

Hitoshi NAGAI (Nihon University)

## Why Isn't Consciousness Real? (2)

### Day 2: Why Are We Zombies?

#### **The contrast between the phenomenal and the psychological is progressive.**

This lecture will be based on the problem raised by David Chalmers in *The Conscious Mind* (Oxford: Oxford University Press, 1996). I would like to develop the problem in what I think is the right direction, the direction we outlined in the first lecture. In my view, Chalmers should have developed his argument in this direction. In the course of this second lecture, however, details left out in the first lecture will be filled in by the use of his problem setting. The central theme will be the true meaning of the concept of a 'zombie'. I would like to begin by expounding Chalmers' argument.

Chalmers' argument starts from the distinction between the 'phenomenal' and 'psychological' concepts of a mind. It corresponds to the distinction between 'consciousness' and a 'mind' in the previous lecture. Put most simply, the distinction is that between the private and public aspects of a 'mind' in the broad sense of the word. Chalmers puts it as follows:

When we wonder whether somebody is having a color experience, we are not wondering whether they are receiving environmental stimulation and processing it in a certain way. We are wondering whether they are *experiencing* a color sensation, and this is a distinct question. (*The Conscious Mind*, p. 15, original italics)

By 'experience' Chalmers is referring to the 'phenomenal' aspect. All functions of a mind could have a role in causal relations even without being accompanied by an 'experience'. Yet they *are* accompanied by experiencing for some reason, and this, according to Chalmers, is the problem.

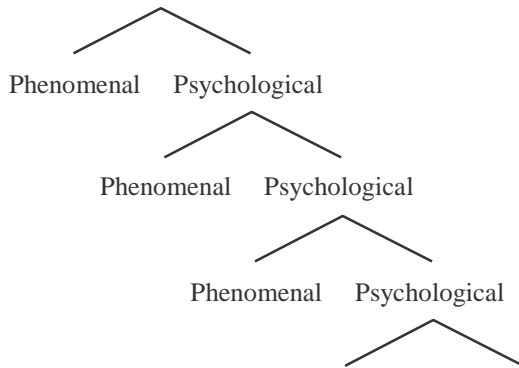
Take a 'pain', for example. What if this had all the functions of a pain, and yet lacked the phenomenal quality of pain? A person of whom this is true would be, as it were, a 'pain zombie'. He could have psychological pain but no phenomenal pain. (If we say that he feels psychological pain but no phenomenal pain, it would be that the word 'feel', too, has both the phenomenal and psychological meanings.) It is crucial to understand this contrast properly. Perception, for example, can be understood as a purely psychological process. For even a robot without consciousness could avoid obstacles to reach a destination, or perform

other actions by perceiving the environment. It could also obey the command 'Bring me a red piece of chalk'. Our perception, in contrast, is accompanied by phenomenal qualities. If a certain phenomenal quality is missing, the perceiver would be a zombie with respect to that quality. What Chalmers calls a 'zombie' is a human (or a human-like creature) that lacks all phenomenal qualities, thereby lacking consciousness itself. We will discuss the details later.

Concepts of mental phenomena can be divided into those for which the phenomenal aspect is essential and those for which the psychological aspect is essential. Sensations are among the former, and perception and thinking are among the latter. So, whereas sensations can dispense with psychological causal relations, perception and thinking can dispense with phenomenal qualities. But this only means that particular phenomenal qualities are not essential for there to be particular instances of perception or thinking, but not that perception or thinking can lack them altogether. There is no phenomenal quality peculiar to mentally calculating the sum of two and five, or the feel of '2+5', (or even if there were, it would not play an essential role). But it does not follow that the subject may lack consciousness.

I already have a small question at this stage. It leads to an immense problem. Chalmers discusses the contrast between the phenomenal and the psychological. Although this contrast might be actually vivid and valid for him, who is speaking of it, how does he already know that the same is true for those to whom he is speaking? That is, how does he know that there is the same contrast for others? Why is it presupposed that the contrast between the phenomenal and the psychological survives beyond the contrast between himself and others? I think that he fails to ask this most important question. It seems to me that the communication by language has transformed the phenomenal into the psychological, or into the phenomenal subsumed under psychological concepts. I suspect that the linguistic communication has downgraded Chalmers' contrast into a contrast between psychological concepts.

This can be viewed in a converse manner, which is more precise. The contrast is actually living and effective for me, who read and understood Chalmers, and precisely because of this, I cannot share with anyone else the contrast I grasp by myself. If so, Chalmers' own contrast, from which I learned the contrast, would have already been downgraded into a contrast between psychological concepts. This means that the contrast inevitably incorporates a progression with the following structure.



(This goes on endlessly.)

Only at the top row is the contrast actually living and operating. As it were, the rows below it, which repeat it by language, are merely its shadows in a cave. So although the contrast is living and effective for me, the content that I can communicate by talking like this now naturally falls into the second or lower row. It must be downgraded into a contrast between psychological concepts when communicated.

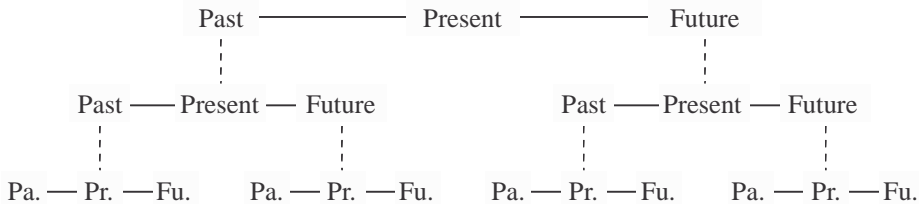
In my view, it is the essence of the contrast that it inevitably incorporates such a progress. I have said, 'It is crucial to understand this contrast properly'. But, in fact, it cannot be understood 'properly'. For it is indeterminable at which level the contrast is to be understood, and this gives the contrast an unstable structure. My view is that this progression of the contrast is precisely where the essence of the concept 'consciousness' is hidden. What is invented in order to lay all contrasts in the same plane is the general concept of a 'self', which I discussed in the previous lecture.

Before I said, 'It is crucial to understand this contrast properly', I said the following: 'If we say that he feels a psychological pain but no phenomenal pain, it would be that the word 'feel', too, has both the phenomenal and psychological meanings'. However, if the word 'feel' has both phenomenal and psychological meanings, is it not rather natural that the word 'phenomenal' itself has a double meaning? If so, is it not that the same is true of 'experience', 'consciousness', 'qualia' and so on?

What happens when reading Chalmers' discussion of the contrast in question is in fact the same as what happens when reading Descartes' doubt and his conclusion 'I think, therefore I am'. Descartes says that even if everything is doubted, the existence of the 'I' who is doubting cannot be doubted. Is Descartes' 'I think (or I doubt)' phenomenal? Or is it psychological? This question is of the same kind as the question I have just posed to Chalmers.

We are now naturally led to the analogy with time that we developed in the last lecture.

Below is a simplified version of a diagram I used in *The Opening: A Philosophy of Actuality* (*Philosophia OSAKA*, No. 3, 2008, p. 23).



(This goes on endlessly.)

Let us briefly explain the diagram. The ‘past’, ‘present’ and ‘future’ of the top row are actual. However, as we said in the previous lecture, there was a present time at any time in the past, or will be a present time at any time in the future, with a past and future centred at that present time. This is expressed by the rows below the top one, which are repeated *ad infinitum*. We have discussed the problem about this understanding of time.

Here let us analogize the ‘present’ with the ‘phenomenal’, and the ‘past’ and ‘future’ with the ‘psychological’. Then the top row in the first diagram corresponds to the top row in the above diagram. That the ‘phenomenal’ spoken of by Chalmers transforms into a class under psychological concepts corresponds to the ‘present’ dropping below the top row such that a present time in the past and a present time in the future also have to be ‘present’. In both cases, what is actual and absolute degenerates and transforms into what is possible and relative. To say that it ‘degenerates’ and ‘transforms’ does not mean that what happens here should be criticized, but rather that it should be noted.

### **What separates logical supervenience and natural supervenience**

We have presented the viewpoint from which we will critically reconstruct Chalmers’ theory. Before we proceed, however, there are still points to be introduced. In what follows, we will clarify what ‘supervenience’ means.

If, in all possible situations, the property *A* cannot be absent so long as there is the property *B*, the property *A* supervenes on the property *B*. To consider the world as a whole, global supervenience can be defined as follows: If there is no world in which the property *A* is absent and the property *B* exists, the property *A* supervenes on the property *B*. For example, if a world physically identical to our world also has to be biologically identical to our world,

biological properties supervene on physical properties. Thus, if consciousness supervenes on physical facts, two creatures in a strictly identical physical state will have a strictly identical conscious experience.

At this point, Chalmers draws an important distinction between logical supervenience and natural supervenience. If, in all logically possible situations, the property *A* exists so long as there is the property *B*, the property *A* logically supervenes on the property *B*. If, additionally, the property *A* exists so long as there is the property *B* in all naturally possible situations, the property *A* naturally supervenes on the property *B*. While natural possibility concerns what is possible within the restrictions of the laws of nature governing our world, logical possibility simply concerns what is logically possible with no such restrictions. Hence, there are plenty of things that are logically possible and naturally impossible, but nothing can be naturally possible and logically impossible.

According to Chalmers, biological properties logically supervene on physical properties. That is, no two worlds can be physically identical and biologically different. If photosynthesis is occurring in a world, the same photosynthesis must be occurring in another physically identical world. In addition, psychological properties also logically supervene on physical properties. No two worlds can be physically identical and psychologically different. If there is a world in which an organism is perceiving in a psychological sense, then in another physically identical world with that same organism in it, it would be perceiving too. However, according to Chalmers, the fact that there is a conscious experience in a world does not necessarily mean that there is a conscious experience in another physically identical world. Consciousness, unlike psychological properties, does not logically supervene on physical properties. The supervenience relationship between consciousness and physical facts is not a logical or conceptual relationship, but is a natural and contingent one. For if a psychological functional model explains how pain arises, for example, one can always pose the question why such a function accompanies *that* sensation of pain. (This will be important later when it is used as the grounds for the claim that a zombie does not actually exist but is logically possible.)

The difference between logical and natural supervenience may be clearer if we imagine God's creation of the world. If the property *A* logically supervenes on the property *B*, then God's creating a world in which there is the property *B* would thereby entail there being the property *A* in the world. But where God created a world with the property *B* where the property *A* only naturally supervenes on property *B*, God would have needed to additionally create a law to make the property *A* supervene on the property *B*. According to Chalmers, materialism is right if we take the property *B* as physical properties and the property *A* as all other properties, and if God need not have performed additional creation as in the former

case above.

Two objections could be made to Chalmers' view from two opposite directions. One objection is that consciousness logically supervenes on the physical. The other is that consciousness does not even naturally supervene on the physical. (The materialist could reply that it would be enough for materialism if consciousness naturally supervenes on the physical, but I do not see any significance in this objection.) I will simultaneously support these two opposite objections later.

What is most remarkable at this stage is that one can ask why any function accompanies *that* sensation of pain. Why is this question possible at all? What is 'that' sensation of pain like? How could one know that people feel 'that' pain? I cannot help but suspect that there is groundless conjecture here. I very much wonder why Chalmers and many other philosophers do not probe into this most essential point. In fact, this point is critically effective in the dispute over the possibility of zombies, covertly serving as the grounds of the arguments. In those disputes, too, the groundless generalization of 'that' plays an essential role. For unless it does, the general 'phenomenal pain' would not exist, and pain has to be assimilated to psychological pain, which is explainable in terms of its causal function. Thus the point here is crucial, and requires extremely careful handling.

### **Turning two-dimensional semantics into three-dimensional semantics**

What does it mean to simultaneously support the two opposite objections? To clarify this, we need to introduce another point put forward by Chalmers – i.e. his two-dimensional semantics.

Chalmers' argument is founded on Kripke's. Kripke drew a sharp distinction between the epistemological contrast between the *a priori* (i.e. knowable before empirical investigation) and the *a posteriori* (i.e. knowable by empirical investigation), on the one hand, and the metaphysical contrast between the necessary (i.e. that which cannot be otherwise) and the contingent (i.e. that which can be otherwise), on the other hand. This distinction is very important. On the face of it, the *a priori* coincides with the necessary whilst the *a posteriori* coincides with the contingent, and that had been thought to be the case.

According to Kripke, there are necessary truths that cannot be known *a priori*, examples of which are 'Water is H<sub>2</sub>O' and 'Heat is molecular motion'. They were known *a posteriori*, that is, as a result of empirical investigation, but once known, a reversal occurs (as discussed in the first lecture) such that water is H<sub>2</sub>O in all possible worlds; they become necessary truths. Then, the fact that water is *that* sort of clear and drinkable liquid, which fills lakes and rivers and falls from the sky, is demoted to a contingent fact, although it has been *a priori*.

*That* wateriness of water appearing to us becomes a property that the real essence of water, i.e. H<sub>2</sub>O, only *happens to* have.

The following point should be noted: what is meant by 'that' here, unlike in the case of the phenomenal qualities of pain, sourness, etc., is not private. It is not that individual persons describe their own private sensation using 'that', but that we, the inhabitants of this world, describe the way water is in *our own world* using 'that'. The privacy is not privacy among persons, but, as it were, privacy among worlds. We must not confuse them. (This concerns the contrast between primary and pre-primary intensions that we mentioned in the first lecture. We will discuss it in detail later.) Nevertheless, the demonstrative 'that' here is indispensable. For, after all, water is *that* kind of thing.

In the semantics of Kripke and Chalmers, the way the reference is fixed depends on whether we consider the actual world or counterfactual possible worlds. In other words, a concept has two kinds of intension. The primary intension, which is epistemological, is a relation that fixes the reference to the actual world, and is determined by the way the world actually is. The secondary intension, which is metaphysical, is a relation that fixes the reference to possible worlds. When the reference in the actual world is already determined, the secondary intension, by presupposing it, determines the reference in counterfactual worlds. In the case of 'water', its primary intension picks out *that* sort of clear and drinkable liquid that fills lakes and rivers. If it is discovered to be XYZ in the actual world, then 'water' refers to XYZ. But if it is discovered to be H<sub>2</sub>O, then 'water' refers to H<sub>2</sub>O. This is the secondary intension. Then it is no longer possible for water not to be H<sub>2</sub>O. A world in which there is no H<sub>2</sub>O is a world with no water, even if there is something that has *that* kind of watery appearance. However, the opposite was the case at the stage of the primary intension; water might not have been H<sub>2</sub>O. The liquid was clear and drinkable, and was grasped as *such a thing*, so it might not have been H<sub>2</sub>O. But once there is the second intension, a reversal occurs such that it becomes possible for water, which is H<sub>2</sub>O, not to be watery in *that* kind of manner. This is Kripke's (and Putnam's) account.

We briefly discussed the reason why this conversion occurs in the previous lecture. The reason, to put it more generally here, is that we have the desire to locate in the world the way the world appears to us, making it something that is not the world itself, or something contingent about the world itself. More precisely, the reason is that the device we developed and call language, in essence, is a mechanism that inevitably contains the tendency towards the same direction as that desire.

Now, according to Chalmers, because the primary intension, though not necessary, is *a priori*, it is determined independently of the way the world really is, which is known by investigation. Then, it is possible to bring the 'necessary-contingent' relation back to the

epistemological stage, superimposing it on the ‘*a priori-a posteriori*’ relation. For example, we can regard the fact that water is a clear and drinkable liquid as a fact about the actual world, and can regard different ways in which that fact can be discovered to be as possible worlds. This enables us to conceive of the possibility that the clear and drinkable water was not H<sub>2</sub>O but XYZ. That ‘water is *that* clear and drinkable liquid’ (or that ‘heat is *that* hot thing’) becomes an *a priori necessary* truth. It is not relevant to consider here how the actual world has really been discovered to be. Whatever way the actual world is discovered to be, it is an *a posteriori contingent* truth, because it is merely the way the world happens to be.

The above line of thought allows us to think as follows: if *that* water was actually discovered to be XYZ rather than H<sub>2</sub>O, water would be XYZ in that actual world, and so would be XYZ in all possible worlds conceived of in that actual world (which is a possible actual world). We become able to think of the relation between the actual world and possible worlds as a possible relation. That is the respect in which the present way of thinking is advantageous, but it is undeniable that, in another respect, it is a superficial grasp of the world. What is *a priori* in our knowledge, e.g. that ‘Murasaki Shikibu wrote *The Tale of Genji*’, *ipso facto* becomes necessary, so we will not be able to deal with the possibility that the person named Murasaki Shikibu had not written *The Tale of Genji*. This possibility corresponds to the possibility that H<sub>2</sub>O did not appear *that* way (i.e. in such a way that we could recognize it as water). Such things become simply impossible.

To sum up in Chalmers’ terms, it is logically possible but metaphysically impossible that water is not H<sub>2</sub>O, whereas it is logically impossible but metaphysically possible that water does not appear *that* way. (It seems to me that it would be precise and clear if we replace ‘logically’ with ‘epistemologically’.)

We can now link the two-dimensional semantics to supervenience. There is logical supervenience based on the primary intension and that based on the secondary intension. Chalmers denies the logical supervenience of consciousness on physical properties by treating it as based on the primary intension. We will discuss this in more detail in relation to the problem of zombies. (P. 60)

### **Things that do not supervene on the physical except ‘consciousness’: indexical facts**

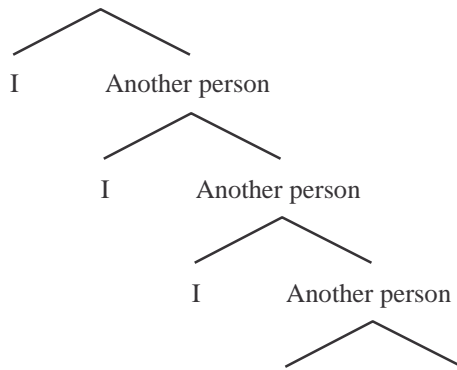
According to Chalmers, nearly everything except consciousness (i.e. phenomenal properties, experiences, or *qualia*) logically supervenes on the physical. If God creates the world physically, everything else, through supervening on it, automatically comes into existence. What Chalmers regards as candidates for things that do not logically supervene on the physical are: (1) conscious experiences, (2) indexical elements and (3) causality.



Before we proceed, I would like to explain (2) and (3). We will then return to the problem of consciousness, relating it to zombies.

Indexicals are such words as 'I', 'you', 'here' and 'tomorrow', whose references vary with the speaker and the context of the utterance. 'Here' means 'the place at which I am', and 'tomorrow' means 'the day after the one which exists now', so, ultimately, 'I' and 'now' are essential indexicals. But why do they not supervene on physical properties? If 'I' could be understood as 'the one who is making this utterance', namely as 'the reflectively conscious utterer', 'I' would supervene on physical properties. It would be that even a purely mechanical robot could utter the word 'I' and refer to the robot itself. The same would hold for 'now'. If 'now' could be understood as 'the time at which this utterance is made', 'now' would supervene on physical properties. Even a purely mechanical robot would be able to utter 'now' and refer to the time of the utterance. Chalmers, however, does not think this way. In fact, he has in mind the meaning of 'I' and 'now' which I explained by describing them as 'sole' and 'actual' in the first lecture, distinguishing it from the other meaning. (See *The Conscious Mind*, p. 85, ll. 6-8.) 'I' in this sense is 'I' as opposed to 'one reflectively conscious of oneself'. That is why he gives 'I' as an example of something that does not supervene on the physical.

Indeed, 'I' and 'now' in this meaning do not supervene on the physical. That is plainly true in the case of 'now'. A present fact ceases to be present and becomes past, remaining physically (or even in all respects) identical. The property of 'being now' does not supervene on anything! The same would be true of 'I', if the analogy we introduced in the last lecture holds. If there was a person who is physically identical with me (and is also even psychologically and phenomenally identical with me), he would not thereby become me – just as events that are precisely identical with those occurring now would not *ipso facto* create now. The property of 'being me' also does not supervene on anything. (Nevertheless, 'I' or 'now', by the same process as that by which the primary intension converts to the secondary intension, can convert to the meaning that supervenes on a particular person or event. We will discuss this process in the next lecture.) Below is a diagram that is parallel to the previous two diagrams.



(This goes on endlessly.)

Of course, only at the top row is the relation between the sole actual I and another person. The property of ‘being me’ at the top row alone does not supervene on the physical properties of the world. The I’s below the top row do not concern an actual fact, but are reducible to a formal property, i.e. self-relatedness. So those I’s can be regarded as supervening on physical properties. I can say, ‘I do not supervene on physical properties’. This means, ‘The fact that this person is me does not supervene on this person’s physical properties’. Nevertheless, this statement uttered by me would be understood by others in the meaning of the second row or below. Then they could say, ‘It does supervene on physical properties’. (The *statement*: ‘The fact that this event is happening now does not supervene on this event’s physical properties’, could be criticized in the same manner at other points in time.) On the other hand, *the same statement* as uttered by me could be uttered by any other person. Then I could say, ‘It does supervene on physical properties’. There is no objective fact as to which contrast is at the top row. Rather, interpretations of the above diagram will get caught up in the conflict it illustrates. That is, the property of ‘being at the top row’ in the diagram is itself relativized in the way that ‘I’ or ‘now’ illustrated by the diagram is. Since language begins with this relativization (or, since language *is* this relativization), there remains no trace of what is eliminated by the relativization inside the linguistic world. However, I, and consequently we, always live both what is constantly eliminated by the process of linguistic relativization and what is constantly generated by it. In fact, this is where the reason can be found as to why the term ‘phenomenal’ has to have a double meaning.

What is Chalmers’ view on the relationship between the indexical facts under consideration and consciousness? He does not think that there is an essential relationship between them. He only places the two on a par with each other. I, of course, think that there is an essential relationship. There should be no room left for misunderstanding, but put that way, my thought might seem to be this: whilst even a mechanical robot could refer to

itself or to the time of the reference, 'I' or 'now' describable as 'sole' and 'actual' cannot be referred to without consciousness. This is not my view at all. It is rather the opposite: unless there exist the indexical elements describable as 'sole' and 'actual', there could be no 'consciousness'. If consciousness is not understood this way, how could it fail to supervene on physical properties? One who grasps 'consciousness' as an objective fact, as Chalmers and most others do, should be able to sufficiently allow it to logically supervene on physical properties.

Chalmers himself writes as follows:

Most obviously, there is an epistemological problem about consciousness – the problem of other minds. This problem arises because it seems logically compatible with all the external evidence that beings around us are conscious, and it is logically compatible that they are not. We have no way to peek inside a dog's brain, for instance, and observe the presence or absence of conscious experience. ... [T]he mere *prima facie* existence of the problem is sufficient to defeat an epistemological argument ... for the logical supervenience of consciousness. By contrast, there is not even a *prima facie* problem of other biologies, or other economies. (*The Conscious Mind*, p. 74)

It is with regard to not just dogs' brains, but people's brains that there is no means to peek inside them for the purpose of observing the presence or absence of conscious experience. So it is question-begging to use the pronoun 'we' in presenting the problem. However, the problem could not be publicly presented in the first place without begging the question. This structure carries within it the very peculiarity of the problem. Chalmers says that there is at least one *prima facie* problem, but he is not right. It is *the only* problem. The 'problem of other minds' is not a problem that happened to arise from consciousness. Rather, what generates that problem is consciousness. To use the notion of the 'privacy of consciousness', consciousness is not something that happens to have the characteristic of being private, but the characteristic of privacy hypostasized is consciousness. However, the problem of privacy here cannot be presented as a problem of general privacy possessed by all creatures, or as a problem lying in one plane at the same level. As I said earlier, this itself is a truly baffling problem, and is also the reason why the character 'phenomenal', for example, has to be progressive. Thus, if it is said, 'Observation of the brain belonging to oneself would enable observation of the relationship between brain states and conscious states', there will be the aforementioned double meaning of 'oneself'. And that is the very problem.

### **Things that do not supervene on the physical except 'consciousness': causality**

Let us briefly discuss the third candidate of what does not logically supervene on the

physical. How could a world differ from our world if it is identical with our world in every detail of a microphysical fact? Thus far I have given two answers on Chalmers' behalf – i.e. absence of consciousnesses (or their being in a different way) and absence of myself (or my being a different conscious subject). I have contended that they are, in fact, not two separate matters. Now, the third candidate is causality.

It is true also of causality that no more than the regularity of the connections between events is observable externally. That is, there is a problem of causality corresponding to the 'problem of other minds'. If there are two worlds that are permanently identical in how all the particles in space-time are distributed, they may differ only in whether there is causality. A world without causation (which is, as it were, a causality-zombie world) is a world in which everything is in fact contingent. One might think that the existence of physical laws entails the reality of causation, but there being such patterns is compatible with everything being contingent. Moreover, the same problem arises as to the very existence of laws; it is possible for a world which is physically identical with our world throughout its spatiotemporal history to have a different set of laws. In Chalmers' comic example, that world has a physical law which will change two hundred tons of pure gold, if assembled in a vacuum, into lead. But since no such thing has happened in that world, it has followed the same history as our world. This clearly implies that our world could be such a world. Thus, laws of nature also do not logically supervene on a collection of individual physical facts. Of course, there is a big problem of whether it is possible to separate individual physical facts from laws of nature, a problem stemming from a conflict between Hume and Kant. I think as follows: to the extent that that is impossible – and only to that extent – would the existence of causality and laws of nature be presupposed.

Chalmers, however, says that although the existence of causality and laws defies reductive explanation, the problem is not as important as that of consciousness. Whereas consciousness is something elusive that demands an explanation, he says, causality and laws are mere postulates to account for the existing physical phenomena, that is, for the regularity existing in the nature. But if he can say this, he should also be able to say that consciousness, too, is only postulated to account for the regularity existing in the nature. Why can't he say this? I think that there is only one true reason. In the case of causality and laws, on the one hand, although there is a problem corresponding to the 'problem of other minds', there isn't the crucial asymmetry between self and other. In the case of consciousness, on the other hand, I can say, 'No matter what anyone says, I have consciousness, because I have this'. I can be certain of at least one instance even if no one else agrees with me, and whether other people also have *the same kind of thing* constitutes the 'problem of other minds'. In the case of causality and laws, however, there exists nothing that could correspond to the sole instance

whose certainty is guaranteed. As it were, everything corresponds to 'another mind' from the outset. To use the terms from the previous lecture, the second counterattack is directly made without the first counterattack. That is why it can be safely said that causality and laws are only postulated to account for the existing regularity. With respect to consciousness, there is actually one instance given of which such a thing cannot be said. Moreover, objective agreement can never be reached as to which that sole instance is. In that sense, there is a most unusual phenomenon that is incomparable with anything else.

If so, however, it is in fact possible to see the same structure in causality and laws. As regards causality, the sole instance for which I can say, 'No matter what anyone says, I know that this exists', would be the causality of free will. For example, I can raise my hand, let out my voice, and so on. As regards laws, it would be the private semantic rules of language which I follow when I speak. Then causality, laws, rules and meaning would each generate a contrast analogous to that between the phenomenal and the psychological, and so it would be possible to insist that that which corresponds to the 'phenomenal' fails to logically supervene on physical properties.

Whether something supervenes on physical properties indeed has no significance outside the range of problems associated with Chalmers and contemporary 'philosophy of mind'. A far bigger philosophical problem will be formed if we ask what the difference is between the idea of a zombie pertaining to consciousness or phenomenal qualities and the idea of a zombie pertaining to causality, laws, rules or meaning. Put another way, if I say, 'Whatever anyone says, I am not a zombie', the meaning of the validity of my certainty should be different between the case of consciousness and that of rules or meaning. But we will not go into this problem here in order not to deviate from the subject and obscure the continuity of the whole discussion. Hoping that my view on that problem will be clear from this series of lectures as a whole, I would like to go straight to the problem of zombies in its commonly understood sense.

### **Finally, here come the zombies!**

Chalmers argues that it is possible for there to be a creature that is physically identical with me but lacks conscious experience – i.e. my zombie duplicate. Here 'I' is presented as an example of something that is certainly conscious, so anything that is conscious will serve his official purpose. Therefore, a global expression of the same thought would be that it is possible for there to be a world that is physically identical with our world but has no consciousness in it at all. Naturally, all the creatures in that world are zombies.

Now I will exist in that world. Let us consider that person in the world in which there is

no consciousness, or my zombie duplicate in a world in which there are conscious creatures. He is precisely identical with me down to the level of molecules. He processes information about the external world in the same way as I do, and responds to stimuli in the same way as I do. For example, he seems to enjoy the taste of beer, distinguishing it from the tastes of other beverages, listens to language, and speaks. He is functionally identical with me. However, there is something crucial lacking in him; he has no inside. Conscious experience is completely absent. To employ the contrast between the phenomenal and the psychological, although there is in fact nothing 'phenomenal' for him, he and I are precisely identical psychologically. Therefore, it is indiscernible from the outside which of us is a zombie. He will also say, by the same mechanism as mine, that he is not a zombie. So, although he 'lacks consciousness' in the experiential, phenomenal sense, we can say that he 'has consciousness' in the functional, psychological sense.

There is the following objection: the conceivability of something does not entail the possibility of it. Chalmers, in response to this objection, appeals to the point which I said I would discuss 'in more detail in relation to the problem of zombies' (p. 54). His argument goes as follows. In the case of a necessary phenomenon discovered *a posteriori*, the objection is right. For example, there is no longer a possibility that water is not H<sub>2</sub>O, because water is H<sub>2</sub>O in all possible worlds. However, *a posteriori* necessity is irrelevant here. For consciousness is the primary intension rather than the secondary intension. Just as it is possible that water *qua* the primary intension is not H<sub>2</sub>O, so it is possible that consciousness does not supervene on the physical states inside the body. A zombie world is impossible in the way that it is impossible for water *qua* the secondary intension not to be H<sub>2</sub>O, but is possible in the way that it is possible for water *qua* the primary intension not to be H<sub>2</sub>O. Chalmers regards the problem of the logical supervenience of consciousness on physical properties as pertaining to the primary intension, thereby denying that logical supervenience.

The above defence of zombies based on two-dimensional semantics is indeed quite impressive and, as a form of argument, deserves emphasis. However, it does not stand. For either in the case of water or in the case of heat, the primary intension is not a phenomenal quality, or a *quale*. This should be obvious in the case of water. *That* clear and drinkable liquid that fills lakes and rivers and that sometimes falls from the sky is already an objective 'thing' that has been discerned from other things. Even 'hotness' in the case of heat, when it is the source of public investigation, is never a phenomenal quality, but is *that* property possessed by such things as fire, which is communally discernible by everyone. 'That' in this case does not mean 'that' phenomenal quality, which is meant by each person referring to their own private sensation, but is 'that' by which we, the inhabitants of this world, refer to the way water is in our own world. If 'that' here is the primary intension, the phenomenal

quality itself is, as it were, the pre-primary intension. Of course, the pre-primary intension inevitably accompanies the progressive structure. The primary intension must not be confused with this pre-primary intension. Even in the case of pain, sourness or yellowness, the primary intension, as the starting point of the investigation of the secondary intension, is never the pre-primary intension. Indeed, if it was, the very public investigation would be impossible.

In fact, the same is true of 'consciousness'. Unless cases where consciousness exists and cases where it does not exist are discernible from each other communally and objectively, the investigation of the secondary intension of 'consciousness' (i.e. that of its microphysical essence) could not have begun. Put in terms of the distinction between the phenomenal and the psychological, both the primary and secondary intensions cannot but be psychological from the outset. That is, 'consciousness', from the outset, cannot but be something that even a zombie can have.

If so, what should be said of the 'possibility of phenomenal zombies' advocated by Chalmers? To state my answer in advance, what he intends to say, in fact, bears on the pre-primary intension, and therefore *cannot be said* (unless the progressive structure is taken into consideration). In my view, this is essentially Wittgenstein's insight. It is utterly puzzling that this truly heart-warming insight seems to have been forgotten by everyone so quickly.

### **Objections from two opposite directions at the same time**

Do you remember when we said, in response to Chalmers' denial that consciousness logically supervenes on physical properties, that two objections can be made at the same time from two opposite directions? We can now easily carry out the task we have been postponing. The two objections from opposite directions were, 'Consciousness does logically supervene on the physical', and, 'Consciousness does not even naturally supervene on the physical'.

Let us begin with the objection that 'consciousness does logically supervene on the physical'. The truth of this statement is clear from the point that even the primary intension of 'consciousness' can only be psychological. In this sense, zombies are literally utterly impossible. The reason is simple: even before the formation of the microphysical secondary intension – that is, from the outset – our concept of 'consciousness' is and must be objectively determined by, for example, whether or not it is possible to partake in the game of 'losing consciousness and recovering consciousness'. Even a zombie could lose consciousness, say, by getting hit on the head, and recover consciousness afterwards. Therefore, a zombie has to have consciousness. A game of this sort is the only home of our concept 'consciousness'. The investigation of the secondary intension has to be conducted on the basis of such a game.

Furthermore, there would not occur the ‘first counterattack’, which we dealt with in the first lecture. For if there occurs a situation for ‘consciousness’ corresponding to the situation where ‘even though I had not eaten anything sour, my mouth suddenly became full of a sour taste (and my face looks as if I have just eaten something spicy)’, it would be a situation where I am conscious but cannot behave as though I am ‘conscious’ at all. So, it would be completely impossible for me to report that that situation has occurred. If the opposite situation occurs, it would be a situation where I suddenly become a zombie. Then I would not be able to say that *it* has occurred. Even if a zombie utters that it has, there must be no possibility that he has said so by referring to *it*.

Let us turn to the objection that ‘consciousness does not even naturally supervene on the physical’. Chalmers, presenting the ‘problem of other minds’, says that ‘the mere *prima facie* existence of the problem is sufficient to defeat an epistemological argument ... for the logical supervenience of consciousness’. However, if this is his argument, it would clearly apply to the case of the natural supervenience of consciousness. This argument should have the consequence that zombies are possible even naturally, let alone logically. That is, there may be a lot of zombies existing normally in this world, although it can *never* be known whether there are. The ‘problem of other minds’ is precisely this sort of problem, so it is simply puzzling that Chalmers, using the argument in question, believes without a doubt that other normal people in this world are conscious.

However, conversely, why could there be such an argument? If the home of the concept ‘consciousness’ resides in whether or not it is possible to partake in the game of ‘losing consciousness and recovering consciousness’, it would be impossible *a priori* that others, in so far as they partake in that game, lack consciousness. Moreover, in so far as the relevant neurophysiological processes are occurring in others, it would be necessarily impossible that they are not conscious. For both the primary and secondary intensions are functional and psychological. Chalmers might say that there will nevertheless remain the problem of whether others have ‘phenomenal’ consciousness. But what is ‘phenomenal consciousness’ that is neither the primary nor secondary intension? What does it mean to ask whether others have *it*? The intuition is presumably as follows. ‘I certainly have *this*, which is consciousness. But do other people also have anything *of this sort*?’ Yet it can never be known, by definition, whether other people also have anything ‘of this sort’. (If this can be known, those people would not be others.) The possibility of others being zombies is, in that sense, necessary.

However, if the problem assumes something that cannot be known by definition, is it not a pseudo-problem? If we start saying that ‘other people’ do or do not have something ‘of this sort’, are we not giving a logical tautology or contradictory statement disguised as an empirical factual statement? We are. Other people do *not* have anything ‘of this sort’. That



is why they are others. Then, are others zombies? In one sense, they *precisely* are. Isn't that right? There are people whose external behaviour and internal states of the brain and nerves are entirely normal, but they feel no pain, sourness, anxiety, or melancholy. Who are they? The answer to this riddle can only be 'other people'. This is a simple, indubitable fact. The home of the concept of zombies can only reside in the eerie contrast between me and people who are not me. However, who are 'other people'? Why can I, expecting approval, speak about *this* to you, other people?

The problem here, again, concerns the progressive multilayeredness of the self-other relation. In a sense, *everyone* can ask, reflecting on themselves, 'I certainly have consciousness, but do other people also have anything *of this sort*?' It would be that I have also just asked this question. *Just as* he is 'I' and she is 'I', so I am 'I', and there are others for each of 'he', 'she' and 'I'. In this case, too, it cannot be known, by definition, whether others also have something 'of this sort', so the possibility of others being zombies is necessary. However, in another sense, they are not *actual* others. 'I' issuing in 'I am "I"' is different in meaning from 'I' whom everyone is, therefore the meaning of 'others' would also be different. This vividly points to the relation between the progressive structure of the 'phenomenal' illustrated in the diagram on p. 49 and the progressive structure of 'I' illustrated in the diagram on p. 56. Thus, the home of the concept of a zombie, and therefore that of the concept of 'consciousness', resides in the progressive self-other structure.

(translated by Shogo SHIMIZU)

