Mind & Body

Wednesday, 11/2

The Mind-Body Problem has been an enduring one in philosophy. Recently it has attracted increased interest because many people believe that it must be solved before we can decide whether there is such a thing as artificial intelligence.

Today we will define the problem and consider one proposed solution to it, materialism as presented in the article by Elliot.

MIND-BODY PROBLEM

1) We directly experience our minds and their operations. List some operations of the mind and some characteristics of the mental states that accompany these operations. (For example, our conscious experience about our physical surroundings is made up of visual images – as the computer software people would say, our minds have a graphical interface with the external world.)

2) We have physical organs called brains. List some of its parts and the ways they work to process information from the external world and to cause our behavior.

3) What is the relation between mind and brain? Where in the physical brain do we find the visual images of which our mind are aware? How do our mental states cause the brain operations that lead to our behavior?

MATERIALISM

As pointed out by Elliot, materialism is a theory about the relation between mind and brain (that is, mind and body), not a commitment to the acquisition of material possessions.

Elliot says that materialism is grounded on three propositions:

- the Law of Universal Causation,
- the explanation of teleology (that is, behavior that is done to achieve a future purpose) in terms of prior causes, and
- monism (that is, the denial that mental entities exist and the assertion that all existing things are purely physical).

Clearly understand each of these propositions:

1) The Law of Universal Causation [p. 375 (bottom) - p. 376 (top)]
2) The explanation of purposive behavior in terms of prior causes [p. 376 (bottom) - p. 378 (top)]
3) Monism [p. 378 (bottom) - p. 381]
Are you convinced these propositions are true? Why, or why not? Philosophers who object to materialism say that we cannot deny the reality of our mental experience, and that it is totally implausible to say this mental experience does not really exist. Do you agree?

FRIDAY, 11/4

Some of philosophy's most famous puzzles relate to the mind-body problem. Today we will look at some of them. But first you need to review Joad's main point:

Last time we saw that Elliot argued that only brain exists. This leaves unexplained the mental images of which we are all consciously aware, but a materialist like Elliot would say that this experience is nothing more than a product of the brain's functioning (even if we cannot fully explain at this time how completely physical processes give us experiences that "feel" immaterial).

Joad presents a dualist position. What is this? (Look at the first paragraph on p. 382, the next to last paragraph on p. 386, and the first paragraph in the section titled "Summary of Argument" that starts on p. 386.)

First puzzle: If a tree falls in a forest and no one is there, does it make a sound?
   How might a materialist answer? A dualist?

Second puzzle: You are on a space journey and encounter an entity that displays many characteristics of being a human (for example, Data on Star Trek: The Next Generation), but you are not sure. What evidence would settle the question for you?
   How might a materialist answer? A dualist?

WRITTEN HOMEWORK #9 – Due in class Friday, November 4.
   For either the 1st or 2nd puzzle, answer the question from both the materialist and dualist perspective.

MONDAY, 11/7 – NO CLASS

WEDNESDAY, 11/9

We have been looking at different theories about what we should believe regarding our own minds – for example, the materialist view that it is an illusion that has no connection with the brain that constitutes the only reality, or the dualist view that it is real and has a specifiable connection with the underlying brain that also is real. Independent of this issue is the question of “other minds.” Do we have sufficient evidence to believe that there are minds other than our own? And, if we do, what should we believe about these? Hospers discusses this issue.

1) We directly encounter our own experiences and feelings, for example, pain and heat. Whatever we believe about their real status – that is, whether they are mental entities or, on the contrary, physical brain states – we have immediate and non-inferential knowledge about them. However, we do not have direct knowledge of the same type of
other people’s experiences and feelings. We can only infer that they have them from their behavior. How confident can we be about our knowledge of other people’s experiences and feelings? Can we verify it? Confirm it? Provide any kind of strong evidence for it?

2) Moreover, if we do not have direct access to the mental experiences and feelings of other people, can we even be sure that they have them at all? Or even that they have minds? What if we admit that we can only infer the existence of other minds from people’s behavior? Would this require us also to admit that we have to say mind is present in ALL cases where the same type of behavior is present?

FRIDAY, 11/11

We have a strong sense of self-identity – that is, we believe in a "me" that continues pretty much unchanged in spite of changing experiences, beliefs, thoughts, memories, and values. Parfit tries to show that this is mistaken, and he suggests we should replace the notion of self-identity with the notion of psychological continuity.

1) What does Parfit mean by the notion of psychological continuity, and how does it differ from self-identity?

2) Consider some of the examples he uses to support his notion:

a) If your brain were transplanted into someone else's body, would the resulting person still be you?

b) If half your brain were transplanted into one person's body and the other half into another person's body, which one – if either – would be you?

c) If a technique were developed for replacing brain cells, what percentage of your brain cells could be replaced before you changed into another person?

3) Parfit's examples seem somewhat far-fetched, but people who undergo major changes in what they believe or in how they behave often say such things as, "I'm a different person now," "I've changed," and "That was the old me." Do they mean these things literally? What do they mean – that although they have psychological continuity with their former selves, they do not have identity?

MONDAY, 11/14

Evans' article departs from Turing's claims about artificial intelligence and the famous "Turing test," and Beloff’s article also addresses the question of artificial intelligence.

In 1950 A. M. Turing published "Computing Machinery and Intelligence," the classic paper on the issue of artificial intelligence. Turing claimed that the question, "Can a machine
think?" can be answered by putting a computer to a practical test. The test is this: can a computer beat a human in the "Imitation Game?"

Turing defines the "Imitation Game" by giving an example. Consider a game played by three people – a man, a woman, and an interrogator who may be of either sex. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman; the object for the man is to imitate a woman so well that the interrogator incorrectly identifies him as the woman; and the object for the woman is to help the interrogator correctly identify her as the woman. In order that tones of voice and other sensory cues may not help the interrogator, the interrogator is isolated in a separate room and communicates with the man and woman by means of a teleprinter (today this would be a keyboard and monitor).

It is clearly the case that some men could win this game; they could successfully lead the interrogator to incorrectly identify them as women. What would a win by the man prove? There appear to be at least four possibilities. 1) A win by the man proves that he is a woman. 2) A win by the man proves that his mind works like a woman’s mind. 3) A win by the man proves that he can simulate a woman’s thought processes. 4) A win by the man proves that he can produce linguistic behavior that is indistinguishable from that produced by a woman. The first possibility is, of course, ridiculous. The second and third possibilities go far beyond the evidence – they require a number of questionable assumptions. The fourth possibility is really the only one supported by the evidence. So, the most we can say about a win by the man in the man-woman "Imitation Game" is that the man can hold a conversation with an interrogator that is indistinguishable from a conversation that a woman might hold – even when the interrogator is trying to trip him up (that is, the man must be able to follow the conversation wherever the interrogator leads).

Turing then defines a second version of the imitation game. An interrogator plays against a computer and a person. The object of the game for the interrogator is to correctly identify the computer; the object for the computer is to lead the interrogator into incorrectly identifying it as the person; the object for the person is to help the interrogator correctly identify him or her as the person. Turing then asks, "What will happen when this version of the "Imitation Game" is played? Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'"

It is at least conceivable that a computer might one day pass the Turing test and win the person-computer version of the "Imitation Game." What should we conclude about the computer is it does?

Here are the possibilities as stated by Turing:

1) The computer is an artificial person, that is, a machine entity which is identical with the person in every way.

2) The computer has an artificial mind, that is, a machine-hosted mind which is identical with the person's mind, where the human mind consists of a set of determinate states such as emotions, consciousness, self-consciousness, and so on.
3) The computer is an artificial thinker (as in, "Can machines think?"), that is, a machine capable of performing thought which is identical with a person's thought, where human thought is a process of a determinate type, utilizing a given set of operations, structures, and so on.

4) The computer has artificial intelligence, that is, is capable of producing behavior which when displayed by person is taken to be a product or accomplishment of their intelligence.

Turing dismisses the first three possibilities out of hand because they rest on false assumptions and/or require solutions to problems that have proven intractable. He says that the fourth possibility is the correct thing to say about a computer that passes his so-called Turing test. Note, however, exactly what this fourth possibility is: The term “artificial intelligence” is defined in terms of “computer output” and has absolutely nothing to do with “thinking,” “intelligence” as we ordinarily apply this to humans as part or our humanness, or the presence of “mind.”

_Evans_ begins his article with a lengthy rebuttal of various arguments against the possibility of computers being able to “think.” Ask yourself if any of the arguments that Evans rejects would count against Turing’s position. Then, read Evans section “Towards the Ultra-Intelligent Machine” and ask if he has interpreted Turing correctly (or, at least, as I have interpreted him above).

_Beloff_ argues that machines will never be able to perform all of the mental processes of humans and, therefore, can never possibly possess artificial intelligence. I think Turing might read Beloff’s article and say, "That's right. So what? Artificial intelligence does not require machines to process information in those ways." So, if you are with Turing, you will see Beloff’s arguments as irrelevant, but if you think Turing missed the boat, you will see Beloff’s arguments as decisively demonstrating that artificial intelligence is impossible.

**WEDNESDAY, 11/16**

Searle's article presents a different perspective on artificial intelligence. He admits that the computer can produce many results that are identical with the results produced by human intelligence. However, he claims that computers and humans produce these results in fundamentally different ways. This means that if what humans do is "thinking," computers do something else, and it is therefore incorrect to say "computers think."

Searle argues for his point by giving the parable of the Chinese room. In class we will discuss this thought experiment, and I will ask you to undertake another.