Free Will of an Ontologically Open Mind

Jan Scheffel Department of Fusion Plasma Physics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

Abstract

Combining elements from algorithmic information theory and quantum mechanics, we have earlier argued that consciousness is epistemologically and ontologically emergent. Accordingly, consciousness is irreducible to neural low-level states, in spite of assuming causality and supervenience on these states. The mind-body problem is thus found to be unsolvable. In this paper the implications on free will is studied. In the perspective of a modified definition of free will, enabling scientific decidability, the ontological character of processes of the cortical neural network is discussed. Identifying conscious high-level processes as ontologically open, it is asserted that conscious states are indeterminable in principle. We argue that this leads to freedom of the will.

1 Introduction

Must we have the thoughts we have? Do our thoughts only happen, rather than being created by ourselves? Does determinism hold our will into an iron grip? The free will problem presumably is the most important existential problem and has generated shelf kilometers of literature throughout the centuries. One reason for the problematic situation could be traced to the most common definition of free will: 'the ability to act differently'. Indeed, it is hard to see any opportunity for scientific methods to determine whether we actually can 'act differently' or not. How do we know whether an individual's actions are autonomous or predetermined? And why should even a free consciousness act differently in two identical situations? Many arguments about the will thus lead to uncertain terrains.

In Scheffel (2018) it is argued that consciousness cannot be represented by a theory and, as a consequence, that the mind-body problem is unsolvable. The associated epistemological emergence of consciousness is related to the problem of free will since if, on the other hand, a theory for consciousness could be designed, then its behaviour would be computable or could be simulated. It would thus be predictable and not free. The argument is subsequently carried a step further to show that consciousness, as a high-level property of the mind, is *ontologically emergent* with respect to the low-level neural states. Although the latter form the basis for consciousness, it is argued that consciousness is not ontologically reducible to these properties. The reasoning is based on elements of algorithmic information theory (Chaitin 1987), and that the limited quantum mechanical information and computational capacity of the world (Lloyd 2002 and Davies 2004) represent an unsurmountable obstacle. The main argument is that if properties of a complex system, being the result of for example long term evolution, can only be manifested by the system itself - that is if nature for reasons of limited information storage capacity cannot accommodate a representation of the system - then the system features ontologically emergent properties. Thus, since a relation to its constituting low-level neural components cannot be expressed, consciousness in a sense comes as a surprise to nature.

In this paper we argue that the ontologically emergent character of consciousness dissolves the deterministic contradiction we have been facing for freedom of the will. The paper ends with discussion and conclusion.

2 Alternative definition of free will

As reasoned above, standard characterizations of free will like, for example, 'the ability to act differently' are problematic. In Carnap's (1950) view, a transformation from the prescientific explicandum to a scientific explicatum would have the advantage of rendering free will a concept available for scientific study. Along these lines the following definition will be employed in this work: *A conscious individual has free will if its behaviour takes place according to its intentions, the intentions are not subconsciously generated and if the individual's mind is an ontologically open system*.

By 'will' we refer to preferences by a cognitive system for certain desires or future actions. Furthermore, by '*ontologically open system*' is meant a causal, physically closed high-level system the future of which cannot, even in an *a posteriori* sense, be reduced to the states of its associated low-level-systems.

We motivate this definition of free will as follows. Experience tells us that basic, lowlevel phenomena are causal and essentially deterministic. Quantum mechanics tells us, however, that certain corrections of a statistical character must be taken into account, as discussed below. We will assume that account has indeed been taken of these effects when we henceforth make use of the term 'deterministic'. If also the high-level neuronal functions and processes being associated with consciousness are deterministic, it is quite natural to draw the conclusion that expressions of will are governed by processes outside its conscious control. This is a feature of the classical, deterministic argument against free will. On the other hand, behaviour related to ontologically open conscious systems is not directly reducible to earlier physical neural states. As will be argued here, this is a characteristic related to ontologically emergent properties. It should be noted, however, that ontological emergence does not straightforwardly imply ontological openness; even if high-level properties cannot be simply reduced to those of low-level it must be shown that downward causation is possible. Since ontological openness, as argued in this paper, is possible also for systems that are essentially deterministic at low-level, the notions of compatibilism and incompatibilism become irrelevant in this analysis.

The concept of 'reduction' is central for the argument. As discussed in Scheffel (2018), it is problematic that 'reduction' is a widely debated concept among philosophers and that there is limited consensus when it comes to details (van Riel and van Gulick 2018, van Gulick 2001). It is in our view reasonable to assume, as van Riel and van Gulick do, that *ontological reduction* should entail "identification of a specific sort of intrinsic similarity between non-representational objects, such as properties or events". An ontologically irreducible property, if it exists, could not be determined by its low-level-properties or behaviour; it could not be characterised by a statistical or law-like behaviour in relation to its low-level components. It is not implied by nature. It was argued in Scheffel (2018) that even assuming causality, the extreme complexity of consciousness, in an ontological sense, 'shields' the dynamics of high-level conscious activity from that of its associated low-level properties are not ontologically implied by the system. The system becomes a mere vehicle for these properties. As a consequence the system is uncontrollable in principle.

At this point we need to distinguish between open and closed systems. Phenomena relating to classical *open physical systems* are generally causal, but indeterminable. These systems are open to external influence, and they are thus not guaranteed to evolve identically when repeatedly started from the same initial conditions. The associated dynamic processes should not be regarded as random or chancy; the point is that the system itself does not contain sufficient information about its future external states. This

becomes clear if we extend the size of the system to also include all of its external influences. Such a system may indeed be *physically closed*, causal and deterministic. We will, in the next section, argue that consciousness is ontologically open in spite of being a physically closed system.

For the sake of completeness and accuracy we should, when discussing the dynamics of open and closed systems, account for that quantum mechanics shows that determinism does not fully apply at the very micro-level. The uncertainty principle of quantum mechanics implies that nature is 'blurry' at the sub-atomic and atomic particle levels in the sense that, for example, the position and velocity of a particle are quantities that cannot, even ontologically, be assigned exact values. For larger clusters of particles, however, like the molecules that make up the neurons, this effect is of much less importance, because of so-called quantum decoherence. The concept of 'adequate determinism' has been coined to emphasize that the statistical determinism that results and is used here, in essence is correct in the macroscopic world, even if quantum phenomena are important on the micro-scale.

Returning to the definition of free will stated above, it is emphasized that the desired actions of a free consciousness must not turn into anything other than intended; behaviour must be based on its intentions. By 'intention' we here adhere to the everyday definition 'determination to act in a certain way'. Now, if I wish to consider what to eat for dinner, such a reflection must be possible. My choices and actions must consistently and adequately follow my will. The phrasing '*takes place according to its intentions*' is deliberately somewhat vague in the sense that the precision we may strive for in our actions is sometimes not achieved; this is not because the will is not obeyed but rather from our physical and psychological limitations. Note also that we assume conscious individuals; it is not meaningful to talk about 'will' for other systems.

Finally, the condition that 'the intentions are not subconsciously generated' is needed to ensure that the individual's brain does not contain any hidden systems that manipulates it in a manner that consciousness, in spite of being controlled this way, experiences intentions as its own. So-called 'character decisions', being decisions based on our experiences and consolidated positions that we make without active reflection, we treat in this context as conscious. We will return to these.

There is a subtle, but important, observation to be made. Even if our conscious desires and decisions would be completely ruled by subconsciousness, the latter has, if the mind constitutes an ontologically open system, capacity for choices that are not predetermined. Thus, even if subconsciousness rules the mind, the individual can be regarded as morally and legally responsible for its activities due to the ontologically open character of its mind. It has, over time, had the ability to integrate the consequences of its actions into its considerations. Hence the debate concerning to what extent subconsciousness rules our decisions is essentially irrelevant as far as moral and legal matters are concerned if the human mind behaves as an ontologically open system.

To sum up, we have cast the characterization of free will as 'the ability to act differently' into an alternative, scientifically decidable formulation in order to improve the methodological conditions to address the free-will problem. A simple, but less precise, condensation of the definition could be something like 'An individual that can realize conscious, unforced choices has free will'. The task is now to address the, as it seems, inhibiting circumstance that the mind must feature a deterministic character in order to enable coherent thought processes and consistent performance of its intended actions, while simultaneously feature an ontologically open nature in order to permit self-caused actions. If this potential contradiction can be dissolved, there is room for free will. It is at this point the ontologically emergent character of consciousness plays an important role.

We next aim to show that the associated ontological irreducibility of consciousness to low-level neural states renders consciousness ontologically open. The individual's cognitive processes, such as actions of the will, are not reducible to low-level neural states, thus rendering them undetermined in principle.

3 Consciousness and determinism

In the following it will be argued that consciousness is an ontologically open, high-level global system. The question thus arises how it would be possible for the brain's essentially deterministic, molecular low-level activity to lead to ontologically open behaviour at the higher inter-neural levels related to consciousness, considering that man and consciousness are of the physical world.

Experience shows that causality applies in the physical world. This means that a current state of a typical physical system, in terms of the positions and velocities of its microscopic constituents, provides a sufficient condition to take it to a subsequent state; cause results in effect. We usually endeavour to find and express these regularities of cause and effect as laws of nature. If new physical states can be found uniquely from previous states of the system, we talk about determinism. Stated equivalently: determinism implies that the evolution of a system, if repeatedly started from the same initial conditions, will always be identically the same. Everyday events, such as when the billiard cue hits the cue ball which subsequently knocks down the yellow ball in the hole, tempt us to believe that causality and determinism are equivalent concepts. But they are not. The future of a specified causal physical system may actually be undeterminable, even disregarding the statistical nature of quantum mechanics. This happens when the system is open in some sense, that is when external phenomena may have an influence.

Let us consider the behaviour of a hypothetical single conscious individual placed in a closed room, without contact with the outside world. We are interested in the specifics of the individual's behaviour in a certain future time interval. For the sake of argument let us first consider an imagined case that we could deem as fundamentally undeterminable with respect to the individual's choices and actions. If the individual, before taking a decision, had the magic ability to consult a clever genie inhabiting some dimension otherwise unrelated to our physical world, the individual's future would clearly not be deterministically given. The influence of the genie's advice on the individual's behaviour would be comparable to the case of external signals influencing the dynamics of an open physical system. Since the individual's decisions are not immediate consequences of its present physical state of mind, we must infer that the will of this individual is not limited entirely by a deterministic dependence on its initial set-up and conditions in the physical world. In discussions of determinism, in a similar vein as that of Laplace in Essai philosophique sur les probabilities (1814), it is often asserted that given the positions and velocities of all particles in the universe, its future would be in principle determinable. The argument implicitly assumes the continual action of the laws of nature. Here the appearance of the genie violates this assumption.

Returning to reality, the genie of the thought experiment can, with a similar result, be replaced by the individual's ontologically emergent thought processes in combination with preferences being acquired during its earlier history, now stored in its memory. Will is about planning and experiences play central roles. Experiences are personal and rated subjectively, whereafter they are remembered and used as a basis for subsequent preferences. The stored preferences are consulted, similarly as the genie, before decisions are taken. Furthermore the formation of new preferences are the result of *ontologically emergent* processes where subjective positive or negative connotations have been related to various events, actions and choices.

Alternatively formulated, consciousness acts as an open system in the sense that the memories associated with subjective preferences, as neural processes, are ontologically detached from the current physical low-level situation. The fact that in principle one can, atom by atom in a Laplacian sense, build the individual's entire network of coupled neurons is not relevant here. The system has built in subjective preferences, the character of which are ontologically unknown (memories have no ontological meaning considered at low-level) and, as we have argued, function in the same way as when conferring with an independent genie. Ontological emergence is crucial in that it decouples the physical low-level state of the individual as a system from its subjective properties and behaviour, thus enabling downward causation.

To sum up, we have argued that consciousness is an ontologically open system and thus undeterminable and uncontrollable in principle. Conscious will is, rather than being determined by low-level neural properties, the result of ontologically emergent high-level processes including subjective experiences, stored as memories in the mind.

We may wonder: how could such complex behaviour evolve in humans? Perhaps the most competitive evolutionary aspect of consciousness is its ability for planning in order to avoid dangers, to gain advantages and to optimize long time survival. Planning requires alternatives to compare with. The alternatives manifest themselves to us humans primarily through experience; we are not born with fixed perceptions about the world and cannot be since, for example, our environments differ depending on where we are born. Our experiences need storage, or memory, to manifest themselves as conscious alternatives when we are about to make choices. Associated with these objective experiences, we have also stored subjective impressions. In the process of planning, when making our choices, it is precisely the subjective impressions that influence our choices or decisions. These individual impressions also define our characters; so-called 'character decisions' will be discussed in the next section. Ontologically emergent brain processes helps to store personal and subjective impressions for subsequent use in decision making processes. The effect of memory to continuously modify consciousness results in that consciousness may respond or act differently, even if external conditions are unchanged. Repeatedly facing identical external conditions, conscious individuals can make new and different choices each time, as a result of recollections of subjective experiences of earlier instances. This does not mean, however, that two identical conscious systems, provided with identical external conditions, will act differently. This would violate causality, which we assume. The conscious, subjective choices made by these systems will be identical. But they are not predetermined; there is no ontological reason for the systems to think in a certain way.

It is essential for free will that consciousness is ontologically emergent rather than merely epistemically emergent. In the latter case an imagined powerful demon, with access to all information in the universe including all details of the individual's consciousness, could control and manipulate the individual to act in any specific way by engineering its low-level neurons. But an ontologically emergent consciousness is without reach for the demon, it is free in the sense that it cannot, even in principle, be controlled.

We have, from a physicalistic and thus monistic position, argued for that the mind is an ontologically open system. Interestingly, the same result seems to follow from a dualist perspective. To show this, assume for a moment that dualism holds; there is both a material and a somehow separated 'mental dimension'. What characterises activity in the mental dimension? Certainly not randomness; scientific analysis of mental behaviour speaks against this. But if the mental dimension features regularity and law-bound processes we face a similar question as when taking the physicalistic stance: what is the maximum freedom that can be excerted by the will, given the laws of nature? Thus a natural conclusion is that dualism does not appear to provide conscious will with higher degrees of freedom than those found within physicalism.

4 Willed intentions and the role of subconsciousness

Free will requires, in line with the definition employed here, that individual behaviour takes place *according to the individual's intentions*. This condition is not really problematic; it is satisfied by our experiences. The individual's everyday functioning is completely dependent on that she consistently carries out what she decides. Does she want to make herself a cup of coffee, she does it. The exceptions that can be identified, such as shortage of coffee or that she is interrupted, are not about principal mental limitations but of properties of the outside world.

So far, we have presented arguments for that consciousness/subconsciousness as a combined system meets the requirements for free will. But few would regard this as sufficient; if our volitional decisions, in spite of their onthologically open origin, are unconsciously dictated to us it would be difficult to speak of free will. There is evidence, however, that consciousness in a number of situations exerts its will without significant influence from mind processes that we would refer to as subconscious. First, it should be noted that there is a spectrum of degrees of collaboration between the two. Our experiences of dreams show that subconsciousness may be active when we are not consciously aware. Driving a car along a well-known road is a well known example of symbiosis between consciousness and subconsciousness. And participation in an intense discussion is an example of consciousness mainly acting on its own. But the independent role of consciousness and the will has been strongly questioned over the past few decades and some authors talk of "the illusion of free will". Support has been partly found from neuroscience. A 'readiness potential', being activated unconsciously well before we make conscious decisions, appears to reveal that the main decision-making takes place beyond consciousness. A pioneer in the field was Libet (1985). Experiments in this field has, however, many possible sources of error, thus criticism comes from several places (Klemm, 2010 and Baumeister et al, 2011). We briefly consider some of these arguments.

In certain practical situations it is, from an evolutionary point of view, crucial that consciousness may act undisturbed. The need for rapid and well balanced decisions, as when we are driving a car and we suddenly need to consider how to avoid a car that suddenly wobbles into the roadway, is one example. In a very short time we need to perform a large number of considerations, including how to avoid colliding with people while at the same time ensure our own safety. The subconscious mind would not, with the associated delay that Libet's and other experiments show, have time to gather all the relevant information in order to survey the situation and in a short time deliver adequate decisions that do not conflict with our conscious perception and handling of the situation. Certainly, if conscious decisions would not be important in situations like these, evolution would likely have provided us with a mechanism that automatically disconnected consciousness in favour of subconsciousness, like when we react reflexively.

Furthermore it is well known that, upon learning new knowledge and skills, performance is gradually taken over by the subconscious as we become more knowledgeable and skilful. But for the beginner who sits down at a piano, the subconscious mind is completely unprepared. There is no way for the subconscious to control the finger movements because it does not 'know' what should be done (Klemm, 2010). Obviously more research is needed to identify to which degree subconsciousness impacts on our actions. In the examples given above, however, the subconscious cannot reasonably have a significant role.

The cooperation between consciousness and the unconscious points to a second argument why consciousness is not controlled by the subconscious. Neuroscience shows that a significant part of the 'processors' of the brain used for conscious thought are also used for unconscious processes (Dehaene, 2014). This supports the idea that also subconscious neural processes are ontologically emergent. Thus, whereas deterministic processes contribute to low-level communication between consciousness and the unconscious, these systems can both, on high-level, be regarded as ontologically open systems that do not deterministically control each other. As pointed out, experience shows that we can consciously cancel impulsive intentions, using "free won't" (Libet, 1985).

From another perspective, we do not necessarily need to make a distinction between consciousness and subconsciousness as separated global systems. Already individual neurological subsystems associated with the mind appear to be sufficiently complex to render their interaction ontologically emergent. In the subject of game theory similar results have, interestingly enough, been found. Emergent behaviour has been observed in simulations of nonlinear interaction between two players, who both act in order to optimize their game while trying to act unpredictable for the opponent, if players are allowed to make use of the game's history (West and Lebiere, 2001).

A complication related to the distinction between subconscious and conscious choices is what might be called 'character decisions'. Based on previous experience and reflections, people accumulate different, often conscious, positions or traits of character that could lead to routine behaviour in certain situations. Facing an approaching threatening individual, for example, certain people will normally escape while others preferably stay to deal with the danger. This behaviour does not necessarily constitute an active conscious choice of the type we have discussed so far, but may rather be a result of the individual's *disposition* to act in such situations. Since the individual normally is aware of her traits of character, we here consider the nature of character decisions to be conscious rather than unconscious.

Our feelings, thoughts and choices do not simply happen to us. They arise from basic neural processes related to our minds and are developed emergently in a cooperation between consciousness and the unconscious. But how, then, can our thoughts and feelings take form in a structured and coherent way? How can the individual carry out her intentions unruled by the subconscious? These important questions are not analyzed here; of prime interest for the question of free will is that thoughts, feelings and choices arise in a manner which is indeterminable in principle.

5 Discussion

Our analysis is consistent with non-reductive physicalism where mental states supervene on physical states but cannot be reduced to them. Thus there are similarities with Davidson's theory of anomalous monism (Davidson, 1970) in which it is claimed that there are no strict laws on the basis of which mental events can be predicted or explained by other events. We may ask what the consequences are for *causal closedness*, that is the thesis that no physical events have causes beyond the physical world. Our answer is that, in the present view, the physical world indeed is causally closed in the ontological sense. Causality holds; any physical state leads, in accord with the laws of nature, to new states. For simpler, low-level systems, new states are in principle predetermined and sometimes even computable. The human brain employs deterministic low-level processes at the neural level for thought processes, carrying out certain actions (somatic nervous system) and for reflexes (autonomic nervous system). But as we have shown, this does not mean that all systems in the physical world are predetermined. Emergence can alter the situation. Consciousness, which we argue to be ontologically emergent, is such a highlevel system. In an epistemologic perspective this means that the possibility for conclusions about the causal functionality of mental systems are limited. The situation is reminiscent of that of mathematics for which Gödel proved that there are true theorems in the system that are unprovable because of their complexity.

We can now explain why emergence does not cause overdetermination with regards to the causal situation for consciousness (Kim, 2006). It has been argued that if the dynamics of consciousness is determined by its current state and the laws of nature, then emergent phenomena cannot exist independently; they must be a result of the complete set of conditions already provided. Otherwise we seem to be facing an overdetermined problem. The solution to this dilemma is that the emergent properties are of the same nature as the new conditions that may present themselves when a closed system is transformed into an open system. Hence they are additional conditions, being governed by associated additional relations. Mathematically speaking, just as many new equations are added as new variables. Thereby overdetermination is avoided. In our example of the person being placed in a closed room, this could correspond to the door being opened. Emergent properties have thus, as far as deterministic control is concerned, the same impact on the development of the system as external influences have on an open system. This solution to the problem of overdetermination also explains how downward causation (Campbell, 1974 and Kim, 2006) can take place. Interacting emergent phenomena can determine the development of the system (in this case, the mind) to a large extent independently of the causal situation at lower levels.

What is then the implication for *compatibilism*; the position that determinism is compatible with free will? Interestingly, whether compatibilism or incompatibilism holds is not relevant here. Even though neural low-level processes are essentially deterministic, we hold that high-level cognitive phenomena are ontologically emergent in spite of being defined by low-level processes. The point of transition is determined by the amount of information that can principally be stored and processed in a physical system (Scheffel, 2018). This also illustrates the point that in a physicalistic view of the world, determinism or indeterminism has no bearing on its reductive character; it is emergence that renders physicalism nonreductive.

Finally, how do these results relate to *epiphenomenalism*, the notion that mental states are only by-products of the physical states and unable to causally influence these? To answer this question, we need to note that the form of non-reductive physicalism assumed here is not a form of property dualism. Although mental states are not deducible from basic neurological states, they certainly correspond to physical states; they supervene on these. Non-reductionism follows because of the emergent character of mental states, not because of lack of correspondence between physical and mental states. Thus epiphenomenalism is ruled out here.

6 Conclusion

We argue that high-level cognitive processes are ontologically open, even though underlying physical laws and low-level neural processes are essentially deterministic. By an 'ontologically open' system we mean a causal, physically closed high-level system the future of which cannot be reduced to the states of its associated low-level-systems. In consequence the activity of consciousness is not determinable, not even in principle. To consider the impact on volitional processes, a scientifically and methodologically more applicable definition of free will than the traditional 'ability to act differently' is suggested. The three associated requirements for free will are all argued to be satisfied; that the individual's actions take place on the basis of its intentions, that these intentions have not been subconsciously forced onto the individual and that the individual behaves as an ontologically open system. Thus the will, as defined here, is free.

Acknowledgements

Many thanks go to Mr Keith Elkin and professor Erik J. Olsson for constructive discussions on several aspects of the work.

References

- Baumeister, R. F., Masicampo, E. J., & Vohs, K.D. (2011). Do Conscious Thoughts Cause Behavior? *Annu. Rev. Psychol.*, 62, 331-361
- Campbell, D. T. (1974) Downward causation in hierarchically organised biological systems. In Francisco Jose Ayala and Theodosius Dobzhansky (Eds.), *Studies in the philosophy of biology: Reduction and related problems*, pp. 179–186. London/Basingstoke: Macmillan
- Carnap, R. (1950). Logical Foundations of Probability. (University of Chicago Press)
- Chaitin, G. J. (1987). Algorithmic Information Theory. (Cambridge University Press)
- Davidson, D., (1970). Mental Events, reprinted in *Essays on Actions and Events* (Oxford: Clarendon Press, 1980), 207-224
- Davies, P. C. W. (2004). Emergent biological principles and the computational properties of the universe. *Complexity* 10,11-15
- Dehaene, S. (2014). Consciousness and the Brain. (Penguin Books, New York)
- Kim, J. (2006). Emergence: Core ideas and issues. Synthese, 151, 47-559
- Klemm, W. R. (2010). Free will debates: Simple experiments are not so simple. *Advances in Cognitive Psychology*, 6, 47-65
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8, 529–566
- Lloyd, S. (2002). Computational Capacity of the Universe. Physical Review Letters, 88, 237901-1-4
- Scheffel, J. (2018). On the Unsolvability of the Mind-Body Problem, PhilPapers:
- https://philpapers.org/rec/SCHOTS-24
- West, R. L. & Lebiere, C. (2001). Simple games as dynamic, coupled systems: randomness and other emergent properties. *Journal of* Cognitive *Systems Research*, 1, 221-239
- van Gulick, Robert, Reduction, Emergence and Other Recent Options on the Mind/Body Problem: A Philosophic Overview, *Journal of Consciousness Studies* 8: 1–34, 2001
- van Riel, Raphael and Van Gulick, Robert, Scientific Reduction, *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), URL =
 - <https://plato.stanford.edu/archives/sum2018/entries/scientific-reduction/>