

Hitoshi NAGAI (Chiba University)

## The Opening: A Philosophy of Actuality (1)

### Introduction

The point of departure for the following inquiry is this question: why is it that, rather than other human beings, this human being, and only this human being, is me, and why is it that I exist in this period, from the twentieth to the twenty-first century, and only in this period?

In the world, there have been, and are, a large number of human beings. However, it is only this human being who has the property of 'being me'. Other people each have the 'property of being me' too, but that only means that they all also have self-consciousness and call themselves 'I'. They are others after all. What is the origin of the property of 'actually being me', which is utterly different from the property of 'being me' that others have? What does it mean to say that there exists only one human being with that property?

As a consequence of evolution, human organisms emerged, and they have become self-conscious. That is what science can explain. But that explanation alone cannot explain why one of the organisms has to have the property of 'being me' in the sense just described. Science cannot explain why such a phenomenon exists.

In my childhood, when that question first struck me, I was overwhelmed by the wonder of 'existence'. I, whose existence is not necessary, started to exist as I do for some reason, and will someday cease to exist, or die. I remember that I felt it from the depth of my heart, and was seized by a sense of solitude as if I had been all alone in the eternal, infinite world, and by a staggering fear. The feeling I now have towards life has not changed much since then.

I am writing this introduction in the library. I do not know the person who is sitting next to me, reading a magazine. I could suppose that he is me. That is, I could suppose that he has the property of 'being me' in the special sense described above. Even on that supposition where the person is me, the person Hitoshi Nagai, sitting next to me, would be exactly as he is; he was born of the same parents as he actually was, at the same time on the same day, has the same genes as he actually does, followed the same life story as he actually did, and is writing this introduction precisely as he does. I will give the expression 'I!' to the sense of 'I' that can be used in such a supposition. By using this expression, the situation under consideration is described as one in which Hitoshi Nagai sits next to where I! sit. I could also think of a situation in which, although Hitoshi Nagai exists as he does, I! am not him, and I! do not exist in the world. However, I! actually exist, and I! am not reading a magazine but am writing an introduction. What is it that makes this fact obtain?

I will consider this problem mainly on the basis of the fruits of traditional philosophers, especially modern philosophers from Descartes to Kant. Contemporary philosophers tend to suppose that post-Fregean work on indexicality by J. Perry and others has destroyed the presuppositions about self which earlier philosophers took for granted. I do not think so. On the contrary, I think that most contemporary philosophers have forgotten their marvellous fruits that have posed life's deepest problem.

## Chapter 1 New Cartesian Meditations

### Section 1 Let Us Begin with the So-Called Other Mind Problem

#### **The Robot Problem**

Let us imagine a robot that looks just like a human being. He appears to think and have intention. He also has such human emotions as happiness and sorrow. Of course, he uses language, and when hurt, he grimaces as though in pain. He bleeds and sometimes even sheds tears. How does he differ from human beings? The answer is simple: he does not have a mind. He behaves as if he did, but inside him there does not exist that absolutely private and essentially internal element: the mind. Is this not a crucial difference between humans and robots?

Then, let us imagine that he is given a mind. Needless to say, no change in appearance results. He thinks, speaks and sometimes cries as he did before. There seems to be nothing unusual. But now, such behaviour of his is backed with an inner mind! This is a significant change. However, why does this change, which cannot be observed from the outside, seem meaningful even for those who are outside? Exactly the same problem arises in a case precisely opposite to this too, a case in which a human being's mind is removed. One night, a philosopher named K incurs the wrath of God, and his mind is pulled out. Seen from the outside, he has not changed at all in an observable way. When the morning comes, he takes breakfast with his family, and goes to the university to lecture on philosophy. No one notices the change that has happened to him. But, in fact, he no longer has a mind. That is, as it were, he is already dead. How can it be that this change seems meaningful, even though *no one* could know it?

This problem takes us in two directions that lead to what I think is the real problem of others. The first is this paradoxical situation: even though the characterisation 'others' is given from the outside, there necessarily exists a limit that absolutely refuses approach from the outside. We can say that all sources of the difficulty of the 'problem of others' reside in this situation. That others exist is nothing other than the insight reached from the outside that there exists a limit to what can be

reached from the outside. A change that cannot be noticed from the outside satisfies the definition of 'others'.

Of course, there might be a view that having one's mind removed is not a meaningful change and therefore is not a change in any sense at all. One who holds such a view, however, should imagine a case in which it happens to oneself. One night, you incur the wrath of God and your mind is removed. But seen from the outside, you have not changed at all in an observable way. When the morning comes, you take breakfast with your family, and . . . Even if *no one* notices the change that has happened to you, you, in fact, no longer exist. Therefore, it is merely for the sake of convenience that I have described the situation using 'you' as the subject of the sentences. The one who is having breakfast with your family is no longer you. He/She is only an imitator who mimics you and has taken over the personality that you once had. It should be quite obvious that a big change exists here. It is the change from your existing to your not existing, and if God had forewarned you of this situation a week before, you would have spent the week in the same dread as one who was about to die.<sup>1</sup>

If you acknowledge as meaningful the change that happens to you, then you must also acknowledge as meaningful the same change when it happens to others. The reason is simple: others should also be able to be a 'self' in the sense in question. It was on this assumption that the 'robot problem' was posed. But then, the problem becomes more curious. The original problem was posed concerning the difference between robots and persons who are others. Everyone understands the meaning of the problem in that way. Nevertheless, because of the fact that *everyone* understands it in that way and because that fact is tacitly assumed in posing the problem, what should be the solution creeps into the problem beforehand. In understanding the robot problem, you have to begin by adopting a perspective that views self and other equally. This structure is found not only in the 'robot problem', but throughout the area of the philosophical problem of others. It is only that its strangeness becomes prominent in the case of the problem of the existence of other minds, e.g. the 'robot problem', rather than the epistemic problem of them. And this is the second direction that leads to the real problem of others. Let us next consider this structure.

### **A Structure**

Expressions such as 'I' and 'self', when used in the philosophical problem of others are, in most cases, rendered a double meaning. It can be said as follows:

---

<sup>1</sup> We should note that the one who gives and removes mind has to be God or someone comparable to God. That condition is extremely important. An engineer, for example, is not capable of giving mind to a robot in this sense. That is so in principle, despite any imagined improvement in robotics. For however a robot reacts, it remains as a mystery whether or not it has a mind. Robotics can never understand what it is to bestow a mind in the sense considered here.

Others and self are different. One well knows things about oneself, but not much about others. What one thinks and feels now is known effortlessly. But what others think and feel is not known effortlessly. This asymmetry of others and self is a fact that everyone knows.

‘Self’ that occurs in the above passage means a self for everyone. It says that *for everyone* what one thinks and feels now is known effortlessly, whereas what others think and feel is not known effortlessly. In fact, the passage is concluded with the sentence, ‘This asymmetry of others and self is a fact that *everyone* knows’. But if so, something a bit strange occurs. Since most of ‘everyone’ are others, it follows that the writer of the passage has raised both the problem of others for others and the problem of others for himself/herself at the same time, and, *ipso facto*, some asymmetry of others and self—that is the problem of others—has already been tacitly passed over.

In fact, in posing the problem in the way suggested by the above passage, the original awareness of what was first felt to be a problem, i.e. what might be called a philosophical anxiety, is completely left out. How, then, should the original awareness of the problem or the philosophical anxiety be expressed? The following expression, for example, is at least better:

Others and I are different. I well know things about myself, but not much about others. What I think and feel now is known effortlessly. But what others think and feel is not known effortlessly.

The only difference is that the word ‘I’ is used this time. That difference cannot be said to prevent the problem of others from being tacitly passed over. This passage too has a double meaning. For it is always possible to question, ‘Who is that “I”?’ However, it is always possible to answer, ‘It refers to only one person, namely *this* I’, though, of course, there must always be a further question, ‘Who is that “*this* I”?’

Nevertheless, the fact that a further question must always be posed gives the problem of others a very special feature. That is, that very fact enables it to be formulated as a problem shared with other people, i.e. as an objective ‘problem’, but, because of that fact, the original awareness of the problem cannot help presenting itself in a distorted form in the formulated ‘problem’. You could have encountered the problem as an example of a philosophical puzzle in a textbook. But if you directly experienced the problem, it could not be a problem for everyone. Still, a formulated problem is already deformed into one for everyone. What is of importance is this deformation.

For example, Husserl’s *Cartesian Meditations* and Merleau-Ponty’s *Phenomenology of Perception* are almost consistently written in the singular first person. Let us cite examples firstly from the former and secondly from the latter.

In connexion with that and, indeed, motivated by it, there occurs a *universal superaddition of sense to my primordial world*, whereby the latter becomes the *appearance* ‘of’ a determinate ‘objective’ world, as the identical world for everyone, myself included.

Accordingly *the intrinsically first order* (the first ‘non-Ego’) is *the other Ego*.<sup>2</sup>

What is given and initially true, is a reflection open to the unreflective, the reflective assumption of the unreflective—and similarly there is given the tension of my experience towards another whose existence on the horizon of my life is beyond doubt, even when my knowledge of him is imperfect.<sup>3</sup>

Although we have had passages where the problem of others is dealt with, the essential point does not reside in their contents. Each of the claims, of course, has a certain meaning. However, the problem relates not to the truth of what is asserted but to a structure implicitly shown in the passages asserting it. To whom does ‘my’ in such sentences refer? Even if the authors each wrote the sentences having only themselves in mind, readers should reinterpret them as talking of ‘my’ for themselves (in so far as they do not intend to read Husserl or Merleau-Ponty’s autobiographical novel). Indeed the authors wrote, anticipating the readers taking such a position. The reinterpretation has always already begun. ‘My’ in such sentences is, therefore, already ‘my’ for everyone, and the problem discussed is tacitly *solved* in advance within sentences that discuss it.

It was in Descartes that this structure was plainly shown for the first time. In the third paragraph of ‘The Second Meditation’, where the methodic doubt reaches its climax, he writes as follows; introducing a deceiver of ‘supreme power and cunning’:

. . . let him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something. So after considering everything very thoroughly, I must finally conclude that this proposition, *I am, I exist*, is necessarily true whenever it is put forward by me or conceived in my mind.<sup>4</sup>

We can construe that ‘I’ in the first half and ‘I’ in the latter half have different functions. We can take ‘I’ in the first half to refer to Descartes himself for Descartes, whereas ‘I’ in the latter half to refer to a self for everyone. However, if it does not refer to Descartes generally (i.e. to Descartes for everyone), but refers to Descartes himself *for Descartes*, does that not mean that it refers to a self for everyone? Let us illustrate the difference between three kinds of referential act.

---

<sup>2</sup> Husserl, E. [1960], *Cartesian Meditations*, trans. D. Cairns, Martinus Nijhoff, §49, p. 108.

<sup>3</sup> Merleau-Ponty, M. [2002], *Phenomenology of Perception*, trans. C. Smith, Routledge, p. 41.

<sup>4</sup> Descartes, R. [1984], *Meditations on First Philosophy in The Philosophical Writings of Descartes*, trans. J. Cottingham, R. Stoothoff, and D. Murdoch, Cambridge University Press, vol. 2, p. 17.

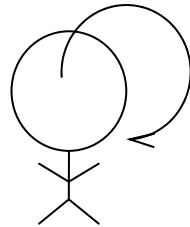


Figure 1

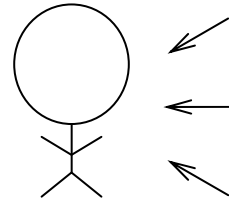


Figure 2

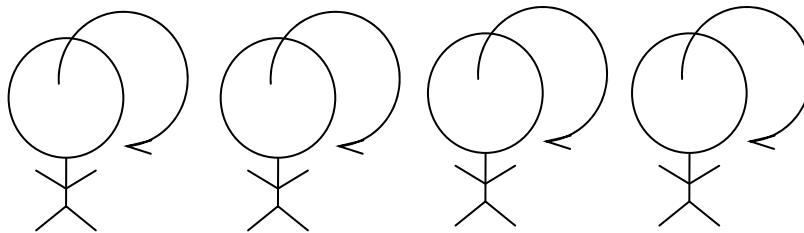


Figure 3

Figure 1 illustrates the referential act by which Descartes refers to himself. Figure 2 shows the act through which Descartes is referred to generally. Figure 3 shows the act through which a self for each person is referred to. The problem lies in the point that the referential act of Figure 1, when communicated to others (in this case, to people other than Descartes), could be reinterpreted as that of Figure 2 or Figure 3. In the first half of the passage above, Descartes obviously utters ‘I’ intending to perform the referential act of Figure 1. At that step of the methodic doubt, other human beings (or other minds) standing on a par with him no longer exist. Moreover, it should be that even a person who has personal identity and is referred to by the proper name ‘Descartes’ has already become what is subject to the doubt and has been abandoned. Therefore, even if there were, at this step, another person who maintained just the same as Descartes, Descartes could not have agreed with him/her. ‘I’ whose existence is indubitable is not ‘I’ in general, but must solely be ‘this I’, and should not even be a person named Descartes.<sup>5</sup> However, when Descartes says in the latter half of the passage that ‘this proposition *I am, I exist*, is necessarily true each time it is expressed by me, or conceived in my mind’, Figure 1, presumably, is already deformed into Figure 3. When attempting to communicate Cartesian meditations to others, this deformation, I think, is unavoidable, and I think that it is this deformability itself that makes the referential act of Figure 1 possible for the first time.<sup>6</sup>

<sup>5</sup> For that reason, Figure 1 (and, of course, Figure 3) is not strictly correct in that a person seems to be referred to. For this point, see Section 3.

<sup>6</sup> So please see Section 3 for details of this process too.

Nevertheless, there seems to exist an extremely peculiar fact such that if I intend to talk of it, I have to adhere to the 'I' of Figure 1. In my view, Descartes' meditations shine brilliantly in the history of philosophy precisely because they have revealed this fact. I believe that, unless the starting point is set at the existence of I! (who is neither a person who happens to be me, nor a so-called 'self' that everyone has) the depths of the problem of others cannot be plumbed. What should be noted here is that the notation I! does not express any sort of semantic claim on the referential act of the word. The notation has been introduced for the ontological claim on the structure of the world, and if Descartes, for example, had to adhere to the referential act of Figure 1, it was the result, but not the cause, of his attempting to refer to I!.

Then, what is 'I!'? In order to answer this question, I can no longer look for a clue in others' thoughts including Descartes'. That is because I! for Descartes is not I! for me, and therefore is not actually I! but only possibly I!. In what follows, just as Descartes does in the midst of the methodic doubt, I will talk ignoring the problem of communication with others (Section 2). The discussion will be communicated in a paradoxical way, and within what is communicated (and only within it), the mode of existence of others, i.e. the otherness of others, will be revealed (Section 3).

## Section 2 I! Exist

Merleau-Ponty, in the passage quoted above, says 'another whose existence on the horizon of my life is beyond doubt'. However, as already seen, that others exist means that there exists a limit that absolutely refuses approach from the outside, that is, the limit that can never appear on the 'horizon of my life'. Others refuse approach from the outside and can never appear on the 'horizon of my life' because each of them is *another origin* that opens a horizon of life different from mine. To grasp the otherness in that sense, in my view, there can be no other way than to pursue the course I will discuss in what follows; to put it briefly, the existence of others cannot be reached without passing through solipsism.

Hundreds of millions of years ago, the earth was a cold planet with no life. Presumably ruled by the laws of nature, organisms emerged on the surface of the planet, and some of them have developed to be conscious, and then to be self-conscious. Let us suppose that this whole process has been governed by the laws of nature. We could assume that the arrangement of substances which generates consciousness and self-consciousness was wholly determined by the laws of nature. And we could also assume that the brain states which generate conscious states, memories, and personality are also wholly determined by them. But, on this supposition, there is something that is never determined by the laws of nature. It is whether or not there is *this* I among individual persons, or minds, thus yielded and, if there is, which of them is or is not me. What the laws of nature can

determine is, at most, that a certain arrangement of substances yields a human being with certain properties. Whether or not I am among the human beings thus yielded and which of them is me cannot be determined by these laws.

Reference to the laws of nature is not necessary here. After a long, long time, the universe produced life, and then human beings. When it became the late twentieth century, I was born, and am here now. The problem is that, however detailed the description of this process, there is no place to be assigned for the fact that a human being who was born there at that time was me. Even if the universe had followed a historical progress precisely the same as it actually did, it was possible for a human being born of those parents there at that time and named 'Hitoshi Nagai' not to be me. But, for some reason, it was me. Why it was me and why a different human being was not me is a sheer mystery. What is clear is that it was possible for a historical progress precisely the same as the actual one, which yielded many human beings including Hitoshi Nagai, not to yield me, and therefore, it could have been possible for a different human being to be me.<sup>7</sup>

It is not necessary to refer to the history of the universe. Thinking of the present state of affairs alone would suffice. I exist now. But I did not exist a hundred years ago. Nor will I exist in a hundred years' time. The fact that this difference is clear means that the difference between the state of a world in which I exist and that of a world in which I do not exist is clear. If I do not exist, the world takes the form described in Figure 4. There only exists a number of people standing on a par with one another.

---

<sup>7</sup> S. Priest says as follows: 'That something is me is not an empirical fact. Being me is not just the fact that a particular being exists. Nor is being me any modal fact, for example, that a particular being is numerically distinct from all other beings, or that some being is self-identical. Even if some of these facts are necessary for something's being me, none is sufficient. There is always more to something's being me than something's being.' (Priest, S. [2000], *The Subject in Question*, Routledge, p. vii.) I entirely agree with this view. However, the critical problem is that two persons who entirely agree with what this literally says do not necessarily have the same view. My argument below is brought forward to show that point.



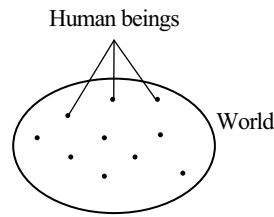


Figure 4

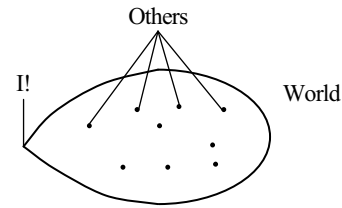


Figure 5

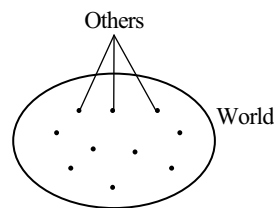


Figure 6

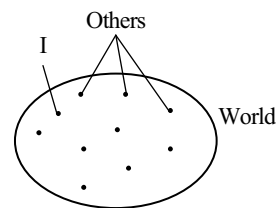


Figure 7

If I do exist, the world takes the form described in Figure 5.<sup>8</sup> (Figure 6 is a cross section of it seen from my viewpoint.) Figure 5 (or Figure 6), however, is transformed into Figure 7 for my being identified with Hitoshi Nagai, a human being in the world. (This transformation corresponds to the deformation of the referential act of Figure 1 into that of Figure 2.) Figure 6 and Figure 7 are two ways I represent the world.

The first thing that should be noticed here is that the fact that I was born and exist here and now is indeed a miraculous event. In the history of the human race, which has continued for a long time and will probably continue into the future, I, for some reason, exist in this period of time, and perhaps only in this period of time called the late twentieth century. It was possible for me to exist in the thirteenth or the twenty-third century, or not to exist at all. However, I was born and exist (only) here now. Is this fact not astonishing?

It should be noted that what is astonishing is neither the fact that there exists a human being composed of various properties in a particular way nor the fact that there exists the individual Hitoshi

<sup>8</sup> Figure 5 is a modified version of a figure in §5.6331 of Wittgenstein's *Tractatus Logico-Philosophicus*; it is modified such that it expels the eye to the limit of the world. §5.6331 and §5.633, according to the organization of *Tractatus*, are comments on §5.633, and §5.633 on §5.63 respectively. §5.63 is a proposition, 'I am my world'. The comment on it, to sum up, says only that the subject that thinks and entertains ideas, 'I' as the metaphysical subject, does not exist in the world, but manifests itself only in the fact that the world is my world. It is for this reason that the figure above is drawn with an explanation that 'the form of the visual field is surely not like this'. The form of the visual field is, of course, like Figure 6. And since the side view of Figure 6 is Figure 5, it should be right to say that Wittgensteinian solipsism is represented in Figure 5 and Figure 6. In Wittgensteinian solipsism, the 'world' and 'I' coincide in virtue of their both being 'my world', and 'I' here is none other than 'I!'.

Nagai as a substratum for the properties, who is a referent of a rigid designator. I am not saying that it is astonishing because such a thing might not have existed. This miraculousness can be represented spatially rather than temporally; why does only one human being, among a large number of human beings existing now, exist in the special way of being me? Why do I have to be this man? Here there is no necessity. I happened to be this man. The essential difference between Figures 5, 6, and 7 and Figure 4 is not whether or not Hitoshi Nagai exists. For it was, and is even now, possible that Hitoshi Nagai had not been me while he existed in exactly the same way as he actually did. The world in which Hitoshi Nagai exists and is merely one of the human beings in Figure 4 should be perfectly conceivable.

But how does that world differ from this actual world? Surprisingly, although such a world is for me obviously different from this actual world, this obvious difference is not at all a difference for anyone else other than I. I have said that it is *even now* possible that Hitoshi Nagai were not me while he exists in exactly the same way as he actually does. If that is true, it should be imaginable that he ceases to be me. At this moment, I vanish from the world. But Hitoshi Nagai still exists and keeps on thinking of this problem and writing this chapter. No change has happened to him from an objective viewpoint. His wife would come into his room and have a conversation with him without having any suspicion. How is it, then, from a subjective viewpoint? If 'subjective viewpoint' means the subjectivity of the person Hitoshi Nagai, it might still be said that nothing unusual has happened. For he will continue to write this chapter as he hitherto has. Nevertheless, he is not me any more.

I shall assign the notation 'I!' to 'I' in the sense discussed above in order to distinguish it both from the 'I' that refers to Hitoshi Nagai for himself and from the 'I' that means one's own self for everyone. It should be clear from what has been said that the fact that a person is 'I!' is independent of any property of the person. Let us suppose that there is an enormous volume that contains all objective facts in the whole universe that did, do, and will obtain. All of what are true of Hitoshi Nagai, including such subjective facts as his apparent memory and present bodily sensations, are written in this volume. But no description of I! would be found on any page of it. So reading it in the most careful way would not result in finding out who I! am. That is, it is meaningless to search for I! in the world of Figure 7. Only in the world of Figure 5 do I! exist!.

### Section 3 Depersonalised Self

However, the problem is that anyone who reads the above can understand and even agree with it, although I tried to say the contrary. Let us consider using more figures again. What does 'I!' mean if it does not mean Hitoshi Nagai as a person or self-consciousness for anyone? We can think as follows. 'Hitoshi Nagai' is a proper name for Hitoshi Nagai, and refers to the human being Hitoshi

Nagai. The first person pronoun 'I' used by Hitoshi Nagai is normally used in order (for Hitoshi Nagai himself) to refer to the human being Hitoshi Nagai. So, in this sense, in so far as there is the name 'Hitoshi Nagai', there need not be the pronoun 'I' in place of it. If everyone knows his name, Hitoshi Nagai would always call himself 'Hitoshi Nagai' and would have no trouble. This is the act of the arrows that point to the bodies in the figure.

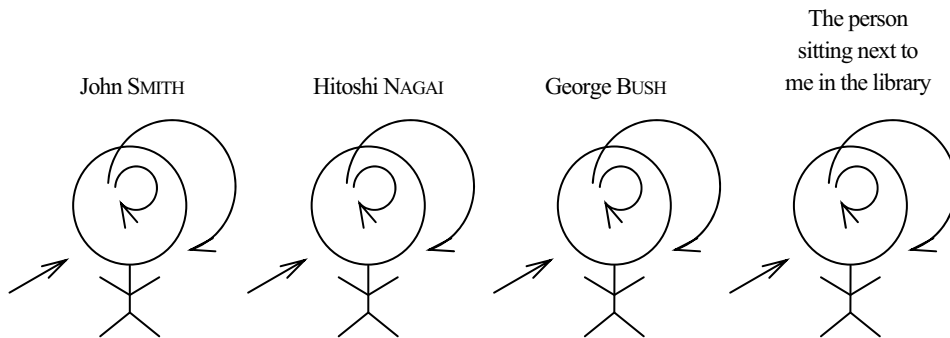


Figure 8

However, it is sometimes necessary for Hitoshi Nagai (or anyone else) to talk of himself in a way that is separate from (or independent of) his being Hitoshi Nagai. For example, he would have to answer the question, 'What do you want to be in your next life? Do you want to be a human being (or a man, or Japanese...) again?' This question presupposes that the one who will be born next time is certainly not Hitoshi Nagai. When it is answered, 'I want to be a Lion (or a woman, or an Arab) next time', 'I', of course, does not refer to Hitoshi Nagai. This is the act of the arrows that point inwardly within the heads in the figure. Let us call it 'depersonalised self-consciousness'. It is undeniable that *anyone* can have depersonalised self-consciousness.

The problem is in the next stage. The question I wanted to raise was this: do those two acts suffice? Does it follow that *this I* exists if a human being called Hitoshi Nagai exists with the function of depersonalised self-consciousness? It should not. It *is* conceivable that, even though Hitoshi Nagai exists and has the function of depersonalised self-consciousness, he is not *this me*. Therefore, in order that I exist, something extra not accounted for by these two facts should be necessary, and moreover, it could even be said that *that extra something is the opening of everything*. What is that extra something?

If there are two kinds of self-consciousness, i.e. personal self-consciousness and depersonalised self-consciousness, it follows that there are two kinds of thing referred to with the first person pronoun 'I'. When self-consciousness is not separated from a particular human being and is not depersonalised, 'I' refers to John Smith, Hitoshi Nagai or another particular human being. In this case, the first person pronoun is used instead of a name only when one refers to oneself. However,

when self-consciousness is separated from a particular human being and is depersonalised, 'I' refers not to John Smith, Hitoshi Nagai or another particular human being but to a depersonalised self. That is, the first person pronoun is no longer used instead of a name.

Now, the problem I raised was this: does it follow that I exist if there are both of those two? The person Hitoshi Nagai exists, and is writing this chapter. Due to the contents of the chapter, he sometimes imagines, 'What if I were not Hitoshi Nagai?' That means that he has the function of depersonalised self-consciousness. Then, do those facts make him me? Of course not! They do not suffice. That Hitoshi Nagai exists and has the function of depersonalised self-consciousness does not make it necessary that he is me. In order for him to be me, something needs to be added, and furthermore, it could even be said that what is added is *the opening of everything*. I take that opening to be something miraculous.

To reinforce this argument, it might serve to consider the problem of a 'person just like me'. Suppose there is a human being perfectly identical to me, i.e. a person with all the same properties as me. Is he me? There could be many ways this problem could be discussed, but here I will consider a human split. I come into a coffee shop, and decide to sit in the left of two chairs placed back to back with each other. Suddenly, my body splits into two, and the other one sits in the right hand chair. He and I are completely identical both mentally and physically. Not only with respect to how we look, but with respect to everything we remember and feel, we are wholly alike, except with respect to where we sit (and things related to it such as what we see). Nevertheless, he and I are fundamentally different. For one is me, and the other is not me. No difference can be as fundamental as this. If such a split happens to another person, i.e. if the two are someone other than me, it would be very difficult to distinguish between them. However, in my case, the distinction is most clear. Now, what is it that *makes* one of the two me and *makes* the other one not me? Of course, a miracle is necessary here again.

But that argument, too, can be understood and even agreed on by everyone. Why is this?

I would like to consider the four people in Figure 8 again. I tried to think of Hitoshi Nagai's being me in terms of something extra, which is neither his being Hitoshi Nagai nor his being self-conscious. Yet it seems that the fact that *everyone* is one's own self means that everyone has this extra component.

Thinking carefully, I have to admit that that is a matter of course. Suppose you are separated—i.e. depersonalised—from a particular person such as John Smith. What would you depend on in order to find out which of the selves (illustrated as what are pointed to inside the heads of the four in Figure 8) yours is? It should be clear that being a particular person and being a depersonalised self do not suffice. For being depersonalised means not being a particular person, and being a depersonalised self is a property common to everyone.

What, then, would you depend on in order to *find out* which self yours is? Presumably, the

answer is simple: you would depend on the fact ‘it is *this* me’. And that would be true for everyone. Here, the problem of how to distinguish *this I* from *other Is* (who are not *this I*) should not even arise. The distinction could be made extremely easily. It is, as it were, a matter of course that, *for everyone*, one of the selves is oneself. In fact, what is difficult is to do the converse, namely to situate oneself in a position of placing your own self and another self on the same plane (or in Figure 8, it is for one of the selves to view all of the four equally).

The new problem that has come up here is that of the ‘individuation of the depersonalised self’. It could be stated as follows. For a self to be depersonalised is to cease to be a particular human being. (In Figure 8, it is a case where the bodies disappear and only the arrows inside them remain.) In that case, how are the selves (or the arrows) distinguished from each other? Since each of them already lack distinctive features and only have the property of being self, it is impossible to distinguish them objectively by viewing them from the outside. It is only possible to distinguish between ‘my self’ and ‘other selves’. It is impossible to distinguish between other selves. Individuation of depersonalised selves who do not have any continuing outward features, blood types, personal memory or any other virtue by which persons are distinguished, could only be accomplished in that way.

The above argument suggests that I have been talking of a general concept of an individuated depersonalised self when I thought I was talking of I!, which is absolutely unique. In trying to raise the problem of the existence of I!, it seems to deform into that of individuation of a depersonalised self. When this argument is read or heard, the problem must be initially taken into the dimension of the individuated depersonalised self. Only in that way could this problem be transmitted to others. If so, the difference between the problem I wanted to raise and what is transmitted has to be inquired next.

‘Individuated depersonalised self’ is only a concept. So there can exist any number of instances that this concept applies to. But my existence is not a mere instance of the concept. Anyone can be *an* instance. How can I express the difference between a mere instance and me—the most radical difference?<sup>9</sup>

#### Section 4 What Is Mind?

##### **Applying Theory of Mind to Oneself**

Before we proceed, let us make a rough sketch of how the concept of ‘mind’ is formed—a sketch whose real meaning will be understood only after reading through Chapter 2.

The settlement of the other mind problem resides not in arguing that I can regard others as

---

<sup>9</sup> This will be discussed in Chapter 2.

having minds as I do, but, conversely, in arguing that I must place my own mind, and everyone must place his/her own mind, on an equal footing with numerous minds (as a mere instance of mind).

There is this well-known question: ‘Suppose a child puts chocolate in a drawer, and while she is playing outside, her mother puts it in a cupboard in the kitchen, where do you think she will, after coming back home, look for the chocolate?’ In this case, the child’s belief disagrees with the objective fact. But in order to be able to understand that her belief disagrees with the objective fact, you must be able to understand ‘another person’—in this case, the child—as a being with mind (i.e. a being who *represents* the world). That is to acquire a ‘theory of mind’. Humans are said to acquire it about the age of four.

However, the really important task would be to grasp oneself as one example of just ‘another person’. Unless one understands what one believes as only a way one human being represents the world, instead of taking it to be the way the world itself is, a ‘theory of mind’ is not completed.

The same thing can be said when we take into consideration the argument on the relationship between mind and brain. Whereas I can observe the relationship between my consciousness and my brain (as it appears in the consciousness), I cannot observe the relationship between another brain (which can also appear in my consciousness) and another consciousness (which cannot appear in it). In fact, if there is a peculiarity in the mind-brain problem, it is because it cannot help bearing on the other mind problem.<sup>10</sup> Do we not name the position that the asymmetry between self and other manifests itself ‘mind’—or ‘consciousness’? Or rather, is it not that consciousness is an entity whose existence this asymmetry requires and necessitates?<sup>11</sup>

### **Mind-Brain Relationship and the Malicious Brain**

The concept of ‘brain’ has a double meaning, and therefore the ‘mind-brain relationship’ has two meanings. One is the relationship between a brain whose function *is* the opening of the world—e.g. a brain in a vat that creates the world itself—and the world it opens. And the other is the relationship between brains in the world and the consciousnesses they produce. Assimilation of the former relationship into the latter relationship makes up the so-called mind-brain problem. It is because the former is projected onto the latter that we feel a philosophical problem in it. Only in this way does a metaphysical flavour in the philosophical mind-brain problem remain. So the essence of the problem is missed when the problem is alleged to be solved, and the problem is never solved when we cling onto the problematic of the problem.

---

<sup>10</sup> The essence of the hardness of Chalmers’s ‘hard problem’ and the essence of the mysteriousness of McGinn’s ‘mystery’, in fact, lie only here. Indeed, the ultimate ground for their argument resides here.

<sup>11</sup> The reason that this asymmetry will always remain as asymmetry, i.e. that it will not become symmetrical, is that I! exist ultimately. This problem will be discussed in detail in Chapter 2.

Let us suppose that a self-appointed scientist calling himself ‘the missionary of brain’ turns up with a vat and says, ‘It is the brain in this vat that creates your world’. And when he fiddles with the brain, I feel a sharp pain in my left toe, or the world becomes entirely red, or a huge dinosaur emerges.<sup>12</sup> Now, should I believe what he says?

I may have no choice but to believe it. Of course, there is something strange in my believing it. For the structure here is twisted such that the outside of a dream intrudes into the inside of a dream to tell the truth, the truth being that it is a dream. Since it is told in a dream, that is also a dream after all. Does that not mean that the brain in the vat I see cannot be the real brain in a vat that creates it?

However, in the case of brain, this twist, namely the secret art of the Incarnation, is not at all a secret art, but only a very common scientific fact. For it is, in principle, possible for me to perceive my actual brain in my perceptual world, and it is, in principle, possible for me to fiddle with it and cause a sharp pain in my left toe, or to make the world entirely red, or to make a huge dinosaur emerge. If I were a brain scientist, those things would be actually possible to some degree.

If so, however, it would also be possible, by stimulating the brain, to alter a perceptual image of the brain itself. It would be possible, by stimulating it, to make it seem as if there were no such brain, or to make myself believe so, or even to make myself convinced that a different (counterfeit) brain is the brain in question.

Then, what if the brain has a mechanism that produces such untrue states of affairs *from the beginning*? That is, what if it is a ‘deceiving brain’ from the beginning? Is ‘malicious brain doubt’—as opposed to ‘malicious demon doubt’—possible? Is it possible for me to say to the malicious brain, as Descartes’ ‘I’ did in confronting the malicious demon’s deception, ‘let him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something’, even though the malicious brain produces my consciousness itself?

If I were a brain scientist, I could, by observing my own brain, research the relationships between brain states and conscious states. However, since the brain states are perceived, they are part of the conscious states. I would never reach the brain states themselves which are producing the perceptual states. The malicious brain problem is a device to exaggerate the problematic of this situation; it is not a mere joke or a sort of invented paradox, but an actually existing problem.

The problem is reversed when others’ mind-brain relationship is in question. If I were a brain scientist, I would be able to, by observing another person’s brain, research the relationships between brain states and conscious states. However, I cannot directly perceive another person’s conscious states. His/Her conscious states which I can perceive would have to be present physically, or behaviourally, as his/her brain is. I would never reach the conscious states themselves which the brain states are producing. But that very fact enables me to perceive the brain itself. In other words, I

---

<sup>12</sup> So the ‘missionary’ might have said, ‘It is the brain in this vat that creates *the* world’. But I shall set this point aside for the time being.

perceive the brain itself even if it is a malicious brain.

Both in my and another person's case, I cannot juxtapose mind and brain to make them equally observable objects on the same plain. It is this discordance that makes the mind-brain relationship one of the biggest philosophical problems. (If I were the only person in the world, this problem would not arise.)

### **The Robot Problem Revisited—A Passage to Chapter Two**

Now, let us consider the robot problem from a different point of view. Suppose that one night I incur the wrath of God, and my mind is removed, would I be a kind of robot—a phenomenal zombie—then? Not necessarily. For it would be sufficient just if the person who was me ceases to be me. The person who ceases to be me can continue to have a mind (which is, of course, his). He can have his self-consciousness. It would only be that he becomes another person or a mere person (who is not me). The world becomes a centreless one which merely has a lot of minds and self-consciousnesses.<sup>13</sup>

If so, that kind of thing could never happen to others. For they are not me from the outset. However, when it is thought, through some ingenious device, that that also is possible, the concept of mind will become possible. This is probably the origin of the concept.

In any case, the God who is capable of such a deed is of a higher order than a God who is capable of removing someone's mind. The ordinary God, whilst he can penetrate minds of all persons, knowing everything about their thoughts and feelings, does not know which of the numerous persons in the world I am. Neither does he know which of the numerous times now is. Although he can see into all persons' minds, he does not know which of them I am. So although the ordinary God can remove a mind, he cannot remove from a person the fact that the person is me. Neither can he remove from a time the fact that the time is now. Of course, this kind of God would not have the ability to create me.

Can we imagine a God who is capable of such a weird deed? Even God cannot do such a thing; that could have been Kant's insight. On the other hand, Leibniz can be interpreted to have said that it is a possible thing for God to do.

---

<sup>13</sup> Of course, another person or a mere person (who is not me) could also think of what is exactly of the same form as that, as a phenomenal zombie could also judge that he is conscious. In fact, the latter is the situation resultant from generalising the former in some way. Even if the unconceptualisable (or unfunctionalisable) evolves from 'I' into 'consciousness', it remains true that the concept of 'unconceptualisable' (or the function of unfunctionalisability) has to be invented. The same is expressed as follows in the Leibniz's terms in Chapter 2: although the distinction between God's will and God's intelligence is essentially different from distinctions made within intelligence, it can be understood by intelligence, and when it is understood, it is assimilated into distinctions made within intelligence.



Chapter 2 The Leibnizian Principle and the Kantian Principle

Chapter 3 What Is the Impossible Private Language?

(translated by Shogo SHIMIZU)