

Lecture 20: 1 April, 2021

PHYSICS 419 - Spring 2021

I. Introduction

Last lecture, we discussed Bell's inequality and its implication for local realism in quantum mechanics. Bell's inequality establishes a quantitative test for quantum mechanics or an additional theory which looks like quantum mechanics but with the added ingredient that the outcome of all measurements are pre-determined by classical variables as proposed by EPR (Einstein-Rosen-Podolsky). That is, ultimately, linear superposition has nothing to do with reality. Measurements on the correlations between the polarization of photon states has confirmed that the world is governed by quantum mechanics not any theory that adopts local hidden variables. What these experiments do not provide is a definitive test using a single measurement. This loop hole was fixed by the Greenberger-Horne-Zeillinger (GHZ) state. Unlike the two-particle states on which EPR is based, the GHZ state is a three-particle state, either composed of spins or photons. Lets take the example of spins. An example of a spin-GHZ state is

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|\uparrow\rangle_1 |\uparrow\rangle_2 |\uparrow\rangle_3 + |\downarrow\rangle_1 |\downarrow\rangle_2 |\downarrow\rangle_3), \quad (1)$$

in which $|\uparrow(\downarrow)\rangle_i$ indicates the spin of particle i . This is an entangled state as it cannot be written as a simple product of two different matrices. Lets put the spin of particle 1 along the x - axis and the spins of 2 and 3 along the y - axis. In the supplementary material (which provides the mathematical details on how to write the components of the spin along any axis you want) provided for this lecture, you will see that local hidden variables predicts, based on a cyclical permutation of the spins, that the product of the x - components of all three spins, $X_1 X_2 X_3 = -1$ but quantum mechanics predicts that $X_1 X_2 X_3 = 1$. Hence, a single quantum entangled state has the complexity needed to falsify the EPR claim. This is a truly amazing result, one not anticipated by Bell. No inequality is needed, just a single measurement on a

single quantum mechanical state. We live in a world in which $X_1X_2X_3 = 1$ not -1 . Hence, quantum mechanics wins. Some have argued that the experiments testing the Bell inequality represent the most important in science because it settles the reality debate. If this is so, then one could argue that the GHZ state is the most significant result in science. You can judge for yourself.

1 Quantum Mechanics and Consciousness

The failure of the EPR experiment to inject objective reality into quantum mechanics has left us right where we started. The problem with quantum mechanics is not that one cannot know the state of a system unless a measurement is made. The problem is deeper. Namely, the state of the system is completely indeterminate before the measurement. And this indeterminacy (ontic indeterminacy) is intrinsic to a quantum mechanical description of the system. It is what gives rise to the interference pattern we discussed in the two-slit experiment. If there is no interference, then there is no indeterminacy. It is this indeterminacy that EPR thought was unreasonable. That is, they did not think that a system acquired a sharp value of a particular observable at the time of measurement. They suspected that the value must be sharp in the first place otherwise it would not be sharp after the measurement. Wigner and von Neumann argued that consciousness led to collapse of the wavefunction. Precisely what consciousness is will be detailed in a moment. All advocates of consciousness theories, do not hold that consciousness collapses the wavefunction, however. D. Chalmers (see *Phil. Sci.* vol. 66, p. 370 (1999)) for a critique of the Chalmers account) has recently put forth a no-collapse version of quantum mechanics based on his theory of consciousness. Hugh Everett argued for a no-collapse version of quantum mechanics stating that the world really is not in any particular definite state, but that it just appears to be that

way. Chalmers invokes that “a superposed brain state should be associated with a number of distinct subjects of discrete experience.” That is, a brain could be in a superposition of perceiving a voltmeter having more than one value or a cat being dead or alive. The emphasis here is on our perception rather than the fact that voltmeters typically give one particular value for the voltage rather than some superposition of values. Both Chalmers and Wigner argued that there is something special about consciousness. They claim it lies beyond the laws of physics in the sense that it is not reducible and accountable by what is entailed in the physical. Notice this is similar to the fact that liquidity is irreducible because it is entailed from the molecular structure, liquidity is a global property or condition of the system (an emergent property). For Chalmers, consciousness is what is left over from an experience that does not supervene with metaphysical necessity on the physical. One cannot make the same statement about liquidity. That is, consciousness and the physical are not connected in a necessary sort of way. For example, there could be a copy of me in some possible world who does everything I do (for example, writing this lecture) without thinking or feeling anything. These are called unconscious zombies. Analogously, non-liquid water is just another state of water itself. For von Neumann human observation collapses the wave function, so a superposition is never observed. Chalmers argues for no collapse given the contingent relationship between consciousness and the physical.

This is a bit hard to argue with since (shades of Berkeley) we don't have much access to a world devoid of consciousness. However, there are some serious difficulties: The whole proposal requires putting people at the center of the existence of the universe. How does that square with everything else we know, e.g. evolution? The world we see shows overwhelming evidence of having once been free of consciousness. Were the laws of physics entirely different then? Who (bacterium, amoeba, monkey, Wigner, ...) was finally conscious enough to collapse the wave function and make positions, etc of particles exist? Just how did Wigner

get there before anything had positions?

While there is no evidence that consciousness plays some role distinct from any other phenomena involving macroscopic masses and times, it is worth taking a closer look at consciousness. In so doing, we will be able to determine why the Chalmers argument is not quite right.

2 Consciousness

What do we mean by consciousness? Conscious experience is a widespread phenomena. Whenever, we say that 'I really like that line of poetry,' or 'ouch that hurts' or 'I feel uncomfortable when you stand that close to me,' we are talking about mental states. Mental states (not to be confused necessarily with brain states) are what we mean by consciousness. One thing Descartes got right is that you cannot deny that you are conscious because such denial would be in part expressing a truth of your conscious experience. Notice that statements about one's conscious state are framed from the first person, that is, there is a perspective. Consider reading a line of poetry. A physical act occurs, namely your eyes scan the line. However, if you say after you read the line, 'ah that line really moved me', you are expressing more than what happened physically. You are talking about the mental state that you experienced after (or while) you read the line of poetry. You are expressing the subjective aspect of the experience. A subjective component to an experience is typically unique. Only you can have the particular reaction that you had. While many might feel moved by the same line, your particular feeling of being moved is truly your own. No one else will feel particularly the same way. Hence, consciousness generally refers to what is left over once we subtract the objective (or physical) aspect of an experience. An equivalent paraphrase of what consciousness is, is that it is the qualitative aspect of an experience. An

organism has conscious experiences if there is something that it is like to be that organism. This is the standard definition of consciousness put forth in the essay by T. Nagel, "What is it like to be a bat". We introduce the term qualia to refer to the general class of subjective components that accompany any experience, feelings of glee, joy, love, dejection, collapsing the wave function, etc. All of these are subjective and hence are left over once we subtract the fact that some physical act occurred. The hard problem of consciousness is to explain experience. For example if someone were to say that although you have determined the crystal structure of an amino acid, you have not answered the question of what it is like to be an amino acid or how is it an amino acid, they would be nonsensical. There is nothing that it is like to be an amino acid. However, one could say that while you have offered a physical account of an experience, you still have not explained what it is like to have that experience, that instinct would be correct because the experience is first person. Qualia are not physical things. There are many who would disagree that qualia are anything at all and really distinct from the physical. Herein lies the problem with consciousness: From where does the subjective component of an experience arise? Is it distinct from the physical stuff that causes the mental event. Is it reducible to something purely physical, that is, so many neurons firing, for example? If mental events have physical causes and effects, then why are we not able to strip away the mental talk and describe mental events purely in objective physical terms? Doesn't an objective description of a consciousness experience take you further away from what the experience actually is and hence isn't the physicalist program a non-starter? Many would argue yes. A positive answer here led Descartes to the dualism alternative. His argument was that when I am thinking, I am not necessarily aware of anything physical that is going on in my brain. Hence, there is no necessary connection between the physical and the mental, and Cartesian dualism was borne. Here is his argument in syllogistic form.

1.) When I think about my own knowledge as rigorously as I can, I realize that I cannot be completely sure that any physical thing or event exists.

2.) But I am completely sure that I think

3.) So, when I reflect as rigorously as possible, I am able to think of thought as distinct from any physical thing or event. (from 1 and 2)

4.) Two things that can be thought of as distinct under the conditions mentioned in 3 are essentially distinct - i.e. they are not identical, even though they may happen to always occur together.

5.) So a thought (any conscious state) is essentially distinct from a physical state. (from 3 and 4)

But dualism does not seem to be tenable. So the central problem in consciousness is that there seems to be some sort of disconnect (that is, an explanatory gap with no apparent bridge) between a physical description of a mental state and the actual mental state. That is, if I were to look in someone's head, I do not see a desire to eat ice cream or thoughts of Marxism or miniature pictures of the Mona Lisa. I see wiring, biochemistry, and living tissue. In this sense, the mind-body problem is fundamentally different from the stomach-digestion problem, or the lung-respiration problem or the internal-engine-combustion problem. No phenomenal talk is necessary to describe any of these problems. However, once we talk of consciousness, a physical description seems to be entirely insufficient. I will put off the subject of physicalism proper (that is mental states are identical to physical states and that is all there is to the problem) until next class and focus primarily on the computer model of mind—that is, functionalism.

Epiphenomenalism

This view purports that mental phenomena might be real alright but mental events cannot cause physical events. This really does not seem to correspond to our commonplace

notion of the relation between the mental and the physical. However, the epiphenomenalists do seem to have a point. On some level, one can argue that physical stuff has physical causes. In fact, the only facts that follow necessarily from physical facts are other physical facts. But what about cases in which mental desires and states seem to affect our behaviour. Consider, for example, the real but fuzzy notion of personal space. We have all encountered someone, typically when we travel to a foreign country, who we do not know all that well but stands uncomfortably close to us. Our natural response is to take a step back. But they take a step forward. This illustrates two things: 1) mental states seem to cause behaviour and 2) the same physical situation can give rise to vastly different mental states. By the latter I mean that in the situation discussed above, one person feels uncomfortable and the other person feels totally at ease when standing 1.5 ft. from a total stranger. This reinforces the subjective aspect of an experience.

Behaviourism

On this view, mental states are just patterns of behavior and dispositions to behavior. By behavior we mean bodily movements. As a result on this view, there is really nothing mental about the mental: there is only the physical. Language is just noise coming out of your mouth. Here again, there can be no causal relationship between physical states and mental states. We have the intuitive notion that our beliefs cause us to act in certain ways. For example, the belief that an umbrella keeps you from getting wet makes you use it to shield yourself from the rain. Behaviourists must deny that this sort of chain of events takes place. They would say that my belief that it is raining will be manifested in carrying an umbrella only if I have a desire to stay dry. My desire to stay dry will manifest itself in my carrying the umbrella only if I believe the umbrella will keep me dry. So to analyze why you took your umbrella, a behaviourist would have to bring in two subsidiary mental states. This process has no end.

Functionalism

Functionalism is a subset of physicalism that was quite popular in the '80s. Physicalism is the assertion that the world is entirely physical.

The key assertion is that the brain is a complicated computer program which generates lots of outputs. One of the outputs is simply conscious states. There are several ways in which the argument generally goes from here. One is the strong artificial intelligence stance. On this stance, it is argued that mental processes are identical to program processes of the computer. Hence, on this account, it is entirely possible to build a computer that thinks, has beliefs, has a sense of humor, and feels sad. Computer programs provide a list of formal commands. They provide a syntax. Do they provide a semantic? The answer is no. We associate with mental phenomena a semantic meaning. The following example of John Searle's illustrates the problem with the computer program view of mind. Consider that computer programmers have written a program that simulates the understanding of Chinese. By this I mean that questions are fed in and out pops reliably correct answers in Chinese. Let's assume that the computer's answers are as good as those of a native speaker of Chinese. The question arises does the computer understand Chinese in the way a native speaker of Chinese does? We can illustrate this with the following example. You are placed in a room in which there are a huge number of boxes containing Chinese characters. Let's assume you do not speak Chinese. Nonetheless, you are given a rule book for manipulating the Chinese characters. The instructions might say for example, take fish-shaped symbol out of basket 1 and place it next to sfsdfs-shaped symbol in basket 34 and so forth. Now suppose that someone keeps slipping cards with Chinese characters under the door of the room and your rule book is expanded to tell you what Chinese symbols you must assemble to slip back under the door. It turns out that the cards being slipped under the door to you are questions in Chinese and the instructions for assembling the Chinese characters generate

the answers to the questions. However, you do not know this. Now let's say you play in this room for quite some time and you become quite proficient in decoding the rules being slipped to you so that you can respond quite easily to the cards being slipped to you under the door. On the outside, it certainly looks like as if you know Chinese. The cards being slipped to you are all answered correctly with great rapidity. The question arises, do you understand Chinese. The answer is no. If the answer is no for you, then certainly the answer is no also for the digital computer. The reason is quite simple. The boxes of symbols are the database; the instructions being slipped to you are the computer program; the people giving you the questions and the answers are the programmers; and you are the computer. You are doing the assembling of the cards. Certainly no one would say that you know Chinese. Note you would pass a Turing test (that is, a passerby could not tell the difference between you and a native speaker). Likewise, one would also have to admit that neither does the computer. The reason is simple. You have no semantic knowledge. All you have is a syntax. The structure of this argument is quite simple: 1) programs are formal (syntactical) things, 2) minds have content (semantics) and 3) syntax is not sufficient for semantics. Hence, programs, however, elaborate are not minds. Note the question is not whether the whole system knows Chinese—the question is does the component that is you (or the computer) know Chinese. The answer is a resounding no. Hence, no amount of silicon chips put together by UI or MIT graduates can generate a computer that can have mental states. Here's another argument against the computational model of mind, also due to John Searle. Let us call those attributes of some x intrinsic attributes if they exist independent of an observer. All other features we refer to as being observer-relative. Consider a steel chair. Steel is an intrinsic property of the chair; however, that it is a chair is observer-relative. Consider computation. Is computation intrinsic or observer-relative? Let's say I am adding two numbers together. That I am adding two numbers together is a property of me independent of what anyone

else thinks. Hence, that addition property is intrinsic. Now consider a digital computer. Is IBM's Deep Blue intrinsically a digital computer? Can anything intrinsically be a digital computer? The answer is no. A process is computational only relative to some observer or user who assigns a computational interpretation to it. So now the question is rephrased, "can we assign a computational interpretation to the brain?" The answer is that we can assign a computational stance to anything that looks to us as if it is computing. Hence, the 'computer' model for mind is ambiguously formulated. It is circular at best. There does not seem to be any clear sense of what is intended when one says the mind is a computer. To quote Wittgenstein, "...machines cannot make claims ... of understanding." We can make such claims. Hence, there seems to be something obviously wrong with the functionalist approach in that it cannot account for the range of mental states we experience.