Neurophilosophy

PHI-405 (Philosophy of Cognitive Science)

Instructor: Sreejith K. K.

sreejith@mail.vidyasagar.ac.in

Introduction: What Is Neurophilosophy?

- "Neurophilosophy" explores the impact of discoveries in neuroscience on a range of traditional philosophical questions about the nature of the mind.
- This subfield aims to move forward on questions such as the nature of knowledge and learning, decision making and choice, and self-control and habits by drawing on data from the relevant sciences – not only neuroscience and clinical neurology but also evolutionary biology, experimental psychology, behavioral economics, anthropology, and genetics.
- It draws also on lessons from the history of philosophy and the history of science, which saw mysteries about the nature of the blood or fire or infectious disease become less mysterious as experimental science began to provide new observations and tested explanations.

- The massive accumulation of neurobiological data from many levels of brain organization and many species of nervous systems is a recent development because neuroscience did not really reach full steam until about the 1970s.
- Although clinical observations had long implicated the brain in mental functions, understanding exactly *why* lesions affected mental functions remained out of reach. This was so because essentially until very recently nothing was known about the microstructure of brains – about neurons and how neurons worked, about how the brain was organized into networks and systems, and about how neurochemicals mediated interactions between neurons.

- Notice that detailed drawings of nerve cells were produced by Camillo Golgi and Ramón y Cajal only in the latter part of the nineteenth century.
- How neurons *interacted* with each other to yield effects such as a behavior was still a profound mystery.
- Chemistry, by contrast, was a vastly more mature science in the early nineteenth century, strengthened by basic organizing principles of atomic theory, as outlined by Dalton in 1805 and a clear appreciation of the fundamental elements – no longer deemed to be earth, air, fire, and water. Instead, the elements were characterized by Mendeleyev in the 1880s in the periodic table – things such as oxygen, hydrogen, tin, and gold.

- As for neuroscience, it is perhaps surprising to realize that the existence of *inhibitory* connections between nerve cells was demonstrated by John Eccles and colleagues only in the 1950s. Physics, far more mature in terms of theory and explanation by that time, had begun to investigate the inner structure of the atom.

Neuroscience is a young science.

- To get a perspective here, note that effective brain imaging techniques came into their own only in the last two decades of the twentieth century. At the micro level, many details regarding the synapse and how neurons communicate are not completely understood even now, nor are the functions and dynamics of neural *networks*. Neuroscience is a young science.

- Because the brain's basic units work by changes in voltage across the cell membrane and by chemicals that regulate such changes, and because the units are not visible to the naked eye, development has depended on a theoretically and experimentally rich physics and chemistry.
- Specifically, neuroscience depends on tools and devices that exploit the knowledge of physics and chemistry, for example, the electron microscope, microelectrodes, nuclear magnetic resonance, monoclonal antibodies, and most recently, optogenetics.
- It is noteworthy that understanding how neurons work required knowledge of electricity, and that knowledge was not in hand until Michael Faraday's discoveries in the first half of the nineteenth century.

- Some philosophers take it as dead obvious that the enduring existence of many puzzles in neuroscience entails that neuroscience can never, *ever* discover much in the way of mechanisms of cognitive function.
- One major reason for this conclusion is that they have generally failed to appreciate the clear historical point that the sciences of the nervous system are very young indeed.

The Relation Between Mind and Brain

- The words "mind" and "brain" are distinct. Even so, that linguistic fact leaves it open whether mental processes are in fact processes of the physical brain. (Remember: water and H_2O are different words, but they do name the very same stuff.)
- A favored theory in philosophical thought, championed by Plato, developed by Descartes, and even now defended by Thomas Nagel (2012), holds that just as the words are distinct, so too are the processes.
- This approach is known as "dualism" – a "two stuffs" theory embracing physical stuff and the utterly different soul stuff.
- Thinking, seeing, and choosing, according to dualism, are processes of the nonphysical mind or soul.

- For dualists, the mind/body problem is the problem of how a physical state of the brain can interact with a totally nonphysical state of the soul.
- By contrast, according to an equally venerable if less popular tradition, there is only the brain; mental processes are processes of the physical brain whose exact nature remains to be discovered. This is known as "physicalism" and found adherents in Hippocrates, Hobbes, Hume, and Helmholtz.
- Physicalists realize that there is no problem about how the mind and body *interact* inasmuch as there are not two things, but only one thing: the brain.
- The mind is what the brain does.

- For them, the important problem concerns how the brain learns and remembers, how the brain enables us to see and hear and think, and how it enables us to move our eyes, legs, and whole body. Their problem concerns the nature of the brain mechanisms that support mental phenomena.
- Interestingly, dualists also have a closely related set of problems: how does soul stuff work such that we learn and remember, see and hear and think, and so forth.
- Whereas in neuroscience physicalists have a vibrant research program to address their questions, dualists have no comparable program. No one has the slightest idea how soul stuff does *anything*.

- Neurophilosophy as a research program has poor prospects unless mental processes such as remembering and attending are processes of the brain. Otherwise, we should just study the stuff that *does* perform attending and remembering and find out how *that* works, stuff such as the "soul stuff" postulated by Descartes.
- At this stage in the sciences, the evidence overwhelmingly indicates that all mental events and processes, including visual or auditory perception, learning, memory, language use, and decision making, are in fact events and processes of the physical brain.
- It is not that there is one single experiment that decisively shows this. Rather, the evidence has steadily accumulated over countless observations and experiments, and no counterevidence raises doubts.

Split-brain studies

- One of the most dramatic observations of mind/brain dependency came from the split-brain studies published in the late 1960s.
- These studies involved patients whose cerebral hemispheres were surgically separated in order to treat drug-resistant epilepsy.
- The nerve sheet connecting the two hemispheres – the corpus callosum – was the structure that was cut, thereby disconnecting the cortex of the right and left hemispheres.
- The aim was to aid the patient by preventing a seizure from traveling from its origin in one hemisphere to the other hemisphere.
- Astonishingly, tests of "split brain" subjects showed that the mental life of the two hemispheres was also disconnected: the right hemisphere might have knowledge the left did not or see something or decide something that the left did not, for example.

- The implications for the mind/body problem were obvious: if mental states were *not* brain states, why would cutting the corpus callosum allow knowledge and experience to be confined to activity in *one* hemisphere?
- Although a defiant dualist might invent some story to accommodate the facts (and a diehard few did this), the best and most reasonable explanation for the disconnection effects was simply that a *physical* pathway was interrupted, a pathway essential for mental unity, and that soul stuff was just not in the game.
- As Michael Gazzaniga (2015), one of the leading split-brain researchers puts it, consciousness can be split.

- The many observations made by clinical neurologists of patients who suffered focal brain damage also weighed in.
- Focal brain damage could result in highly specific losses of cognitive function, such as the loss of the capacity to recognize familiar faces, loss of recognition of a limb as one's own, and loss of the capacity to perform an action on command, such as saluting or waving hello.
- Studies of a few patients who had suffered bilateral damage to the hippocampus (a small curved structure beneath the cerebral cortex) showed them to be severely impaired in learning new things (anterograde amnesia).
- This finding initiated a massive research program to understand the relation between learning and memory and the hippocampal structures.

- Memory losses associated with dementing diseases also linked memory with neural loss and further suggested the tight link between the mental and the neural.
- Important also are studies of attention using brain imaging along with single neuron physiology.
- These varied studies suggest that at least three anatomic networks, connected but somewhat independent of the other, are involved in different aspects of attention: alerting, orienting, and executive control. Moreover, each of these functions has been the target of detailed further study, indicating, for example, that there are strong associations between these functions and awareness, especially between detection of a target (consequent on orienting) and awareness (Petersen and Posner 2012).

- Developments in psychology, especially visual psychology, also implicated neural networks in mental functions, and this work tended to dovetail well with the neuroscientific findings on the visual system.
- Explanations of color vision, for example, depended on the retina's three cone types and on opponent processing by neurons in cortical areas.
- It was well appreciated that much in the world – such as ultra violet and radio waves – could not be detected by our visual system because of its physical organization.
- Perception of visual motion was linked to the behavior of single neurons in a visually sensitive area of cortex known as MT (middle temporal).

- Visual hallucinations were known to be caused by physical substances such as LSD or ketamine.
- consciousness could be obliterated by drugs such as ether (as well as by other substances employed by anesthesiologists, such as propofol).
- No evidence linked these drugs to soul stuff.
- On the contrary, many anesthetics appear to work by altering the normal balance of excitation and inhibition of neurons in circuits.

- Short term memory can be transiently blocked by a blow to the head or by a drug such as scopolamine.
- Emotions and moods can be affected by Prozac and by alcohol; decision making can be affected by hunger, fear, sleeplessness, and cocaine; elevated levels of cortisol cause anxiety.
- Very specific changes in whole brain activity corresponding to periods of sleep versus dreaming versus being awake have been documented, and explanations for the neuronal signature typifying these three states have made considerable progress.
- In aggregate, these findings weighed in favor of the hypothesis that mental functions are a subset of functions of the physical brain, not of some spooky "soul stuff."

- Evolutionary biology encouraged us to dwell on the fact that nervous systems are the product of evolution and that the human nervous system is no exception.
- Comparisons of anatomy, between human and nonhuman nervous systems, have revealed that the functional organization, at both macro and micro levels, has been highly conserved over hundreds of millions of years.
- Although human brains are larger than the brains of other land mammals, we share all the same structures, pathways, innervation patterns, neuronal types, and neurochemicals.

- Neurons in a fruit fly work essentially the same way as neurons in the human brain. Molecular biology revealed that the genetic differences between humans and our nearest relatives, chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*), are very small .
- These evolutionary relationships imply that either no mammals have nonphysical souls or all do.
- Now questions flood in: if humans *alone* do have a soul, where do human souls come from, and why does the soul suddenly appear, some 4 million years after the *Homo* species branched off from our common ancestor with chimpanzees? Did extinct *Homo* species such as *Homo erectus* and *Homo neanderthalensis* have souls too?

- Based on cranial measurements, anthropologists believe that the brains of *Homo neanderthalensis* were typically larger than our brains.
- Neanderthals probably had some form of acoustic communication even though they may not have been able to make all the vocalizations of which humans are capable (Lieberman 2013).
- Moreover, genetic data reveal that they did interbreed with *Homo sapiens* (Pääbo 2014).
- What about *their* souls? Still other questions challenge the idea that the human soul, not the human brain, is the repository of all that makes us clever. How can ravens and rats and monkeys solve complex problems – how can they sleep, dream, pay attention, and so forth – if a soul is needed for such functions?

- By the 1980s, there was impressive, if cautious, agreement among scientists as well as philosophers that the existence of a nonphysical soul that feels, decides, sees, and reasons was improbable.
- Where disagreement flourished unabated, however, concerned whether neuroscience could *explain* those functions, physical though they may be.
- Neuroscientists tended to expect that with new techniques and more experiments, progress would continue to be made. How far we shall get, time and research effort will tell.
- Some philosophers, by contrast, confidently predicted that neuroscience would never explain cognitive functions, a view particularly associated with Jerry Fodor.
- This view tended to be known as the "autonomy of psychology" – autonomous with respect to other sciences, especially neuroscience.

- It is important to understand that this claim about the limits of neuroscience was just a *prediction*, and it was supported by philosophical speculation, not scientific evidence.
- Although highly popular until about 1990, the idea has slowly and systematically been undercut by actual progress in the neurosciences, especially by increasingly suggestive links between data at the behavioral, whole-brain, and neural levels.

- One further reason for ignoring much of neuroscience arose from a misguided analogy. The idea was that cognition is like running soft- ware on a computer, where the brain is analogous the computer hardware.
- Just as you need not know anything about a computer's hardware to understand an application such as PowerPoint, so you need not understand anything about the brain to understand cognition, or so the argument went.
- To anyone who looks at all closely at the brain, the disanalogies between brains and conventional computers are so numerous and so profound that the brain/hardware analogy was not taken seriously in neuroscience or bioengineering.

- Not least among the differences are that brains are parallel not serial processors, that storage and processing in brains are not done by separate modules but by the same structures, and that brains change their structure as they develop from gestation to adulthood and at all stages as they learn.
- The point where influential philosophers are still confident that the mysteries permanently have the upper hand concerns conscious experience. Typically, there are two distinct arguments to support this con- viction.
- The first argument makes a straightforward prediction about where science will go in the future. It is based on current intuitions about the tractability of the problem of explaining consciousness in neurobiological terms. With great confidence it will be claimed that consciousness is so completely and utterly and thoroughly mysterious, it will *never* be explained at all, period (McGinn 2012, 2014).

- By way of illustration, it may be suggested that expecting any science to explain how conscious experience emerges from the activity of neurons is like expecting a rat to understand differential equations.
- Despite its chest- pounding confidence, this prediction should be taken with ample doses of caution because predicting where science will go and what will be discovered is really a rather risky business, to put it politely.
- The second and more influential argument rests on the dualist's belief that although nonconscious events such as memory consolidation and preprocessing in vision are brain events, *conscious* events such as feeling nauseous are not brain events. Hence neuroscience cannot explain them.

- Thus, when I am aware of a pain in my tooth or a decision to kick off my shoes, some philosophers, such as David Chalmers (1996) and Thomas Nagel (2012), consider those conscious events to be extraphysical, merely running parallel to the physical events.
- A methodological point may be pertinent in regard to the dualist's argument: however large and systematic the mass of empirical evidence supporting the empirical hypothesis that consciousness is a brain function, it is always a logically *consistent* option to be stubborn and to insist otherwise, as do Chalmers and Nagel. Here is the way to think of this: identities – such as that temperature really *is* mean molecular kinetic energy, for example – are not directly observable. They are underwritten by inferences that best account for the mass of data and the appreciation that no explanatory competitor is as successful. One could, if determined, dig in one's heels and say, "temperature is *not* mean molecular kinetic KE, but rather an occult phenomenon that merely runs parallel to KE" (Churchland 1996b). It is a logically consistent position, even if it is not a reasonable position.

- In the case of conscious experience, although philosophers such as Chalmers and Nagel express their reservations about the brain, the only thing they really do have are reservations.
- Moreover, their reservations are based on intuitions about how different experience seems to be from states occurring in the physical brain. They have neither competing experiments nor a competing hypothesis with any power or detail; in particular, they have no hypothesis that surpasses let alone competes seriously with the neuroscientific hypothesis.
- How do the dualists address the dependencies – the causal dependencies that suggest identification – between consciousness and brain activities?
- A favored strategy is to propose that conscious states just run parallel to brain states. This proposal may be embellished, perhaps by the idea that conscious states neither cause nor are caused by brain states – the two streams are causally isolated.

- A variation of this opts instead for a one-way causal street – brain states cause conscious states, but conscious states do not cause brain states.
- Traditionally, the view that mental states do not cause brain states is called "epiphenomenalism." Actual evidence is lacking for both hypotheses – both are merely empty denials of the idea that consciousness is a biological phenomenon.
- Historically, the most renowned defender of two-way causal isolation was Gottfried Leibniz.
- Leibniz held this view because he thought that it was inconceivable that completely different substances could interact causally. If they shared no properties – not even spatial properties – how could they affect each other?

- Moreover, with the benefit of contemporary physics, we can see that the causal interaction between *nonphysical* stuff such as a soul with *physical* stuff such as electrons would be an anomaly relative to the current and rather well-established laws of physics.
- More exactly, it would affect the law of conservation of energy. If brains can cause changes external to the physical domain, there should be an anomaly with respect to conservation of energy. No such anomaly has ever been seen or measured.
- The absence of anomalous data suggests either that the hypothesis of a nonphysical conscious stream of states lacks credibility or that the conscious stream of conscious states does not interact with brain states at all.

- When the neuroscientist Josef Parvizi used a tiny electrical stimulus to activate a very specific part of the brain (the middle cingulate gyrus) as part of the preparation of his human patient for surgery, his patient described the emergence of a conscious state consisting of the determi- nation to muster courage to deal with a problem. When the stimulus was off, the feeling vanished (Parvisi et al. 2013; P. S. Churchland 2013b).
- This experiential event was repeatable in that patient. Moreover, a very similar state was also reproducible in yet another patient stimulated in the same region. The reasonable conclusion is that the stimulus caused the change in conscious state.
- Some naysayers may wish to take the option that the brain events and the experienced event happen synchronously without causation: the experience stream and the brain stream are separate.

- What keeps the two streams synchronized? That is the stunning puzzle that emerges from the epiphenomenal hypothesis.
- Here is how Leibniz dealt with the puzzle: God sets up and maintains a "pre- established harmony" to keep mental and physical states properly aligned.
- Leibniz' solution is completely ad hoc, cobbled together to in order to fill an embarrassing silence.
- Chalmers' does not appeal to God, but he does advert to a future physics that allegedly will explain the alignment between noninteracting streams of mental and brain events. A revolutionary new physics, according to Chalmers' (1996) conjecture, ultimately will explain the nature of consciousness as a nonbrain phenomenon.

- Churchland: I have been unable to escape the feeling that this is really the old Leibniz solution suited up in the duds of a future physics instead of theology.
- Chalmers: Consciousness is so extraordinarily mysterious that only a revolution in physics will account for it.
- Churchland: My small sampling of physicists indicates that they do not wish to rush into investing heavily in a new physics just to address *consciousness,*

- If you are a dualist, either you can pretend that the huge accumulation of dependency evidence in neuroscience is not really there (not a realistic option), or you can say something substantial to address them. Rationally, something must be done insofar as this accumulation appears strongly to favor the hypothesis that conscious states are brain states.
- A novel strategy, tendered by Chalmers, claims that neuroscientific data are actually *neutral*, as between his parallel- stream hypothesis and the hypothesis that mental states are states of the physical brain.

- To assess the figures of merit of this "neural data neutral" strategy, try it elsewhere in science and see what results. Consider the nature of light as understood within contemporary physics: light is electromagnetic radiation (EMR) – light visible by humans is just one part of a larger spectrum that includes x-rays, microwaves, and so forth.
- Here is what the "neutral strategy" could say about light: "actually, the physical evidence is neutral between the hypothesis that light *is* EMR and that light is not EMR but a spooky thing. That is, light and EMR run in parallel streams, whose synchrony will be explained by a revolution in physics."

- Here is what the "neutral strategy" says about life: "all of cell biology is neutral between the hypothesis that life is an occult force (vitalism) and the hypothesis that life is the outcome of the biological structure and organization – cells, membranes, genes, ribosomes, mitochondria, and so forth."
- Scientifically, these "data neutral" proposals look counterproductive and more elaborate that the facts require.
- Silly though they may be, they are not, however, internally incoherent hypotheses.
- One bizarre claim that oddly appeals to various philosophers of mind is that if the "parallel stream" hypotheses are not internally contradictory, they are as reasonable as established scientific theories.

- Notice that it is not internally contradictory to say that the Earth is only one hour old, but it would be strange to say that this is as reasonable as saying it is about 5 billion years old.
- The twin predictions regarding mind and brain – that neuroscience will never account for conscious experience and that a revolution in physics will explain why – are generally motivated by emphasizing the difference between a neuron, on the one hand, and a feeling of tooth pain, on the other, for example.
- On reflection, it is argued, the differences appear to be so profound and so complete that surely, *surely* it is inconceivable that the pain in my tooth might really be the activity of neurons in the brain.

- Churchland: One problem with relying on what seems inconceivable is this: what is and is not conceivable is, after all, merely a psychological fact about us – about what we can and cannot imagine given our current beliefs and our capacity for imagination. It is not a metaphysical fact about the nature of the universe.
- An issue that spells trouble for a nonbrain theory of consciousness concerns the fact that the division between awareness and lack of awareness is typically blurry and often fluid.
- One place this really shows up is in the automatization of behavior as a skill is acquired, a commonplace phenomenon. Eg. Swimming, cycling, driving

- Here is the issue: are the many behavioral decisions of which I am unaware just mental brain events that blink out of the mental experience stream until an emergency arises and I must pay attention?
- Why do some philosophers of mind oppose so strenuously the two hypotheses: (1) mental states are states of the brain and (2) probably neuroscience can at least outline the mechanisms of cognitive functions?
- One reason: boundary: A strong assumption in the philosophy of mind is that philosophers are uniquely equipped to set the boundaries of what we can know and to outline the essential and enduring features of concepts that scientists might apply.

- A strong assumption in the philosophy of mind is that philosophers are uniquely equipped to set the boundaries of what we can know and to outline the essential and enduring features of concepts that scientists might apply.
- Philosophical intuition, in this view, is a special trained capacity that can home in on those necessary properties of a phenomenon that science must respect and not challenge.
- In this way, philosophy sets the foundations for the science. And if philosophers characterize necessary properties of the mind that intuition and logic show cannot be explained by properties of the brain, then that is the contribution of philosophy that science needs to honor

- Thus some philosophers of mind believe that they own a problem space that is concerned with conceptual necessities – necessary truths about psychological states and processes, discovered by conceptual analysis and so-called thought experiments.
- A necessary truth can- not, according to this approach, be falsified by scientific data. Intuitions trump data.
- Scientists, not surprisingly, are puzzled by where such *a priori* knowledge might really come from, and they do not want to be bamboozled by philosophical flimflam. After all, intuitions appear to be just strongly held beliefs that are likely grounded in education and reinforcement learning.

- Our intuitions can be mistaken.
- How could our intuitions be misguided?
- Here is how: complex nervous systems are not mere reflex machines or simple conditioning machines; they build models of the external world that are deployed in navigating the world.
- But not all models are equally accurate to the world itself.
- A mouse's model of the spatial world may be sufficient to get it around its environs given its limited goals, but it will not be as accurate as *my* model of the spatial world or indeed that of a wolf.

- I might success- fully update my causal model of the world as I come to realize that cholera is caused not by "bad air" but by bacteria. Somehow that information will modify and reshape my causal model of the world.
- However, a rainbow will still *look* like it has a location in space, even though I know full well that it does not.
- What about the model of attention and mental states generally? The model of mentality may persist in *seeming* to be spooky, even when I know "cognitively" that spooky is not accurate to the facts. This may be owed to deep biological features of the way the neural model works.