does supply is some reminder of the contents in the context of an argument about those contents; the reminder is offered in the course of explaining how one should understand or interpret those contents.

Please try to avoid writing a sermon on any of the topics suggested below. I am always interested in your views on life, but your primary task here is to exhibit your ability to read and interpret the texts read this term.

Suggested Topics:

Satire has targets; it intends to show up the irrationality of certain practices. In addition to specific targets, every satire embodies a general view of what counts as irrational and what counts as rational, and for this reason, every satire has a general target--the source of much irrationality in human life. To identify the general target (in this sense) of a satire is not always a simple task. Compare the general targets of *Erewhon* and *Candide*.

Here is a brief extract from some pages of Darwin's *The Descent of Man* that we did not read:

Man in the rudest state in which he now exists is the most dominant animal that has ever appeared on this earth. He has spread more widely than any other highly organized form: and all others have yielded before him. He manifestly owes this immense superiority to his intellectual faculties, to his social habits, which lead him to aid and defend his fellows, and to his corporeal structure. The supreme importance of these characters has been proved by the final arbitrament of the battle for life.

How well does the notion of dominion invoked here accord with the view of selective advantage expressed in the *Origin*? How apt are the criteria for its use? How valid is Darwin's explanation of the causes of this "dominance"? Compare the notions in

the extract from Darwin's *Descent* with those of Butler on the domination of machines and of Huxley and Wells on "the final arbitrament [i.e., judgement] of the battle for life".

Butler presents two views of mankind's relation to machinery in "The Book of the Machines", one of which is used to plead for the destruction of all machinery in the kingdom of Erewhon, the other used to plead for its preservation. Are the views really opposed? How do they compare with views on machinery of any two other writers read this term?

Butler prefaced the second edition of *Erewhon* with a disclaimer to the effect that he was not poking fun at Darwin in "The Book of Machines". The real target, he said, was the poor use of analogy. We have met quarrels about analogy in Hume and Paley and it must be said that Darwin, despite Butler's disclaimer, relies heavily on analogy to make a case directly opposite to Paley's. Is Butler's disclaimer to be accepted at face value? How well has Butler captured the Darwinian use of analogy?

Erewhon is "nowhere" spelled backwards, with a slight concession to pronouncibility. To what extent is reversal the principle of satire in the book? What are some of its targets and how does the technique of reversal apply to Butler's treatment of them? Is the idea of reversal applied consistently? Does any general sort of target appear to be the main concern (perhaps only implicit, not necessarily explicit) of the book? Compare Butler and Voltaire as satirists. If the object of society is the establishment of a stable, predictable code upon which individuals may depend, what are the comparative merits of Voltaire's Eldorado and Butler's Erewhon. How does a world of rational behavior compare with one devoted to the following of Ygrund?

The judge in Butler's *Erewhon* speaks in condemnation of the man whose wife has died: "You have suffered a great loss. Nature attaches a severe penalty to such offenses, and human law must emphasize the decrees of nature." In the same way, Erewhonians regard luck, health, success (or their opposites, ill-fortune, disease, failure), as if they were ethical attributes, and criminality as non-ethical, something that can be pitied or ameliorated. What is the underlying distinction between the ethical and the non-ethical that is at work here? Is the Erewhonian system of justice consistent? If not, is that a fault in Butler's book?

Compare and/or contrast the relation between the ethical and the natural in any two texts text read this term. How consonant, for example, is the position taken by Huxley's call to resist nature with the position elaborated by Darwin at the end of *The Descent of Man*? We may recall the passage cited in class:

The advancement of the welfare of mankind is a most intricate problem: all ought to refrain from marriage who cannot avoid abject poverty for their children; for poverty is not only a great evil, but tends to its own increase by leading to recklessness in marriage. On the other hand, . . . if the prudent avoid marriage, whilst the reckless marry, the inferior members tend to supplant the better members of society. Man, like every other animal, has no doubt advanced to his present high condition through a struggle for existence consequent on his rapid multiplication; and if he is to advance still higher, it is to be feared that he must remain subject to a severe struggle. Otherwise he would sink into indolence, and the more gifted men would not be more successful in the battle of life than the less gifted. Hence our natural rate of increase, though leading to many and obvious evils, must not be

greatly diminished by any means. There should be open competition for all men; and the most able should not be prevented by laws or customs from succeeding best and rearing the largest number of offspring.

If artificial should replace natural selection among human beings, should we practice eugenics? Comment, using specific passages to illuminate the position that might be taken by Huxley on this issue. (And please refrain from sermonizing upon the issue: we are testing your competence in dealing with these texts, not your ethical fervor.)

The above passage linked "struggle for existence" with "advancement". Certain writers whom we have read do not think that standing still is possible in the life of a species--it is either "advancement" or "decline". Comment with specific reference to authors and texts. Butler's book of machines depends in part on the "willingness" of machines to suffer individual extinction for the good of the species of machines as a whole. Is this consonant with Darwin's view on species? Does this make machinery a danger to mankind?

What is the occasion for Huxley's argument? What view of the implications of Darwinism is he trying to discredit? Huxley opposes the State of Nature and the Human State of Art (that is to say, artifice, the intelligent management of things) and insists that the two are in conflict. What was Aristotle's view of this relationship? How would an intelligent Darwinist argue against Huxley's notion of a conflict? Huxley links the State of Nature with Cosmic Nature and argues that on earth human artifice is the product of cosmic nature. If it seems contradictory, then, to speak of art as resisting cosmic nature, Huxley says, that is just too bad for logic. Is the opposition logically contradictory? How does Huxley attempt to show that cosmic nature itself in the case of human beings transforms the struggle between members of a species into a cooperative effort to contain that struggle. Is his case convincing?

The emblem of artifice chosen by Huxley is the Garden--an image of nature, by implication, as thoroughly under control of human artifice as might be, "where every plant and every lower animal should be adapted to human wants and would perish if human supervision and protection were withdrawn". How does this garden compare with the garden in Genesis? How does its supervision compare with the idea of breeding domestic animals in the first chapter of Darwin's *Origin*? What is the serpent in the garden? Is the idea of transforming the environment of humanity good but impractical, in Huxley's view, or ultimately possible, or a false ideal, inherently wrongheaded? In answering, take into account Huxley's view that direct selection of the best humans would be undesirable not just because we lack the knowledge to select the fittest but also because it cannot "be practiced without a serious weakening, it may be the destruction, of the bonds which holds society together." Also take into account his view that the struggle for existence among human beings in industrial nations is just about over and the real struggle now is a struggle for conveniences and luxuries.

The story of Jekyll and Hyde has been taken sometimes as a warning against science's meddling with nature, at times as a moral tract warning against loosening the hold of the ethical self upon one's baser or more animal impulses. Discuss the validity and/or the limits of either of these views.

Discuss the adequacy of the account that Jekyll gives of his aim in developing his Hyde-creating drug. A good deal goes into it in very brief space--Jekyll speaks of

two elements of his soul and of the advantages of liberating one from the other. How coherent is this enterprise? Jekyll says that "the drug had no discriminating action" and simply liberated the element active when one was taking it, leaving the other unaffected. Does this mean, as Jekyll says, that taken in the right spirit, it would have made him a saint? If so, what becomes of the notion that the moral of the book is never to indulge the evil side of one's character?

Describe carefully the relationship between the two sides of Jekyll's double being, as Jekyll himself understands it. How does each think of the other? Jekyll says that during the last nine months he has labored exhaustively and even enjoyably at the good work of relieving the sufferings of others in order to make up for Hyde's actions. The story, however, will not let this remain a stable position. What unbalances it? Why is Hyde described as having more life in him than Jekyll? In this connection, examine with care the long paragraph (third from the end of the story) in which Jekyll details the final relationship between the two sides of his character. Jekyll's loathing of Hyde might seem complete, yet he describes Hyde in such a way as to lead to the conclusion: *I find it in my heart to pity him*. How adequate is Hyde as a representation of what Jekyll calls "my original evil"? How tempted is Jekyll to become all Hyde? How closely does he identify himself with Hyde at various moments during the course of Jekyll's narrative?

Both Wells (in *The Time Machine*) and Stevenson write stories in which two sides of human nature are conceived as potentially separable--but the division is not conceived in the same way. Mankind is a composite creature, but the elements of the composition differ. Comment.

Wells's *The Time Machine* seems to accept Huxley's viewpoint in *Evolution and Ethics*. Hardship and struggle, foster the human capacity for intelligence (recall Malthus, too). To put this in Huxley's terms, natural selection, in producing humans, has produced something that opposes the work of natural selection in favor of cooperation and the use of natural forces for the benefit of the human species. Paradoxically, however, intelligence, by removing the need for struggle, may remove the selective advantage of intelligence. Hence, the existence of Eloi and Morlocks, neither of which are very intelligent or cultured. Has intelligence thus led to the victory of the cosmic process over intelligence and culture, which seem to be missing in Wells's futuristic world?

A more general version of the same question: "All the ethical points made by Huxley are present in *The Time Machine* but translated into vision." This statement (by the critic J. P. Vernier) is a commonplace of observation about Wells's scientific romance. Comment by comparing selected aspects of the two books.

Wells's *The Island of Dr Moreau* seems to accept many elements of Huxley's viewpoint. Mankind has evolved from animal nature by dint of pain (Huxley would say, hardship and struggle); the force behind evolution (Huxley's "cosmic process", Wells's *Dr Moreau*) is without morality; the creatures thus produced have to create their own morality. Comment on the parallel and how far it can be drawn.

In various ways, the discourses of Huxley and Wells modify the idea of "perfection", a term used with some insistence by Darwin. As Huxley points out, although the process of natural selection may be said to perfect, it does so only in relationship to the particular situation of a given species--its particular conditions of life. Under

many imaginable conditions, therefore, "perfection" by the standards of natural selection can be "degeneration" by the standards of ethics. This notion is not only applicable to Wells's treatment of the Morlocks and the Eloi but also to the distinction between Moreau and his creatures. *The Island of Dr Moreau* raises the question whether the animals, in becoming more human-like, have been "improved", as opposed to merely "altered". And it also raises the question whether Moreau represents a further evolution of mankind, for Moreau seems himself to be like the being that he hopes to create one day—a rational creature free from the "taint of the beast", no longer in bondage to fear of pain and therefore a creature still an animal, "thinking a little less obscurely what an animal feels." Compare and contrast the division implied here between the creatures and the ultimate human with the distinction between Morlock and Eloi or the distinction between Jekyll and Hyde.

In a sense, Wells's *The Island of Dr Moreau* gives us a double dose of pessimism. In the first place, Dr. Moreau expresses a conviction that "the study of Nature makes a man at last as remorseless as Nature." Is Darwinian Nature remorseless? Is Huxley's cosmic nature remorseless? In the second place, if Moreau acts in the manner of Nature, what is the remorselessness of Nature remorseless about? Moreau is devoted to investigate the limits of "the plasticity of living forms", but he seems to have no interest whatsoever in ameliorating the conditions of human life thereby. Is the investigation a worthy goal in itself? What oppresses Prendick about the island at the end of the book—when he suffers, so to speak, from a spiritual hangover—is the aimlessness of things there. Is this aimlessness in line with a Darwinian view of life on planet Earth?

Take the opening episode of *The Island of Dr Moreau*, the account of the struggle for life in the boat, and show how it relates to the central themes of the story as a whole. Or take the last chapter and compare it with the last paragraph of Darwin's *Origin* in order to estimate how closely Wells was influenced by the Darwinian view of things, if not by the *Origin* itself. Darwin contrasts the eternal regularity of cosmic nature with the complex changefulness of biological history, whereas Prendick, although educated by Huxley, seeks solace in the study of astronomy.

Wells claimed to be a disciple of Huxley. He has Prendick doing research with Huxley "as a relief from the dullness" of his leisured existence. In a book one of whose central terms is "aimlessness", this is a significant qualification. He also has Prendick refer (in a long passage concluding Chapter 16) to the aimlessness of things on the island and to the fact that Moreau has wrenched the creatures on the island from their condition as beasts (this is Prendick's usual word), in which their instincts were "fitly adapted to their surroundings" and made into something unfitted for any natural existence. What are the implications of "bestiality" and "mark of the beast"? Why not refer just to "animal nature"? Is this idea of fitness in accord with Huxley's? With Darwin? By the same token: Are the beasts on the island, in their lack of fitness of adaptation, apt models for human beings? In what way? Prendick regards with horror the notion that humans would be turned into beasts, but is somewhat pacified to begin with by the notion that beasts are being turned into humans. Does the course of the story alter this view? What does the story say as a whole about the relation of humanity to nature (in any relevant sense of the word "nature")?

Moreau claims to be a religious man, "as every sane man must be". What does he mean by this? At one point, Moreau does seem to have more in mind than just exploring "the plasticity of living forms". While complaining that somehow he cannot get at the seat of emotions, which, in his view, "harm humanity", he says that he wants to create a creature of his own which is truly rational. Does Moreau himself possess the sorts of emotions that he is complaining about? Is he striving to

create a rational creature “in his own image”? Would this make the creature more or less human? Alternatively, is the liability to emotion a “mark of the beast”, as Moreau claims?

Forster's *The Machine Stops* was written partially in response to Wells's *Time Machine*. Compare the two texts in order to highlight Forster's reworking of Wells's materials and show how this reworking serves Forster's purposes. How does Forster's text capture Huxley's idea of a world entirely subdued to artifice? What are the significant differences? How extensively do Forster's views about the effects that an end to the “struggle for existence” will have on mankind parallel Wells's? How do they differ? Does Forster's book match up with the fears expressed by the author of the “Book of the Machine” in *Erewhon*? What view would the second author in *Erewhon*—the one who wrote in answer to the author of “The Book of the Machine”—be likely to take of the implications of Forster's story?

What are “ideas” in *The Machine Stops*? Why does Vashti “get no ideas” from those parts of the natural world that mankind naturally inhabits? What is meant by “the terrors of direct experience”? If the inhabitants of the Machine have indeed lost all sense of “direct experience” (whatever that is), has their devotion to “ideas” played any role in this? Are there any correlatives in our contemporary world to the “ideas” that are so central to human life in Forster's story? What does Kuno mean when he says that mankind is the measure of all things? Kuno says that people “worship” the Machine. Can a Machine be worshiped?

Each year, writes Forster, the machine is served more efficiently and less intelligently. Is this possible? Can people be part of a machine (or a machine-like process)? Would you regard the workers on an assembly-line as *servicing* the conveyor belt, which embodies an idea of consecutive assembly that the workers do not have to grasp to perform their function? Can one regard the factory itself as a machine in which the workers are functioning parts?

What does it mean to live inside a machine? Is the machinery simply a metaphor, a stand-in for features of human relationships that Forster judged adversely? Or is the story actually about machinery and what happens to human relationships when human relationships become dependent upon machines, particularly in the way of communication? Or can these questions both be answered affirmatively—by raising the possibility that the story is about something in human nature (of which Forster does not approve) which does not depend upon the existence of machinery but which is encouraged by an increasing reliance of human beings to manage their relationships with one another by means of machines?

The question, can computers really think? may be regarded as another version of the question, does nature really select? As nature exhibits design, so behavior (linguistic or otherwise) exhibits consciousness. How far does the parallel go? Discuss in relation to the packet of readings on Artificial Intelligence, selfish genes, or any other text read this term.

How well-suited is the Imitation Game to the purpose for which Turing intends it? Are the conditions of the game adequate to serve as a *decisive* example that machines can think? Suppose that we set up a rock so that it dropped on the foot of anyone who shoved his or her foot through a hole in a wall. Would we be entitled to say that the rock has successfully “imitated” a malicious human being behind the hole and we can confidently expect the rock to take on more human qualities in fifty years time?

In a passage not included in your reading, Turing seems to nod in the direction of such an argument when he writes: "The reader will have anticipated that I have no very convincing arguments of a positive nature to support my views. If I had I should not have taken such pains to point out the fallacies in contrary views." How would you characterize the contrary views? Which of them do you think is the most formidable? How successfully has Turing dealt with it?

A similar set of questions apply to Searle's case against what he calls "strong AI". How would you characterize the AI views? Which of them do you think is the most formidable? How successfully does Searle deal with it? Unlike Turing's adversaries, Searle does not deny that machines can think; he denies only that machines that must be programmed to simulate thinking can think. Why is the idea of being programmed to think, in his view, disqualifying? How much difference does this make to the force of his objections? "In strong AI," he writes, "what matters are programs, and programs are independent of their realization in machines." He hints that this is a version, on the part of proponents of "strong AI", of the view that soul is distinct from body or mind from brain. What is your view of this issue?

How far does the "Chinese Room" example duplicate the situation of those inside Forster's machine? We may recall the reply given to "The Book of the Machines" by the second author in *Erewhon* to the effect that machines are simply extensions of human beings, who may one day (like Wells's Martians in *The War of the Worlds*) spend their lifetimes inside ever-more powerful machines and manipulate the world at their pleasure. How close does this come to the "Systems Reply" in Searle's essay, in which the whole system, not just the little man inside, understands Chinese?

Like Paley, the writer of Butler's "Book of Machines" compares the eye to the microscope and the telescope (the little and big "see-engines"). He writes: "What is a man's eye but a machine for the little creature that sits behind in his brain to look through?" The word "machine" here should not disturb us; it means "contrivance". But what of the "little creature"? If there is a little creature that uses the eye to look through, what does the little creature see with? Starting from this question, compare the various writings on computer intelligence distributed this term. Does a machine or contrivance always serve a purpose? If human beings are composed of machines or contrivances, and if the human brain is also a contrivance (as Butler hints in the paragraph after the one just cited), what is the entity for which it performs a function or a service?

Compare Searle's views with those of Daniel Dennett in "Intentional systems". Searle writes: "We say, 'The door knows when to open because of its photoelectric cell,' 'The adding machine knows how (understands how, is able) to do addition and subtraction but not division,' and 'The thermostat perceives changes in the temperature.' The reason we make these attributions is quite interesting, and it has to do with the fact that in artifacts we extend our own intentionality; our tools are extensions of our purposes, and so we find it natural to make metaphorical attributions of intentionality to them; but I take it no philosophical ice is cut by such examples." Compare this passage with Dennett, who writes: "Any time a theory builder proposes to call any event, state, structure, etc., in any system (say the brain of an organism) a signal or message or command or otherwise endows it with content, he takes out a loan of intelligence. . . . The intentionality of all such talk of signals and commands reminds us that rationality is being taken for granted, and in this way shows us where a theory is incomplete. . . . [W]herever a theory relies on a formulation bearing the logical marks of intentionality, there a little man [within the brain] is concealed." In the light of this remark—which urges us to find explanation for intention in terms of design and explanation of design in

terms of natural selection—where does Dennett stand on the quarrel between Searle and the proponents of what Searle calls “strong AI”? Would Dennett agree or disagree that the use of intentional idioms should “cut no ice” in understanding whether or not machines actually think?

“A hen is but an egg’s way of making another egg,” said Samuel Butler (but *not* in *Erewhon*). Dawkins’s arguments in *The Selfish Gene* and elsewhere seem to elaborate this remark, the spirit of which is already present in “The Book of the Machines”, when its author argues that both humans and machines, considered as species, are involved in a common life-cycle in which humans foster the existence of machines, which in turn fosters human existence, which in turn . . . etc. Butler’s author concludes that the cycle exists for the purpose of producing machines, not the other way around, and it is part of Butler’s satire to make his Erewhonians take this idea seriously. Dawkins takes the opposite tack—the machines are not the beneficiaries of the cycle—but he also argues that we are the machines. ‘We are . . . robot vehicles blindly programmed to preserve the selfish molecules known as genes. This is a truth which still fills me with astonishment’ What are the grounds for Dawkins’s argument that we serve the interests of genes? Comment on the relation between Butler’s author and Dawkins in any way that seems to illuminate the argument of one or the other.

Various aspects of Dawkins’s arguments seem to recall the arguments by Searle and Dennett about the “intentionality” of such things as machines, only here, in the phrase “the selfish gene”, the intentionality belongs to viral-like lengths of chromosome. Compare Dawkins’s argument about the “behaviorist” definition of selfishness with the quarrel about ascribing intentionality to machines in any way that seems useful. How does the issue here bear upon Darwin’s remarks about the confusions prompted by his use of the term “Natural Selection” and his insistence that “selection” is just a metaphor?

Stephen Gould is critical alike of gene-selection and of group-selection as theories describing the targets of Natural Selection. How would his criticisms affect Huxley’s argument in *Evolution and Ethics*? What might Dawkins say in answers to his criticisms?