

Lecture 18: Dualism and the Mind/Body Problem

What is the relationship between mental and physical?

If mind = Brain

Does desire, hopes, feelings, thoughts, beliefs = some brain state (chemical state or chemical reaction)

Descartes's Dualism:

The mind and brain are distinct entities, are two kinds of entities (mental and physical)

But there seems to be causal interactions between them.

Body -> Mind: Aspirin helps a headache, alcohol impairs reasoning.

Mind -> Body: Deliberation precedes action, sadness causes tears, finding something funny causes laughter.

Alternative: Mind/Brain Identity theory:

The mind is the brain.

Examples of identities that we discovered over time:

Water = H₂O

Heat = Mean Kinetic Energy

Influenza = effects of a virus on the system

Mind = Brain is just one that we are coming to understand now (even though we do not fully understand it)

Mind/Body Identity and the Immortality of the Soul.

If the mind is the soul, and the soul is immortal, then the mind is immortal.

If the body is not immortal, but the soul is, then the mind cannot be the soul.

This seems to present an argument for Dualism, but notice that there might be some who doubt that the soul is immortal, and so are not convinced by the argument. It is a quasi-circular argument. (No one doubtful of dualism is likely to believe the soul is immortal).

Identity and Leibniz's Law

If two things are identical, then they have all the same properties.

Or negatively, if one thing has a property another thing does not, then they are not identical, they are distinct.

This is known as the indiscernability of identicals.

The President of the U.S.A. is 5' 9", Bill Clinton is 6'

Therefore, Bill Clinton is not the President of the U.S.A.

Descartes First Argument

I can't doubt my own mind exists

I can doubt my body exists.

Therefore they are not the same

This argument seems valid, but there is something troubling about it.

Descartes seems to claim that being dubitable or indubitable is a property of some thing (or category of things). Then by Leibniz's law (or something like it) show that the two things must be different. But we should be careful about doing that.

Consider the Superman example.

Lois Lane wants to marry Superman.

Lois Lane does not to marry Clark Kent.

Superman is not Clark Kent

(This seems to be of the same form, and it is clearly not a good argument)

If the conclusion is false either a premise is false or the argument is not valid.

I think either might be true.

It certainly seems plausible to say that the conclusion could be false if Lois Lane did not realize that Superman was Clark Kent (if she did it gets harder to explain).

But one might also suggest that though she might say that she does not want to marry Clark Kent, in fact she does, though she does not realize it. I think that is also possible.

Propositional Attitudes.

One reason such examples are difficult is that they involve *propositional attitudes* (PAs) *Propositional Attitudes* are stances or feelings we can have towards some proposition.

They include an attitude (or feeling about) and some proposition (or statement). For example: I want that (attitude) I drink some coffee (proposition).

Thus, I want coffee is really a propositional attitude, in the PA jargon it is translated more precisely as I want that I drink some coffee. That makes it more specific than simply wanting coffee on my pants, to exist in the world, beans in my pocket etc..

By this reasoning if the propositional attitudes are different, then the propositions must be different (if they are both to be true), but it does not mean that the propositions are about different things.

So "Lois Lane wants to marry Superman" is a different proposition from "Lois Lane does not want to marry Clark Kent." But they both are about the same person (Kal El).

Frege Sense and Reference

Frege notes that sentences can be understood in two ways, by the senses (or meanings) or the terms, and by the reference (the things they are about).

Ex. The Evening Star is the Evening Star

The Evening Star is the Morning Star

In terms of sense (meaning) the sentences are quite different. One expresses an a priori claim about self-identity. The other, make a contingent claim that the first star visible at night is the same as the last star visible in the morning. (It is actually the planet Venus).

In terms of reference (the thing talked about) they both refer to the same thing (the planet Venus).

Two names (or descriptions) which refer to the same object are said to be coreferential. To terms which have the same meaning are synonymous.

Synonymous Terms

Ibn Rushd & Averroes

Cougars & Mountain Lions

Coreferential Terms

George W. Bush & 43rd US President

Inventor of the YoYo & Inventor of Parking Meter

Frege thought that sentences were either true or false depending on whether what was said was true of the referents (not the senses). But this poses a problem for the Superman case. According to Frege sentences expressing certain propositional attitudes must really

be referring to the sense. So Lois Lane does not want to marry "the person she thinks Clark Kent is, not the person he really is." Obviously, this account is troubling since we do not always know to whom we are referring. Still it offers an alternative to the idea that we cannot apply logic to propositional attitudes.

Descartes's Second Argument:

Bodies are extended and divisible

Minds are not extended and indivisible.

Minds are not bodies

This is a valid argument. However it may be begging the questions. How do we know that minds are not divisible? If we have no independent reason for believing that, then it seems to be that are either defining minds in such a way that they cannot be bodies, or simply saying that they are not. But is there any independent reason for believing minds to be indivisible? (What about multiple personalities?)

Causality between Physical and Non-Physical:

One of the concerns with Dualism is how exactly it is supposed to work. It seems plausible (essential) to suggest that mental stuff causes physical stuff to happen, and physical causes mental stuff.

Ex. I desire a drink, so I walk to the fountain.

(mental causes physical)

I drink some beers, so I think that I am a really good dancer.

(physical causes mental)

But how that is supposed to happen is unclear. If mental things have no extension, no physical form, that is if they take up no space and non-solid and have no mass what force can they exert on physical things (say our bodies.)

Review Questions:

1. Dualism asserts that there are two kinds of things: mental and physical; and that the mind is not the same as the brain.
Mind/Brain Identity asserts that the mind just is the brain and nothing more. That all mental events/things are really physical things of some kind. (e.g. Brain States)
2. If two things are identical, then they share all the same properties.
The mind seems to have different properties from the brain, it cannot be doubted, it is non-spatial, so they cannot be identical.
3. No. Our ability to doubt something is not a property of it, but a feature of us.
4. It is valid. But the conclusion may be false, if the premise that the Mind is not divisible or another is also false.
5. If mental stuff has no physical component, how is it supposed to make the physical body do anything? This assumes of course that the physical world responds to physical causality. At the very least the answer is unknown or puzzling.

Lecture 19: Logical Behaviorism

This is thesis more about the meaning of certain mental terms than about the metaphysics. Behaviorism asserts that mentalistic terms: belief, desire, thought, etc.. refer to behaviors we see in others and ourselves rather than some internal mental events.

It has a negative and positive claim:

Negative Claim: There is no Ghost in the Machine

Common-sense promotes two theses that, according the Behaviorists are false.

1. Mentalism: Mental states are the inner causes of behaviors (wants, beliefs etc.. cause our physical actions).
2. Dualism: There is a fundamental difference between physical and mental stuff.
(You can be a mentalist without being a dualist)

Mentalism leads to skepticism.

- 1 If mental states are inner causes of behavior, we would not have knowledge of the mental states of others.
- 2 We do have knowledge of the mental states of others
Therefore mental states are not the inner causes of behavior.

This is a valid argument. So there must be something wrong with a premise if you want to deny the conclusion. Premise 2 seems true, premise 1 is a good candidate to check.

Maybe we can know the mental states of other despite the fact that they are inner states. We can do this by an argument of analogy.

I know my own inner states via introspection

When I am in state M, I produce behavior B

I observe another person is producing behavior B

I know that the other person is in mental state M.

This argument is fairly weak, since it infers quite a specific thing from only one example.

An abductive argument might be stronger.

I observe some behavior B in a variety of people. I produce a theory to explain the behavior. It includes saying that behavior B is caused by mental state M. (This will be more complicated in reality, but I am simplifying it.) If M can explain and B and makes a variety of other predictions, then if it is successful, we may be able to say that it is a good theory for explaining behavior by attributing certain mental states.

The Positive Claim: Mentalistic vocabulary describes kinds of behavior.

If someone "wants" something, really that means that they will try to get the thing. That is a person who wants something is "disposed to get it." This is called a dispositional account (it does not rely on success, but rather the likelihood of trying).

Sober has two objections to this account.

1. The dispositional analysis of desire is incomplete.

A single mental state cannot give a disposition to any behavior, but rather batches of them together (beliefs and desires) are needed to lead to even a disposition to behave.

2. A dispositional analysis doesn't refute mentalism.

The fact that something can be described as a disposition does not mean that there is no inner state that is responsible for that disposition. In fact we frequently expect there to be some hidden properties which explain all behaviors of all things, (even if they are not mental things).

Pain without Pain Behavior

A person, you for instance, can be paralyzed by ingesting curare (a drug). This will prevent all behavior associated with pain (screams etc..). We could even give the person, you, an amesiac drug which will make him, you, forget the pain. Would you volunteer to undergo an operation under those conditions? If not, it is probably because you would think that you would feel pain even though you could not express it.

Review Questions:

1. If we have to experience inner states to know of our own mental states, then we could never know about other people's mental states (since we cannot experience their inner states). But we seem to know when someone is happy or sad, or wanting to get inside, etc..
The argument from analogy opens the door to a solution but is weak since it is only one case.
2. The claims that (a) mentalism is false, and (b) mentalistic terms refer to dispositions to behave.
Jones believes that there is rat poison in the gravy. Jones will not eat any gravy.
3. No. Even if you can describe something as a disposition it may be that some hidden inner state is responsible for that behavior. (like the solubility of sugar which is caused by the chemical structure of sugar).

Lecture 21: Mind/Brain Identity Theory

This theory gets its motivation from the advances of science and argues that there are no things that are not physical things. Thus all mental things must be physical things. Things which seem different from physical things are just less obvious. Things we once thought of as unique kinds of things turn out to be some combination of physical phenomena.

Ex. Lightning is one and the same thing as a kind of electrical discharge.

Dualism resembles the old idea of Vitalism (which claims that all living things have some *elan vital*, a substance which makes them alive).

Just as biology discredited vitalism, we will expect science to discredit dualism.

Ex. Feeling Pain is one and the same event as having the c-fibers in one's brain fire.

Feeling pain is identical with some brain state. All mental states just are some brain states. There is some evidence for this. There is a strong correlation between mental phenomena and brain activity. Although neuroscience has yet to show absolute correlation, it does seem to be true that when there is no brain activity there is no thought.

But there could be correlation without the two being identical. Perhaps mental states cause the brain to fire that way.

Parsimony: This principle suggests that we accept as true the simplest of two hypotheses (all things being otherwise equal). On this ground the identity theory seems more palatable since it is clearly more simple (postulating only one kind of thing in the world). Science relies on parsimony for numerous advances, why should we when it comes to mental phenomena?

Review Questions

1. No. It is the view that all things in the universe are physical things and that there are no non-physical things.
2. The belief that a living thing contains *elan vital* (a substance that gives them life). No. Rather DNA shows us that there is no special substance necessary for life, just regular chemicals organized in a certain way.
3. They both posit a special kind of substance to explain a certain difficult-to-explain phenomenon. The problem of life is what is it that makes some things animated and other things inanimate. The problem of the mind is what our thoughts and other mental states are made out of.
4. The simplest answer is usually the best. It requires adding another type of substance (mental stuff).
5. No. It says that since they are identical in predictions we should prefer the simpler theory.

Lecture 22: Functionalism

Type/Token Distinction

Types: Properties or Kind (Blue, Unicorns, Feeling Pain)

Tokens: Instances, Unique objects (This shirt, no thing, yesterday after stubbing my toe)

Mind/Body Application

Token Identity: Every one psychological token is identical to some physical token.

"This pain in my toe" = "These neurons firing"

Type Identity: Each psychological type is identical to some particular physical type.

"Pains" = "Neurons Firing"

Functionalism rejects type-identity. Psychological types can be realizable in a number (perhaps infinite number) of different physical ways. Thinking or having psychological states is a matter not of the specific physical stuff, but rather being able to do certain things (the functions).

Could computers think? Computers and calculators seem to facilitate our own calculations not perform their own. Beware of mentalistic terminology (knowing, remembering, etc..) If *Functionalism* is correct, then there is no reason why computers could not have all the same mental states and abilities that we do.

The Chauvinism of Identity Theory.

Why should something need to resemble or be like our brain in order to handle the same functions?

Note that both *Identity Theory* and *Functionalism* are *Materialist* theories. *Functionalism* is consistent with *Token Identity*., That is my current thought that I have a computer in front of me may very well be identical with some physical state (some brain state), but it needn't be (an identical thought could be instantiated in some other way).

The negative claim of *Functionalism* is that mental types are not identical with physical types. Though particular tokens (instances of them) may be identical with particular physical tokens.

The positive claim of *Functionalism* is that mental states are "functional states." That is, mental events are describable as causal roles between behaviors and other mental states. A "desire" is anything which would cause certain beliefs to lead to certain behaviors. My "desire for a Samosa™" (girl-scout cookie) is when I am in a state that when I have the belief that "Samosas™ are available" I will behave in such a way as to obtain and devour them. If I do not "desire a Samosa™", then my belief that "a Samosa™ is available" will not cause me to behave so as to obtain and devour one.

Of course, not all desires do cause behaviors, nor do all behaviors result in the satisfaction of the desire. But there does seem to be some sense to say that mental states are causally related to behavior.

The problem of *Qualia*:

A functionalist account of pain will include reference not only to pain behaviors (as would *Logical Behaviorism*), but also to *psychological* states. Pains are states that cause among other things "a desire" for the pain to stop. *Functionalism* thus is more complete than *Behaviorism*. However, *Functionalism* still does not seem to account for the sensation of pain. The specific sensations, or subjective experience of something is

sometimes called its 'qualia'. According to Functionalism, a pain simply is some causal state. But the fact that a certain pain has a certain feel to it is not easily explained by the theory. Of course, it is not incompatible with it. But if Functionalism claims to say what pain is, but doesn't include anything about the feel of pain, perhaps it is not a complete theory.

Review Questions:

1. In both cases he is referring to types of things, classes or kinds of things rather than tokens. (At least I think so). If he meant the tokens, then we should expect the family to be short a couch and the son to be not very interested in the meal of regurgitated food.
2. There can be more than one way to instantiate the phenomena. It may be possible to be in the same functional state but be in some other physical (or non-physical) form.
3. They claim that Mind/Brain Identity limits the sorts of things that could have psychological states to the sort that we are most familiar with, that is things with brains. But there is no reason to limit it that way since most functional states can be realised in multiple ways.
4. A functionalist account can say that a psychological state is a functional state, and not just some particular token of a functional state. Thus what makes something a psychological state is not that it happen in the brain, but that it does something.
5. The state is related to other states such that it is caused by some collection of other states, and causes others. It can be described by those relations.

Lecture 23: Freedom, Determinism, and Causality

Most everyone agrees that our minds (however you describe them) cause things to happen in the world. And the world, be it our upbringing, our genes, causes things to happen in our minds.

In fact we tend to explain the behavior and facts about ourselves with reference to our genes and environmental factors.

The Puzzle of Freedom:

The environment you grow up in and the genes you start with are not things within your control. If all of your behaviors result from those things that you did not choose, how can you say that you are free to choose now?

If we are caused by things outside of us to do what we do, how can we be free?

If computers don't think or have minds because they are unfree to choose what to do, how are we any different? If computers are not free, how can we be free?

Distinct Causation Argument:

One is free if one is the cause of one's own behavior.

Our actions are caused by our beliefs and desires.

Our beliefs and desires are caused by our environment and our genes.

We do not cause our own environment or our genes.

We do not cause our own behavior

We are not free

Could have Done Otherwise Argument

One acts freely if one could have done otherwise

Given any fixed current state (beliefs and desires) one cannot do otherwise

One cannot act freely.

Examples of Unfree Acts:

A nervous tic. (this is something a person cannot avoid doing)

Brainwashed behavior

Compulsive behaviors (kleptomania)

These certainly seem unfree even if there is a species of free actions.

Are all behaviors of the same sort as brainwashing and compulsions?

If there are free actions there must be actions which differ from the above examples.

What is the difference between brainwashing and education?

What is the difference between kleptomania and wanton stealing?

Clash of Conceptions

One reason this is such a difficult problem is that it seems that two very strongly held beliefs contradict each other.

We are part of the causal network. (The world affects us and our decisions)

We are free to choose how to act.

Causality.

Remember Hume's skepticism about causality. How do we know that one event is the cause of another. Perhaps we don't know it. But we certainly believe it.

Some features of causality.

What causes something one time may not cause it in different circumstances.

A cause does not have to be a sufficient condition for its effect. (301)

Ex. Turning the key causes the car engine to start. But not if there is no gas in the tank. Or if the battery is dead.

Ex. Drinking 10 ounces of whisky will get one drunk, but not if one is very large, or an alcoholic.

There may be many possible causes for the same effect.

Causes often aren't necessary conditions for their effects. (301)

Ex. You can start a car by "hot-wiring" the ignition.

Ex. You can get drunk by drinking beer, or produce a similar effect by smoking marijuana.

We can talk about the "whole cause" of some effect which will include the circumstances, the proximal cause and any other relevant features.

Determinism.

This is the claim that if we have all the information about the whole cause, we can know or predict with certainty the consequences or effects. In other words, if we know the present we can know the future. More basically, that the present state of affairs determines or causes a particular future state of affairs.

"Determinism is the thesis that the facts at one time uniquely determine what comes next."(305)

Probability:

Die rolls, and other random things seem to be indeterminate. Of course, some people would argue that there really are not indeterminate things, just things which we cannot know well because of our limitations. (We do not understand the physics of a rolling die well enough).

Indeterminism:

Even if some things are random, that hardly gives us freedom. Consider the following option. We are not determined by our genes and environment, because certain random factors in our brains or mind can tip the scales in our actions. How random action becomes "free action" is unclear. (Imagine a computer with a random element.)

Causality is the issue not determinism.

Even indeterminate causes such as random ones seem to rob us of freedom. It would seem that in order to be free we would need to be uncaused, or not caused from the outside, or

something like that. For the sake of simplicity we sometimes conflate the causal and determinist claims, but the causal appears all that is necessary.

Determinism and behavior.

"Given your genes and environment, you couldn't have had a set of beliefs and desires different from the set you in fact possess. And given that set of beliefs and desires, you couldn't have performed an action different from the action you in fact produced." (305)

Fatalism:

This is a variant of determinism. It states that all or some of the facts of the universe, past present and future, are set in stone. (Note that this follows from three things. (1) A set starting point, (2) set rules of causation (no randomness), (3) determinism.) Fatalism frequently suggests that no matter what we want we will end up in the situation the fates determine (c.f. Oedipus). Determinism says only that the future will be determined by the present state whatever they may be. Thus a fatalist about Oedipus will say that no matter what he does, he will end up killing his father and marrying his mother. A determinist will say that if Oedipus has the correct beliefs and desires, he need not kill his father or marry his mother. According to determinism Oedipus could challenge the fates by doing something different than he did.

Another example of the difference:

I became a philosopher rather than a mathematician.

It was fate, no matter what I tried to do, I would end up become a philosopher.

Even if I thought I was becoming a mathematician I would in fact end up a philosopher.

It was determined. I became a philosopher because of various desires I had. I could have had other desires if I were raised differently went to another school, etc. If philosophy had more requirements I may not have become a philosopher. But given the circumstances I was determined.

Review Questions:

1. See Above. They both are very general claiming that none of our actions are free.
2. Brainwashing, Kleptomania
3. Determinism: All events are uniquely determined by their causes.
Indeterminism: Not all events are uniquely determined by their causes (there might be randomness)
They both make claims about what we could know if we had complete descriptions of some event.
4. Adding a roulette wheel to one's brain. It adds indeterminacy to one's deterministic brain. But it does not seem make the individual any more free.
5. That no matter what the circumstances (beliefs and desires) one is doomed to do behave in some way.

Lecture 24: A Menu of Positions on Free Will

Compatibility.

Two claims are compatible if they are consistent, in other word if they can both be true.

Note two propositions can be compatible even if one or both are false. Like validity, compatibility is about the relation of the propositions.

Incompatibilism about free will claims that free will and determinism are not compatible.

In other words: *If determinism is true, then we are not free.*

Incompatibilist Views

Hard Determinism (there is no freedom)

Libertarianism (determinism is false)

Compatibilist Views

Soft Determinism

(they are both true)

Libtertarianism:

This view argues that determinism and free will are incompatible, and that we know that we have freedom. Evidence for freedom is mostly introspective. There are two ways that we can do this. (1) We can act out of character (so our characters do not determine our actions). (2) We can sense our own internal struggle (c.f. Campbell readings.)

Sober Objects to (1) in that it seems more likely that the so called "out of character" actions are really explainable in terms of some characteristic not focused on. All that acting out of character means is that eithe the person is in an unusual circumstance, or that we did not know the character as well as we thought.

Soft Determinist Theories:

Hume's

You act freely if you could acted differently if you had different wants or beliefs.

Free actions are those which are under the causal control of the agent's beliefs and desires.

Unfree actions for Hume are actions you must take regardless of your wants.

1st Objection: Compulsive Behavior

The Kleptomaniac behaes in such a way according to his beliefs or desires. That is the compusion takes the form of the beliefs and desires. We are inclined to think of the Kleptomaniac as unfree since he would not will that he steals and cannot help himself. According to Hume a Kleptomaniac is free despite not being in control of those unwanted desires.

2nd Objections: Locke's Locked Room

Hume requires that you could have acted differently if your desires were different for an act to be free. But we can imagine a free action where that does not apply. I might choose to enter a room and stay there reading for four hours. This would be described as free by most people, including Hume. But if unbeknownst to be someone locked me in the room, according to Hume I would not be free since I would not be able to act differently.

Does Coercion rob us of free will?

Sober makes a good argument that coercion, force or the threat of force, does not rob us of free will. We are able to choose between two choices, death or giving up something. Granted, coercion might limit us. But it does not rob us of free will.

Second Compatibilist Proposal: Second Order Desires

One way of resolving the free will debate is to argue that a person is free if they have the desires and beliefs they approve of; or put negatively a person is not free if they have desires and or beliefs that they do not approve of.

First Order Desires: Wants for things in the world (states or objects)

Ex. I want that Porsche!

Second Order Desires: Wants for our own wants to change.

Ex. I want to want to exercise.

This solution takes care of some unfree compulsions. But it does not answer the question about a compulsive who approves of and wants to be compulsive.

Review Questions:

1. Y and X can both be true at the same time.
X and Y cannot both be true at the same time.
X and Y stand for propositions.
2. Compatibilism: We can be both free willed and determined.
Incompatibilism: If we are determined, then we are not free.
Hard determinism: We are determined, and incompatibilism is true, therefore we are not free.
Libertarianism: We are free, incompatibilism is true, therefore we are not determined.
Soft Determinism: We are free and we are determined, therefore incompatibilism is false.
3. We sometimes act out of character. Not strong.
Such actions can merely be evidence that we did not fully know the character of the agent.
4. We can assent to or reject any proposition and no force compels us to do it.
This seems plausible, though, it is inductive. There might be some propositions we cannot deny (I exist).
5. We are free if we would be able to act differently had we different desires.
Our desires can be determined by outside forces.
We would act differently if we had other desires, but we cannot give ourselves those desires which seem outside and alien to us.
We can't do otherwise, but we are free nonetheless.
6. No. As evidenced by the Locked Room
7. His definition is not sufficient, because even though the compulsive meets his definition he is not free. And his definition is not necessary, because the person locked in the room does not meet his definition but he is free.

Lecture 25: Compatibilism

The Weather Vane Analogy:

A stuck weather vane, is unfree. It points in one direction as opposed to what it was meant to do, point in the direction of the wind. That does not mean that the weather vane is free only if it is uncaused. In fact it must be caused by something (the wind).

An unfree will, or person, is likewise "stuck". They do not operate they are supposed.

Desire Generating Device[DGD] = The Will.

This is a way of describing a person such that if their DGD is operating properly, then they are free. That does not mean that their will is itself uncaused. Rather it means that it is caused in the right way, by the right things. Unfree people's will are in a sense stuck, because they operate in a way other than they are intended.

Functions:

Having a function is not the same as having a property (although functions are properties of sorts). A heart has mass, it makes sounds etc.. but its function is to pump blood, not simply take up space, make a thumping sound or what have you.

On a *naturalistic* account, things have the function that give them some advantage for survival or reproduction (review evolution Lecture 6).

Note: Just because something has a function that does mean that the function is the only valuable or useful thing about it. The lips serve the function of making the mouth air and water tight. But they are also nice for kissing.

What is the Function of the Will (DGD)?

Sober suggests that the will helps us do things which are (or were historically) advantageous to our survival (or reproduction). This seems to make sense since desires help us act to get food, find mates, etc..

A stuck will, or unfree will is one that causes us to do things that are not part of that function. It may even cause us to do things contrary to that function.

New Proposed Definition of Free will:

"An agent freely performs an action if the agent's mind is functioning properly. This includes the idea that the agent's DGD (will) is able to perform the function of representing what is good for the agent." (320)

Replies to earlier objections:

Distinct Causation Argument:

One is free if one is the cause of one's own behavior.

Our actions are caused by our beliefs and desires.

Our beliefs and desires are caused by our environment and our genes.

We do not cause our own environment or our genes.

We do not cause our own behavior

We are not free

This seems deductively valid, so if it is false, it must have a false premise.

The problem is the idea of causation.

To be free I must be the cause of my own actions.

But if my desires are caused by outside forces, then I seem not to be.

Causal Story.

Environment + Genes -> Beliefs and Desires -> Action
 (Not me) (Me) ?

Sober maintains that something can be the cause, even if it is caused by something else.

Ex. (Dog bite causes rabies)

Rabid Squirrel bite dog -> Rabid Dog bites me -> I get Rabies

Even though the squirrel gave the dogs rabies, the cause of my getting rabies is the dog biting me.

Moral Responsibility:

Perhaps more than being part of the causal chain is required. A person is morally responsible for some act, if (1) the person is the cause of the act, and (2) the act reflects the person's character.

This doesn't refute the version of free will since:

Our moral characters do not need to be uncaused (or do they?)

What would be be (as agents) if not the collection of beliefs and desires we have?

Reply 2

Could have Done Otherwise Argument

One acts freely if one could have done otherwise

Given any fixed current state (beliefs and desires) one cannot do otherwise

One cannot act freely.

P1 suffers from the Locke's Locked Room example

If I choose to sit in a room, I do so freely even if unbeknownst to me someone locks the door preventing me from doing otherwise.

P2 is ambiguous. If it means that because of the beliefs and desires one cannot act differently, it may simply not be true. I.e. indeterminism, from quantum states or what have you may make it the case that beliefs and desires do not uniquely determine some action. They merely make some action more probable than others. Even if it is true, it remains to be shown how the inability to do otherwise prevents free will.

Objection: Coerced actions are unfree.

According to the weather vane theory, even if someone is coerced, a gun held to head, to act a certain way, they are still free (so long as their mind is working properly). But that hardly seems free.

Reply:

It is true that coerced actions are still free according to the WVT. But that is alright. The situation of coercion is very different from the cases like Kleptomania or severe OCD.

We can rationally choose to do things even if we would rather not. We do so all the time, coercion is just a matter of degree.

Objection to Weather Vane Theory (Free Self Sacrifice)

The idea that function of the DGD or will is to produce desires that represent what is good for the agent is too vague or it is mistaken. Sometimes we freely rationally choose to sacrifice ourselves (in some way) and there is no malfunction in doing so.

Can the desire to risk or sacrifice one's life itself be something which is in our own interest? It does not seem so.

Other interpretations:

The DGD produces desires that the agent *thinks* represent the good to her.

Problem: The kleptomaniac may think stealing is good for her.

The DGD produces desires that represent the good for the gene (selfish gene theory).

Problem: This needs proof, or at least a reasonable articulation. Isn't universal self-sacrifice also bad for the gene?

Any alternatives?

Review Questions:

1. It is free if it caused by the sort of things that it is intended to be caused by, and nothing else. It is both free and determined by causes.
2. The heart evolved to be the way it is because it serves a particular function, that of pumping blood.
3. Moral responsibility seems to imply that the action was the result of the person's character not merely some accident or thoughtless act.
4. Just because something is caused to be a certain way doesn't mean that it is not the cause of some action. Beliefs and Desires can still be the cause of my actions, even though they are caused by something else.
There is the locked room problem: I may be free even if I cannot do otherwise. And there is the real possibility of indeterminism which would render this condition simply false. Also it is not clear that the ability to do otherwise is needed at all for free actions.
5. Yes. The WeatherVane theory is plausible even though it makes coerced actions free.
6. If actions are free when they arise from well functioning desire generation devices. And DGDs are defined as promoting personal survival, sacrifice seems to run contrary to that goal. So such actions would be unfree. But that seems obviously false.

Russell Readings: Other minds are known by analogy

How do we know other people have minds and mental states if our main evidence for these is our own experience?

Argument by Analogy.

When I have mental states, I behave in certain ways.

Other people behave in those same ways,

Therefore other people have minds.

Although this is not very strong on its own, Russell adds that the very complex behaviors we see add plausibility to the claim.

Smart Readings: Mental Processes are Physical

We should avoid the claim that some mental processes are irreducibly psychical (mental). We admit that all other phenomena have only physical explanations, having long given up supernatural explanations. We also believe that all mental phenomena are correlated with certain physical (brain/neurochemical) phenomena. Why do we insist they are different?

Occam's Razor tells us to accept the simplest answer (between equivalent ones).

This does not mean that all statements about mental events can be translated as statements about physical events. My experience of a painting may be blue, but my brain processes are not blue.

Obj 1. Anyone can have beliefs about the mind, but only a few can understand them about the brain processes. Thus they must be different.

Reply: We can be ignorant of things under one description but not under another.

Obj 2. Our present theories might be wrong.

Reply: That only means that our theories do not mean the same as our mental claims, not that they are wrong.

Obj. 4. The sensation is not in physical space, but the brain process is.

Reply. It is the experience of the sensation that takes place in physical space not the sensation itself.

Obj. 5. Brain processes can be swift or slow, but experiences can't be swift or slow.

Reply. The terms may have a different logic, but that does not mean that they do not refer to the same thing.

Obj. 6. Sensations are private, brain processes are public. A doctor could not look into your brain and 'see' what you are sensing.

Reply: This true now, but we may develop neuroscience to the point that effectively we can tell what sensations are going on by looking at the brain.

Obj. 7. I can imagine myself without a brain, or stone and still having sensations.

Reply. We can imagine things that are false, that does not make them false.

Turing Readings: Computing Machinery and Intelligence

1. The Imitation Game

"Can machines think?" is an ill-formed question.

"Machine" and "Think" are not well defined terms.

Revised Question.

The Imitation Game.

Three people participate. Two a man (a) and a woman (b) and the third (c) the interlocutor (c). (c) tries to figure out through a series of questions and answers who is the man and who is the woman. (a) tries to convince (c) that he is in fact the woman. (b) tries to convince him that she is.

The *revised question* is: "What will happen if we allow a machine to take the place of (a) in the game? Will the interrogator guess wrongly as often as when it was played by a person?"

"Are there imaginable digital computers that would do well in the imitation game?" (351)

2. Critique of the new problem.

Advantage: It clearly distinguishes mental from physical capacities.

Disadvantage: Heavily weighted against the machine.

Requirements: No practical demonstrations (only questions and answers)

3. The machines concerned in the game.

Anything that is not a 'living' computer a.k.a. a human (or brain)

In other words a digital computer.

4. Digital computers.

Note: Computer came from the word for a kind of accountants's aid. A person who computes.

Store = Memory

Executive Unit = Processor

Control = Operating System

One can make a computer entirely mechanical without any electrical parts. It is a function of what it does that makes it a computer.

5. Universality of Digital Computers.

Discrete state machines (let's not get too deep here). The machine is always in some knowable state.

Given the original state + Knowing all inputs. One can know the outputs and all subsequent states.

(Digital computers) can mimic any discrete state machine.

6. Contrary views on the Main Question.

Suppose a computer could convince (c) 30% of the time after 5 minutes.

By 2000, people will say "machines think" and not think it odd.

Objections.

O1. Theological Objection: Machines have no soul, thinking is a function of the soul. Therefore machines can't think.

R1. This limits the omnipotence of God too much. Why couldn't God make a machine that has a soul.

O2. Heads in Sand. The idea of thinking machines is too terrible to imagine. So it is false.

R2. No refutation necessary.

O3. Mathematical Objection: There are certain questions pertaining to mathematics and logic that computers cannot answer. If asked those we can predict the answer the computer must give.

R3. There is no reason to believe that people are also so limited. And. The fact that some people might know something that some computers don't is hardly that impressive.

O4. The argument from consciousness: Computers cannot feel any emotion. So they cannot be the same as a brain.

R4. This rejects the validity of the test. It leads to solipsism since it requires that one know the thoughts from inside (as it were). Viva Vce (test to determine real knowledge from parroting).

O5. Argument from various disabilities: Computers lack all sorts of emotional abilities.

R5. Some of the disabilities are on the part of man, not the machine. Question begging.

O6. Lady Lovelaces Objection: Nothing original can come from a machine.

R6. Can something original really come from nothing in a person? If you mean surprise, then computers can surprise us all the time (try using Windows 2000)

O7. Argument from the continuity of the Nervous System. Small changes in the nervous system could have big changes in the outputs.

R7. This would not hurt the computer in the imitation game. It could be programmed to simulate that.

O8 Argument from the Informality of Behavior. People do not have a set of laws determining their behavior.

R8. This has yet to be determined by science and cannot simply be asserted.

O9. Argument from ESP. A person with ESP could tell the difference between a person and a computer.

R8. Maybe we need a stricter test, that prohibits this sort of playing. No telepathy.

Learning Machines.

Can a machine ever learn?

Possible answers:

(1) Learning is really just performing in some unexpected way. Right now computers don't really learn. But in time they might.

(2) Brain states may be like the skin of an onion. Each layer of mechanism reveals another until there is nothing left underneath.

In designing the computer and program for the millenium imitation game, we should think about what the education of a child requires. What sort of elements do real children get that our learning machine should have?

Hume Readings: Of Liberty and Necessity

Part I. Despite the appearances to the contrary all people actually agree on the compatibility of free will and determinism (liberty and necessity.)

Necessity: It is universally allowed that all events are determined uniquely by their causes.

There is also great uniformity in the actions of men. Human nature is the same in its principles and operations. The same motives always produce the same actions.

To know the desires of the Greeks and Romans, study the French and English. They are all the same.

Contrary to the popular idea that volitions are less determined than physical events.

Consider that a prisoner not think his jailors likely to let him go, because they are not determined. Rather they appear as hard as the stone walls.

Part II. Theological Objections

If we are determined by necessity to do as we do. Why should blame ourselves or others for their actions rather than the real source of their actions God?

Campbell Readings: Has the self "Free Will"?

The question of free will needs to be addressed as a moral question. Otherwise it will get bogged down in metaphysical conundrums.

(1). Real Question When is it appropriate to blame someone for their actions?

Moral Freedom pertains to inner acts (we do not consider Robots to be morally blamable).

(2a). Moral responsibility can only be extended to a person for the part of some inner act for which he or she is solely responsible. We are morally responsible for the acts of which we are the sole author.

(2b). We are not responsible for those acts we could not help but do. We must, in some sense, be able to do otherwise, if we are to be blamed for what we do.

(Note that (2a) and (2b), are elucidations of the Distant Cause and Could have Done Differently Requires for Freedom c.f. Sober Lecture 24, 25.)

(5) Since we accept that we are determined in some sense (heredity, genes). And admit that in some cases, that might be enough to override *prima facie* moral judgments. There must be some kernel of sole authorship aside from that in our actions. This might be the idea of *moral effort* which is not itself affected by outside forces.

(6) We have internal evidence of the moral struggle and the moral effort. But is that evidence sufficient to prove that it is underdetermined?

(7) Can we disbelieve that we could have acted differently?

We seem to be able to doubt that on a theoretical level if not a practical one (we may not be able to decide now that we do not choose).

There seems to be a disjoint between practical and theoretical reasoning, who wins?

Determinism has two sorts of arguments:

Prediction, we can predict someone's actions.

But free will allows for prediction

The meaninglessness of saying a "my choice" without reference to "my character."

Otherwise we must say that a person acts "for no reason at all." Which seems false or unintelligible.

Our mere defining "intelligible" as following from some cause necessarily. If so, that is not really fair to Libertarianism. It is at best a circular argument. But if unintelligible means something else, like cannot be made sense of. Then it needs some proof. If we adopt the inner standpoint (from which we experience our own decision making) then our behavior seems to be clearly *ours*.

See the fivepoint at the end of the reading. (389-390)

Skinner Readings (Determinism Rules Out Freedom)

Hypothesis: Freedom is an illusion.

It need not be predictable (just as the weather is not, though it is determined)

Coercion (force by punishment or threat) is not effective in controlling people. That is why that model of non-freedom is a bad one.

"Restraint is only one sort of control, the absence of restraint isn't freedom." (395)

"By skillful planning, by a wise choice of techniques we increase the feeling to freedom."
(395)