Mind Mach
DOI 10.1007/s11023-006-9045-z

The problem of the many minds

Bradley Monton · Sanford Goldberg

Received: 24 May 2005 / Accepted: 11 September 2006

Abstract It is argued that, given certain reasonable premises, an infinite number of qualitatively identical but numerically distinct minds exist per functioning brain. The three main premises are (1) mental properties supervene on brain properties; (2) the universe is composed of particles with nonzero extension; and (3) each particle is composed of continuum many point-sized bits of particle-stuff, and these points of particlestuff persist through time.

Keywords Mind · Functionalism · Supervenience · Individuation

Introduction

The present paper proposes a new philosophical problem: *the Problem of the Many Minds*. (As will become clear in due time, this problem is distinct from the Problem of the Many.) Our problem concerns the number of minds which exist per ordinary functioning brain. It is normally thought that, at least in standard cases (e.g., not involving multiple personality disorder), at most one mind is associated with a brain as the particles which compose that brain evolve through time. It will be argued here, however, that when a brain is functioning in a normal, awake state, it is actually the case that an infinite number of qualitatively identical but numerically distinct minds

Thanks to the audience at an informal talk given at Princeton University, which included Allison Dawe, Cian Dorr, Benj Hellie, Brian Kierland, Aaron Konopasky, and Chad Mohler. For helpful comments on a previous version of this paper, we thank David Lewis. And thanks to an anonymous referee of *Minds and Machines*.

B. Monton
Department of Philosophy, University of Colorado, Boulder, CO 80309-0232, USA
e-mail: Bradley.Monton@colorado.edu

S. Goldberg (✉)
Department of Philosophy, University of Kentucky, 1427 Patterson Office Tower, Lexington, KY 40506-0027, USA
e-mail: scgold@uky.edu

are associated with that brain. At least, it will be argued that this is so given certain premises presented below. Since the conclusion is somewhat unbelievable, one should think seriously about rejecting a premise. However, none of the premises clearly deserves rejection.

The problem

There are three main premises in the argument. First, mental properties supervene on brain properties. (Perhaps this is because physicalism is true, or perhaps because naturalistic dualism is true; it does not matter for present purposes.) Second, the universe is composed of particles with nonzero extension. And third, each particle is composed of continuum-many point-sized bits of particle-stuff, and these points of particle-stuff persist through time. A more natural term for these points of particle-stuff might be 'point masses', but it is preferable to remain agnostic about whether these points of particle-stuff have any properties besides position. The third premise presents the most natural way of understanding the ontology of a particle with extension. The particle is made up of *something*; the generic term 'particle-stuff' is used to designate whatever this something is. Since there are continuum-many points inside a particle with extension, it is natural to assume that a particle is composed of continuum-many points of particle-stuff.

The argument for the thesis that an infinite number of minds exist per functioning brain is as follows. Assume that the above three premises are true of the actual world. Now consider a *possible* world in which the simples of the universe are *point particles*; that is, in this possible world the simples are particles each of which is located at only a single zero-dimensional point in space. Suppose that there are the same number of point particles in the possible world as there are particles in the actual world. Suppose further that for all time the particles evolve in the same way, in that the distance relations between the point particles in the possible world are always the same as the distance relations between the centers of mass of the particles in the actual world. Finally, suppose that while particles in the actual world have properties like mass and charge, these properties do not apply to the point particles. The only property the point particles have is position (and derivatively, instantaneous velocity).

We submit that mental properties are tokened (instantiated in the brain described) in this possible world exactly when mental properties are tokened in the actual world (instantiated in the brain with which we started the discussion). This claim follows naturally from a commonly accepted (functionalist) view in philosophy of mind, to the effect that mental properties are functional properties: they are individuated in terms of their instantiations' typical causes, causal relations to other mental property instantiations, and typical effects. Such a view, together with the fact that the point particles in the possible world occupy the same causal roles as the particles in the actual world, entail that tokens of the same mental property are realized in each world.

Now consider a single particle of the brain of a person in the actual world. Consider a point of particle-stuff located at some particular distance and in some particular direction from the center of mass of this particle. Taking the other particles into consideration, we see that there is a point of particle-stuff at this location

for each particle in the brain. As the particles in the brain evolve, these points of particle-stuff evolve as well. In fact, the points of particle-stuff evolve in about the same way as the point particles do in the possible world considered above.

(It could be that the points of particle-stuff evolve in exactly the same way as the point particles do in the possible world. Or it could be the case that there are currents of particle-stuff within a particle in the actual world, in such a way that the point of particle-stuff that at one time is at the center of mass (for example) of the particle evolves such that it is no longer at the center of mass. Even if the point of particle-stuff did evolve in that way, though, it would not go very far relative to the point particles in the possible world, because all the particles in the actual world have very small diameters. The movements of particles in the brain, which are relevant for the production of a mental state are, presumably, much larger than the diameters of the particles themselves.)

Let N be the number of particles of the brain evolving in the actual world. Consider a mereological sum of N points of particle-stuff, one point of particle-stuff per particle. These N points of particle-stuff evolve in about the same way as the point particles do in the possible world. It follows that (and here is the important move) just as various mental properties are realized by the N point particles as they evolve in the possible world, so too various mental properties are realized by the N points of particle-stuff as they evolve in the actual world.

So far we have considered just one mereological sum of points of particle-stuff located within the particles of a brain. But since there are continuum-many points of particle-stuff within each particle, there are continuum-many disjoint mereological sums of N points of particle-stuff, with one point of particle-stuff per particle. Hence, it would appear that continuum-many minds, each constituted by a distinct set of mental properties (one corresponding to each of the disjoint mereological sums just described), supervene on a single brain. Each of these minds tokens precisely the same mental properties (and in the same order) as the others, since each mereological sum of points of particle-stuff is evolving in about the same way. What is more, if qualitative states supervene on functional states, then each mind is *qualitatively* identical as well. However, the minds are numerically distinct, since the mental properties that constitute them are realized by different disjoint systems in the physical world.[1]

First objection

Some might say that there cannot exist numerically distinct minds that are nevertheless identical in the respects mentioned above. On this objection, minds cannot be judged numerically distinct on the basis of the physical arrangements that realize the mental properties constituting the minds. Instead, the objection continues, minds are to be individuated, and hence (since it would follow from such individuation) judged numerically distinct, on the basis of the mental properties that realize them. It

[1] We are agnostic about the question of whether numerically distinct minds can supervene on non-disjoint but non-identical mereological sums of points of particle-stuff. If one had a solution to Peter Unger's (1980) Problem of the Many, then one would be more likely to have an answer to this question. It's perhaps worth mentioning, though, that the Problem of the Many Minds is not the Problem of the Many applied to minds—the Problem of the Many Minds focuses on disjoint mereological sums, while the Problem of the Many applies only to non-disjoint mereological sums.

follows that if minds A and B are constituted by the same mental properties (instantiated in the same order), then A is numerically identical with B.

This objection can be met. Consider Fred and Ralph. Both enjoy daydreaming about deserted tropical islands. Suppose that there happens to be some interval of time where Fred and Ralph daydream in exactly the same way. Suppose that all that either of them is conscious of is his own respective daydream, so during the daydream Fred's mind and Ralph's mind are functionally (and, we can add for good measure, qualitatively) identical. Now everyone grants that before Fred and Ralph have the same daydream, Fred and Ralph have numerically distinct minds. Similarly, after the daydream, when Fred and Ralph again have different conscious experiences, Fred and Ralph have numerically distinct minds. It is counterintuitive to say that, during the daydream, one mind disappears, and reappears after the daydream is over. We standardly think that a mind is unique to a person, regardless of which other people exist in the world.

Second objection

Some might say that a mind could not be constituted by mental properties which themselves are realized by a mereological sum of points of particle-stuff in the actual world, since the requisite *causal relations* do not hold between the various points of particle-stuff. For example, a supporter of causal role functionalism would say that, in order for a mind to be constituted by mental properties realized in some particular physical system, the appropriate causal relations must hold between the various subsystems of that physical system. It is reasonable to hold that the causal relations between the physical subsystems supervene on the causal relations between the atomic components of the physical subsystems. The proponent of this objection would say that the appropriate causal relations hold between the N particles, which compose the brain, but not between the N points of particle-stuff.

In reply, it can be maintained that the same types of causal relations hold between the N particles as do between the N points of particle-stuff. Consider a situation where particle A is a cause of particle B's changing its direction. Then we can ask: is each point of particle-stuff in particle A a cause of each point of particle-stuff in particle B changing its direction? We maintain that the answer is *yes*. It is well known that there is no uncontroversial analysis of causation, but counterfactual analyses of causation (e.g., Lewis, 1986) are popular so we will rely on such analyses. Consider the event of a particular point of particle-stuff in particle A moving toward particle B, just before particle A deflects particle B: had that event not occurred, would the points of particle-stuff in particle B still have been deflected? Presumably the answer is *no*. Particle A is atomic; it is physically impossible to divide it into smaller particles. So in the closest possible world where the point of particle-stuff in particle A moves in a different direction, so it does not interact with particle B, *all* the points of particle-stuff in particle A move in that direction. Thus, in that possible world, particle A does not interact with particle B, so particle B is not deflected. It follows that the point of particle-stuff in particle A enters into the same types of causal relations as particle A, and thus the same types of causal relations hold between particles as do between points of particle-stuff.

Third objection

Some might say that a mind could not be constituted by mental properties which themselves are realized by a mereological sum of points of particle-stuff in the actual world, since these points of particle-stuff do not exhibit the requisite *void* between them. They would grant that a mind supervenes on the point particles in the possible world. They would argue, however, that there is an important difference between the point particles in the possible world and the points of particle-stuff in the actual world. In the possible world, the point particles have void between them. In the actual world, however, the points of particle-stuff have other points of particle-stuff very near them. The implicit premise here is that mental properties cannot supervene solely on particle position; they must also supervene on the void between these particles.

But there are two considerations which render the claim made in the objection unlikely. First, even supposing the standard account that at most one mind supervenes on a brain, there is not a lot of void between the particles on which a mind supervenes. There are many particles in the brain, which presumably are not directly involved in the production of consciousness, but are in between the particles, which are. For example, neurons in the brain are held in place by *glial cells*, which probably have no direct effect on the state of the mind. It is agreed that mental properties can supervene on the particles, which compose the neurons of a brain, even though there are many other particles dispersed throughout the brain. Hence, it is at least the case that particles do not need *a lot* of void between them in order to form a mind-constituting arrangement.

Second, a thought experiment can be used to suggest the unreasonableness of the third objection. Consider a series of brains: at one end there is a brain composed of N point particles (as in the possible world discussed above), while at the other end there is a brain composed of N particles with extension, and all the brains evolve in the same way. The next entry in the series, after the brain with N point particles, is just like that brain except with a few extra points of particle-stuff scattered throughout the brain. Fill in the series with brains that have more and more points of particle-stuff: is there some step in the series where mental properties would cease to supervene on the N point particles? A proponent of the third objection would have to maintain that there is, since such a person would maintain that mental properties supervene on the N point particles of the first brain in the series, but mental properties do not supervene on the corresponding N points of particle-stuff in the last brain of the series. The more reasonable intuition, of course, is that there is no step in the series where mental properties cease to supervene on the N point particles, and hence the third objection is incorrect.

But one might continue to worry that there are reasons to doubt whether numerically distinct minds supervene on the same relevant *physical* properties.[2] There are two possible grounds for doubt on this score; we will argue that neither ought to compel.

The first ground for doubt on this score is this. It might alleged that, while it makes sense to speak of a mind supervening on a collection of particles (as these evolve over time), it makes no sense to speak of a mind supervening on a collection of points of ''particle stuff'' (as *these* evolve over time). But above we have already

[2] This objection is owed to an anonymous referee of *Minds and Machines*.

noted that, given functionalism about mental states, we can make sense of a mind supervening on a collection of points of ''particle stuff''—at least we can so long as the evolution of these points has the requisite sort of functional profile. What is more, we noted that this collection of points of ''particle stuff'' will have the *very same* functional profile as something that everyone can agree is a mind—namely, the very mind that supervenes on the collection of particles constituting the individual's brain. Thus it would appear that this first way of spelling out the objection from supervenience succeeds only as a challenge to functionalism itself. (We will return to this below, in the section regarding the role of functionalism in our argument, where we take up the question whether our argument is best seen as presenting a case against functionalism.)

The second (related) ground for doubt on the score of supervenience is based on the contention that, even if it does make sense to speak of a mind supervening on a collection of points of ''particle stuff,'' points of ''particle stuff'' have different physical properties than do particles themselves—leaving it unclear whether it is correct to say that numerically distinct minds supervene on the same relevant *physical* properties. But this worry fails to appreciate the role of supervenience in our argument. Our claim is *not* that each of the infinitely many distinct but qualitatively identical minds that is associated with a given brain *has the same supervenience base*. Precisely not: our claim, that there are an infinite number of *distinct* but qualitatively identical minds per brain, itself *depends on* the claim that each of these minds has its own unique supervenience base. In particular, it is the distinctness in the supervenience base of the mental states constituting each mind that distinguishes each of the (otherwise qualitatively-identical) minds themselves. So while the present version of the objection from supervenience is correct to note that the distinct minds do not supervene on the same relevant physical properties, it is wrong to suppose that this amounts to an objection to (as opposed to an implicit assumption of) our argument.

The role of functionalism in our argument

Above we noted that we are presupposing functionalism. But now it might be thought that the foregoing argument is best treated a *reductio* of that doctrine since, given functionalism (together with the three premises of our argument), we get the absurd result that there are infinitely many distinct qualitatively-identical mind per functioning brain.[3] Such a reaction to our argument might be further bolstered by noting a result from Putnam (1988), to the effect that every physical system implements any finite-state automaton. The claim then is that our present argument is yet another way, in addition to the argument of Putnam (1988), to demonstrate the inadequacy (because over-liberality) of functionalism.

Here is not the place to explore this matter in any detail, but we do want to make a programmatic remark. Despite arguments (such as that of Putnam, 1988) aiming to undermine functionalism, many people in contemporary philosophy of mind remain committed functionalists. Stronger, it is arguable that functionalism remains *the* dominant view. Perhaps our argument will tip the scale against functionalism; but we

[3] We owe this suggestion, along with the reference to the results of Putnam (1988), to an anonymous referee of *Minds and Machines*.

doubt it. Nor is it obvious to us that the rejection of functionalism would be the *proper* reaction to our argument. In particular, functionalists can preserve the reasonableness of their position in the face of the foregoing argument by claiming that, in the absence of a viable alternative philosophical account of mentality, it remains reasonable to retain functionalism, and to go on to regard all would-be arguments to the contrary as problems to be addressed (rather than as reasons to surrender the doctrine itself). Such a move would be no less reasonable here than it is in the sciences more generally; and as we have no case against a move of this sort in the present case, and since we do not pretend to have a novel alternative philosophical account of mentality, we are prepared to grant that functionalism is not the odd man out. Rather, we submit that our argument is better interpreted as yet another problem on the list of problems that must ultimately be dealt with by those who continue to endorse functionalist approaches to the mind.

Evaluating the premises

We conclude that the argument presented is valid; should we believe the premises? The first premise, that mental properties supervene on brain properties, is a widely but not universally accepted one. Discussing its merits is beyond the scope of this paper.

The second premise, that the universe is composed of particles with nonzero extension, may or may not be true. Physics has not yet, and may well never, decide the issue. It is worth noting that according to (some interpretations of) quantum field theory, particles are not a fundamental constituent of reality. Nevertheless, perhaps one could present a related Problem of the Many Minds for field theories. Whether one could do so would depend on how mental properties supervene on fields, and this issue is very much open for discussion.

The third premise, that each particle is composed of continuum many points of particle-stuff, is the natural way to understand the ontology of a particle with extension.[4] There is stuff interior to the particle, and that stuff will be located at various points interior to the particle; presumably there will be stuff at continuum many points. We have also assumed that each point of particle-stuff persists through time in the same way that the particle does. One might claim that points of particle-stuff do not persist. We doubt that this position is coherent, though: it is not clear how a particle could persist if the points of particle-stuff, which compose the particle do not persist. The alternative view would have to be a holistic theory where, even though no part exists for more than an instant of time, the whole comprising the parts persists through time. Such a theory seems metaphysically unreasonable, so we maintain that the third premise is a reasonable one.

[4] Here it may be worth noting what we do not intend by ''a point of particle-stuff'': we do not intend a point-sized portion of stuff that is not itself a material object but which together with some other stuff constitutes an indivisible material object. Such a reading reflects recent work on a dual ontology of stuff and things where the constitution relation is a relation between some stuff and a thing; we do not intend to be read in this way.

Conclusion

Why is it that we call the Problem of the Many Minds a ''problem''? Shouldn't we simply take the claim that there are an infinite number of minds per brain to be an interesting fact entailed by many physical theories? The claim being made in this paper *does* pose a problem for proponents of such theories. The problem is definitely *not* one of empirical adequacy, however. Intuitively speaking, it may *seem* to you that there is only one mind supervening on your brain, but it seems that way only because your experience is constituted by *that one mind*. Your experience does not depend on whether the other minds exist; and it would presumably be impossible for you ever to directly experience whether these other minds exist. Hence you presumably cannot experimentally verify whether these other minds exist.

The problem is simply that the conclusion of our argument seems *unbelievable*. We conclude by mentioning what we take to be the three most viable potential responses to the problem. The first potential response is to maintain that the argument is invalid. But we do not see any problem with its validity (although greater perspicuity in this respect might require greater care in formulating the premises). The second potential response is to reject one of the premises. If one were to endorse this response, we would recommend rejecting the second premise, and instead accepting a theory with a point particle ontology, or a theory which does not have a particle ontology at all. More interesting, though, is the third response: one can simply *embrace* the problem and believe that many more minds supervene on a brain than was previously thought.

References

Putnam, H. (1988). *Representation and reality*. Cambridge: MIT Press.

Lewis, D. (1986). 'Causation' and postscripts. In *Philosophical papers Volume II* (pp. 159–213). Oxford: Oxford University Press.

Unger, P. (1980). The problem of the many. *Midwest Studies in Philosophy, 5*, 411–467.