The Way We Are Free

Weinbaum, Weaver D.R.

Published in:
The Noosphere Outside In

Publication date:
2023

Document Version:
Submitted manuscript

[Submitted to:] M. Lenartowicz & Jessie Eggers (Eds.) The Noosphere Outside In: Consciousness Beyond Autopoiesis (Forthcoming 2023)

The Way We Are Free

On the problem of free will

By
Weaver D.R. Weinbaum

Email*: space9weaver@gmail.com*

Abstract

Modern understanding of determinism collides head on with our experience of free will and leaves little place for it if any as a real phenomenon. This article proposes a new reconciliation between free will and determinism. It traces the experience of choice to an epistemic gap inherent in mental processes due to them being based on physically realized computational processes. This gap weakens the grasp of determinism and allows for an effective kind of freedom. A new meaning of freedom is explored and shown to resolve the fundamental riddles of free will, or at least going a long way towards that end. Freedom as arising from the epistemic gap is no longer a mysterious property of an agent but rather a dynamic property of a wider state of affairs involving the agent, the distribution of computational resources and information flow. Some implications of the way that we are free and how it might be influenced by future technology are briefly discussed.

Introduction

Rational choice is perhaps the most fundamental and meaningful activity recognized to be uniquely humane and therefore characterizing our humanity. We all have a deep experience of being able to choose. We say that we are choosing to act in this way and not in another way out of our free will, when we have the experience that, all other things being equal, we could act otherwise, out of our independent and unconditional volition. When we lack such an experience we say that we are acting without choice, that we are constrained in our actions by conditions which are not within our control. The ability of a person to exercise free will, making free unconditional choices that is, is a fundament of personal responsibility and is the basis of most ethical systems and the legislative institutions of society. It is therefore no wonder why the question of free will is central to philosophical discourse.

The question is particularly disturbing given that our understanding of the physical world as deterministic and causally closed[1] leaves no place to free will as we will see in more detail next. Determinism collides head on with our very common and accessible experience of freedom. It brings forth a deep existential paradox.

Many philosophers have tried to settle the paradox of free will. There is a respectable philosophical tradition of the issue, but no single philosophical treatment has produced an overwhelming sense of satisfactorily resolving it. This article addresses the issue of free will from an angle that, to the knowledge of the author, has not been explored as yet. It traces the experience of choice to an epistemic gap which is inherent in mental processes, particularly in deliberative processes. This gap, it will be argued, allows for a sense of freedom which is not epiphenomenal, yet is significantly diverting from the commonly accepted meaning of free will.

---

[1] Causal closure means in brief that any event in the physical world is necessarily caused by other events in the world. There is no event without causes or with causes external to the physical world.

The first part of the article will discuss the problem of free will. This is far from an exhausting presentation of the philosophy of free will. It comes only to facilitate those aspects of the issue necessary to understand the argument developed here. The second part will discuss determinism and the epistemic gap inherent in it. The third part explains the experience of free will and how it corresponds to the epistemic gap inherent in mental processes. The fourth part further explores the meaning of freedom that arises in the epistemic gap. The fifth part discusses a few interesting implications that arise from the approach to free will proposed here. It will be shown that freedom of choice is not attributable to a particular agent but is a property of a wider state of affairs that involves the agent and its immediate domain of relevant cognition and action. It is shown that issues such as interactivity, privacy, consciousness, cognitive and mental speed are influencing the degree of freedom an agent may enjoy independently of any concrete constraint he may be subject to.

What is the problem of free will?

There are quite a few comprehensive descriptions of the problem in the literature. Here is a brief and simple account made by John Searle (Searle, 2001):

> There is no question that we have experiences of the sort that I have been calling experiences of the gap; that is, we experience our own normal voluntary actions in such a way that we sense alternative possibilities of actions open to us, and we sense that the psychological antecedents of the action are not sufficient to fix the action. Notice that on this account the problem of free will arises only for consciousness, and it arises only for volitional or active consciousness; it does not arise for perceptual consciousness. What then, exactly, is the problem of free will? Free will is typically taken to be opposed to determinism. The thesis of determinism about actions is that every action is determined by antecedently sufficient causal conditions. For every action the causal conditions of the action in that context are sufficient to produce that action. Thus, where actions are concerned, nothing could happen differently from the way it does in fact happen. The thesis of free

> will, sometimes called "libertarianism", states that some actions, at least, are such that antecedent causal conditions of the action are not causally sufficient to produce the action. Granted that the action did occur, and it did occur for a reason, all the same, the agent could have done something else, given the same antecedents of the action.
>
> …
>
> We come by the conviction of the freedom of the will, in my sense, because of the experiences of the gap. So the problem of the freedom of the will can be posed as follows: what reality corresponds to those experiences? Granted that we experience our actions as not having an antecedently sufficient, psychological, causal conditions, why should we take this psychological fact seriously? Is it not possible that the neurological underpinnings of the psychology are causally sufficient to determine the action, even though the psychological level by itself is not causally sufficient? And could there not be unconscious psychological causes determining the act? Even granted the psychological reality of the gap, we still have a problem of free will left over. What exactly is it and how exactly might we go about solving it?
>
> …
>
> The gap is a real psychological phenomenon, but if it is a real phenomenon that makes a difference in the world, it must have a neurobiological correlate. [pp. 276-81]

Searle goes on to examine a sort of modified compatibilist hypothesis that states that there is psychological libertarianism (free will) compatible with neurobiological determinism. But in conclusion he reflects (Searle, 2001):

> The result, however, is intellectually very unsatisfying, because, in a word, it is a modified form of epiphenomenalism. It says that psychological processes of rational decision making do not really matter. The entire system is deterministic at the bottom level, and the idea that the top level has an element of freedom is simply a systematic illusion… If hypothesis 1 is true, then every muscle movement as well as every conscious thought, including the conscious experience of the gap, the experience of "free" decision making, is entirely fixed in advance; and the only thing we can say about psychological indeterminism at the higher level is that it gives us a systematic illusion of free will. The thesis is epiphenomenalistic in

> this respect: there is a feature of our conscious life, rational decision making and trying to carry out the decision, where we experience the gap and we experience the processes as making a causal difference to our behavior, but they do not in fact make any difference. The bodily movements were going to be exactly the same regardless of how these processes occurred.
>
> Maybe this is how it will turn out, but if so, the hypothesis seems to me to run against everything we know about evolution. It would have the consequence that the incredibly elaborate, complex, sensitive, and – above all – biologically expensive system of human and animal conscious rational decision making would actually make no difference whatever to the life and survival of the organisms. Epiphenomenalism is a possible thesis, but it is absolutely incredible, and if we seriously accepted it, it would make a change to our worldview, that is, in our conception of our relations to the world, more radical than any previous change, including the Copernican revolution, Einstein's relativity theory, and quantum mechanics. [pp. 281-6]

Searle has a strong point in stating that the proposed hypothesis makes no evolutionary sense. There is no problem that our experience of free will is indeed an elaborate illusion, but it is deeply disturbing that such an expensive illusion makes no evolutionary sense.

To conclude, the problem of free will can be summarized as following:

1. Our experiences of freedom of will do not correspond to our understanding of the world as causally closed and deterministic.
2. We have to accept therefor that either free will is a kind of an epiphenomenon, or determinism is not true.
3. Given that we do not give up on determinism, we cannot account how the experience of free will corresponds to anything that produces real effects in the physical world. It is thus an epiphenomenon.

4. The thesis that free will is an epiphenomenon is not consistent with our knowledge about evolution. In other words, it is highly implausible and actually incredible that evolution would select such an elaborate and expensive experiential illusion if it cannot possibly represent any evolutionary advantage for the organism. So even dismissing free will as an epiphenomenon does not make the problem go away.

In the following I propose an alternative thesis reconciling free will and determinism. I argue that the meaning of the experience of free will is substantially different from the commonly accepted meaning, yet it does not turn to be an epiphenomenon. Moreover, the evolutionary advantage of the experience of free will can be explained.

Determinism, mental states and the epistemic gap

Determinism has been a part of philosophy for a very long time. When we think about determinism we generally think about a philosophical theory about effects and the way they are necessitated by other effects that we call causes or more accurately the causal circumstances of the effect. Something being an effect, calls attention to the fact that it has causes that necessitate it. The relation between effects and causes and the way they define each other, is at least in part, the subject matter of determinism as a philosophical theory. Without entering into the intricacies of determinism, I will use here Ted Honderich's description (Honderich, 2002):

> …we take an effect to be an event that was preceded by a causal circumstance – such a circumstance being something that would still have been followed by the effect whatever else had been happening… If we understand by 'X' any possible change in things consistent with the existence of a causal circumstance, then what was true of the circumstance is that since it occurred, even if X had occurred, the effect would still have occurred. Nothing else mattered or could have got in the way. [pp.14]

Another important aspect of determinism is the concept of causal chains (Honderich, 2002):

> Causal circumstance for an effect will typically be made up of parts that were also effects themselves. So, as we can say, this whole circumstance was the effect of an earlier one, maybe a long time before. Then the earlier circumstance also necessitated the final effect. This fact about effects – the fact of what you might call causal chains – is very important to determinism. [pp.15]

Implicit in the idea of causal chains is the idea of the world being a causal closure. A simplified description of this idea is that our world is such that all possible effects in the world are caused *only* by causal circumstances in the world. There are no causes or causal circumstances outside the world.

This brief account of determinism is consistent and coherent with our general everyday thinking about the world and events in the world. Whether we think about cars, eggs, rain, global warming, viruses, or any other object involved in our everyday affairs, our understanding of causes, effects and their relations seem to faithfully describe the lawful happening around us. But when it comes to our mental states and more specifically our choices and decisions, we do not think about them as necessitated events, as effects of preceding causal circumstance. It seems to us that we are the originators of our choices and decisions, and being of our origination, our choices and decisions are not necessitated effects. We have that deep sense that at a point of choice, everything else being the same, *we could choose otherwise*. In other words, our decisions and choices, being originated, cannot be accounted by determinism.

Before we continue, a brief discussion about the nature of mental states is in place. The nature of mental states or mental events is undoubtedly the most fundamental question in the philosophy of mind (Lowe, 2000). Mental states are those items that make up our

subjective conscious streams of experience. Though it seems we barely begin to understand the nature of mental states, there is an emerging consensus that they are a natural phenomenon, that they are taking place in space and time continuum and are correlated to the neural activity of the brain (Honderich, 2002), (Searle, 2004). Mental states therefore necessarily partake in the causative closure of the phenomenal world, they are determined by their neural correlates and thus the freedom we experience in choice or decision can be explained away. There is no such freedom.

In my analysis of determinism as applied to mental states, however, I found something quite surprising that seems to have been overlooked in other discussions on the subject. It has to do with how determinism relates to knowledge in practical situations.

An almost trivial consequence of determinism is that if an effect E is necessitated by causal circumstance C, knowing that C occurred at some point in time t1, one can predict that E will occur at some point in time t2 such that either t2=t1 i.e., the effect E is coincident with the occurrence of C, or, t2>t1 i.e., the effect E occurs at a time arbitrarily later than the occurrence of C. In the second case where t2>t1, it can be said that a time interval T=t2-t1 is necessary for the causal circumstance C to develop (possibly through a chain of intermediate effects) into E. If we observe a machine producing an artifact, a plant growing from a seed, or a computer performing a complex computation as well as many other examples, we conclude that the causal circumstances ensuring the final effect are incomplete without the time necessary for the effect to be produced. The time interval T needed for the process of producing E is therefore an integral part of the causal circumstance that necessitates the eventual effect E.

That C as the causal circumstance of E is incomplete without the incorporation of T is somewhat disturbing. We would like to think about C as an event or a compound set of events and conditions. The time interval T is neither an event nor a condition. Time

intervals are measured by counting events[2]. Only if the time interval T is somehow filled by other events such as the displacement of the hands of a clock, or the cyclic motions of heavenly bodies, it can be said to exist. If such events are regular, that is, repeating and countable, then T can be measured by counting them. If no event whatsoever can be observed to occur between t1 and t2, how can one possibly tell that there is a temporal difference between them, that any time has passed at all? T becoming part of C should mean therefore that a nonzero number N of events must have occurred in the course of E being produced from C.

T can be a component of the causal circumstance C in two distinct ways. An *effective* T is the case that T is effective in the production of E from C. Effective T comes to mean that a number of events within T, *not being a consequence of C*, could influence the eventual occurrence of E. For example, events in T could turn out to be disruptive in relation to one or more of the components of C therefore preventing the occurrence of E after all. The second case, a *neutral* T, is the case that T is neutral in the unfoldment of E from C. Neutral T, comes to mean that whatever event that may reasonably have taken place during T is either a consequence of C or unrelated yet consistent with C. Such events just designate the passage of the time interval T.

If T is effective then it is clear that C is an incomplete causal circumstance for E, which is equivalent to establishing that during the time interval T or at least part of it, E cannot be said to be determined by C. T being effective in the production of E is interesting as it reflects our everyday practical sense that whenever there is a gap in time between C and E, even the shortest gap, since we never possess a complete knowledge of all events and their relations in the universe, something unpredictable by us may come up and disrupt the occurrence of E. T may always hide a surprise for us as observers and though such surprises seem not to threaten in any fundamental way our understanding of the universe

---

[2] At this point we need a working definition for an event: an event is any difference in the state of affairs that makes a difference which can be perceived by a conscious observer. More generally, an event is anything that changes the state of an observer.

as being deterministic, it weakens the meaning of determinism. If the knowledge of C that produces E is often incomplete then Hume's fundamental principle of causality - like causes have like effects - (Searle, 2004), becomes much less useful in practice when one tries to predict E given C. The meaning of determinism is then weakened because generally we cannot infer effects from causal circumstances with full confidence. This weakness of determinism, however, becomes more acute if we consider the following.

In the case that T is neutral in regards to the production of E (regardless of any surprises), the status of the relation between C and E during T deserves a deeper consideration. In this case E is indeed determined by C according to our definition because nothing that could have possibly occurred during T is effective in the production of E. Still, one's *knowledge* of E during T is not necessarily established even though the knowledge of C is already complete. To show why, my argument from here will take the following course:

1. In many important cases, including those involving mental states having to do with the experience of free choice, the knowledge that C does not and cannot imply the knowledge that E *prior* to the actual occurrence of E (during the time interval T that is).
2. E might be determined by C as a general principle, but a very important and perhaps critical sense of determinism is lost if in a practical sense the knowledge that C *cannot imply* the knowledge that E at any time *prior* to the occurrence of E. In other words, the logical implication of E from C is not complete once time is considered[3].
3. From 1 and 2 it follows that in some important sense, and particularly in cases that involve mental events, the gap of knowledge during T allows for a novel meaning of the act of choice and origination that are compatible with determinism yet not in the manner proposed by the standard compatibilist arguments (McKenna, 2008).

---

[3] Unless a version of temporal logic is introduced.

The first step is to single out and define those cases for which the knowledge that C does not and cannot imply the knowledge that E prior to the occurrence of E even though E is determined by C. An interesting class of such cases is the class of computational processes. A computational process can generally be specified as a mapping between a given input set and a given output set. For example: multiplication is a procedure that obtains the product of any two numbers. The input set is the set of all pairs of natural numbers and the output set is the set of all natural numbers. Another example is the prime factoring procedure that provides for every natural number the list of its prime factors. Because of their importance, general formalisms for computational processes have been developed, the primary of which is known as the Universal Turing machine (Turing, 2004). What is important for the discussion here is:

1. Given a set of possible causal circumstances <C>, and given a set of possible effects <E>, and given P to be a computational procedure that maps <C> to <E>, any instance E of <E> is determined by executing the procedure P on some instance C of the input set <C>.
2. The general specification of membership in <C> and <E>, and the specification of P are *not* generally sufficient to obtain the instance E that corresponds to an instance C of <C>. The actual execution of the procedure P in necessary.
3. Determining E from C by executing the computational procedure P invariably takes a minimum of time interval T_P. The minimum necessary interval T_P is related of course to the fastest implementation of P known to us at any given moment. Necessarily T_P is nonzero because P always involves a manipulation of physical entities in space-time (Bennett & Landhauer, 1985). Moreover, lower limits on the time of computation can be established for various computational procedures given the specifics of their implementation (Immerman, 2008).
4. For notational clarity, since the execution of P can be considered as a component of the causal circumstance that necessitates E and since P must already be specified at

the beginning of the time interval T when C is known, E can be considered as the effect of the combined circumstances (C, P).

5. Though it can be said that E is already determined by (C, P) at time t1, it is clear that prior to the culmination of the computation after at least a minimum necessary time interval T_P, there is no way to know what is E. E is unknown and cannot be known[4] even when we possess the complete knowledge of its causal circumstance well before its occurrence. The subtle point here is that while E is *determinable* at t1, it is not *determined* until t2.

When we assert that E is determined by C, an important sense of this assertion is that the knowledge of C implies E *even prior* to the occurrence of E. This is a very powerful tool of making sense of the world. If every effect has a cause, and like causes produce like effects, we can know certain effects prior to their occurrence just as we happen to observe the occurrence of their causal circumstances. But this is definitely not the case for computational processes: though E is determinable by (C, P); E *cannot be known* prior to its occurrence and thus is not actually determined. In other words, no matter how complete our knowledge of the causal circumstances that determine E, for computational processes, E will ever be a surprise to us when it occurs. We cannot predict E prior to its occurrence; though it is determinable, by definition it cannot be known until its computation is completed. What is special in the case of computational processes is exactly this: the generalization of causality, namely, like causes produce like effects, is not translatable to any shortcut in predicting E from C, the computation must be performed. An example of the difference might be in place: if I notice that lighting fire produces smoke, according the principle of causality, I can expect that at any future instance of lighting fire, smoke will ensue. Moreover, whenever I see smoke in the future I may be able to infer that fire was lighted. But this kind of reasoning is not readily

---

[4] To assert that E cannot be known, demands that out of all possible implementations of P, the most time efficient one is considered. Here, for simplicity it is assumed that this is indeed the case. Practically, however, E cannot be known even if the *theoretically* most efficient implementation is considered and not only the *practically* most efficient implementation. The issue is further developed later.

applicable to computational processes: if I multiply 2 by 3 and obtained 6, indeed 6 will be the result of any future instance when I will multiply 2 by 3, but when I multiply any other pair of numbers, say for example 4238 times 152398, unless I performed this exact multiplication in the past, I will have to go through the whole length of the multiplication procedure to obtain its unique product. The prior knowledge of the causal circumstance that determines the result does not help me here in the same manner that it helped me in the case of smoke and fire. Why is that?

Computational processes represent a unique kind of determinism in the sense that the lawful relation recognized between C and E and implemented by P is a general mapping of *a class of cause instances into a class of effect instances.* This mapping however does not provide an *a priori* knowledge of what particular instance of effect E is produced by the particular instance of causal circumstance C. To obtain that knowledge the computation must be performed and carried out to its conclusion. Furthermore, though the deterministic relation is understood as a *general* lawful relation, in the case of computational processes, the *unique* instances are the significant ones. Those particular instances, though being generally determined *a priori,* cannot be known prior to concluding their particular instance of computation. It follows therefore that in the case of computational processes, determinism is in some deep sense unsatisfactory. The knowledge of (C, P) still leaves us in darkness in regards to E during the time interval T_P while the computation takes place. This interval represents therefor what I call an epistemic gap; a gap during which the fact that E is determined by (C, P) does not imply that E is known or can be known, inferred, implied or predicted in the same manner that fire implies the knowledge of smoke even before smoke appears. It can be said if so that within the epistemic gap, E is determined yet actually it is unknown and cannot be known de facto.

The core of my argument revolves around clarifying the epistemic status of E while being computed. The question that comes to mind of course is in what sense is E determined

while in the gap? From the standpoint of the theory of determinism, supposedly E is already determined by C, because E is necessitated by (C, P), we just do not know exactly how prior to the completion of the computation P. However, a serious objection can be made here: in what sense E is determined prior to the conclusion of the time interval T_P if nothing and no one in this world can know it before the computation is concluded? If we have to answer this question *before* t2, the endpoint of the interval T_P, we have no answer we can account for. One could say we just have to wait till t2 and let things resolve themselves. But this is certainly not a satisfactory answer. It is missing something fundamental regarding how we understand things that are determinable and thus predictable by their causal circumstance.

If we accept a functionalist theory of mind describing mental processes as correlated to computational processes being implemented on the neural substratum of the brain (Putnam, 1975), (Horst, 2009), (Chalmers, 1995). Even if we do not have a complete understanding of this correspondence and how it is realized, still it is more than plausible that mental states develop in time in correspondence to the computational processes to which they are correlated. In other words, mental processes can be said to be *temporally aligned* to the neural processes that realize them. If mental processes such as deliberation and choice, are temporally aligned to computational processes realized in the brain, then, based on the two steps of the argument that have already been established above, it follows that such mental processes develop correspondingly within the epistemic gap inherent in the computational processes that realize them. Within such epistemic gap, the final content of the mental process which depends on the culmination of the corresponding computation is already determined yet unknown and in fact unknowable. In other words, within the epistemic gap T, the status of any choice, decision or deliberation being the final content of the mental process is determined yet is unknown.

It follows therefore that the experience of the outcome E not being determined while in the gap, that is, while being in a so-called process of choice prior to t2, does correspond

to an effective state of affairs in the world because E is indeed unknown while in the gap. Such experience carries some of the attributes of free will but obviously not all of them. Most notably, the experience in retrospective: “even if all conditions being the same I could still choose otherwise” does not directly contradict determinism because it reflects the fact that while in the gap, E is indeed unknown and could be anything[5]. Yet it is equally important to note that once t2 has been reached, and the knowledge of E is available, the experience “I could choose otherwise” is *only* a retrospective experience that does not correspond anymore to any effective state of affairs in the world. The knowledge of E obtained at t2 by the corresponding computational process P changed both the observer and the world towards this effect. The epistemic gap has closed, it does not exist anymore.

The introduction of the epistemic gap is important because it reconciles freedom with determinism in an elegant manner. The experience of free will turns to be far from an epiphenomenon; however, it does not carry the same meaning as it is believed to carry in our more simplified everyday mode of thinking. Moreover, the epistemic gap carries significant implications regarding how we can relate and understand the freedom and uniqueness of conscious agents.

The experience of free will and its meaning

If we accept that mental processes are temporally aligned to corresponding computational processes carried out by the central nervous system, it follows that these mental processes inherit the epistemic gap described in the previous section. It is argued here that the epistemic gap inherent in mental processes of deliberation and choice is sufficient to render the experience of freedom as corresponding to an actual state of affairs in the world, and therefore the experience of freedom *is not* an epiphenomenon. Further along

---

[5] Not anything at all but rather any one of the potentially viable instances of <E> as determined by the set of causal circumstances < C>.

this line of argument, a new meaning of freedom arises. It is compatible with determinism yet not in the sense that other compatibilist arguments bring forth. There is a real and effective sense to the freedom of will we experience, and it will be shown that this sense preserves most of how we commonly understand free will. Furthermore, following the same line of thought, the evolutionary sense of the mental state of deliberation will become evident. To show all these we first need to revisit the concept of origination that is central to the classic understanding of free will.

What is origination? Ted Honderich's definition goes as follows (Honderich, 2002):

> Origination as an event or occurrence it is the emergence or bringing-about of a mental event such as a decision or an action (1) in such a way that the opposite mental event or action might at that moment just as well have occurred although the person had remained in every respect the same and his or her situation had remained the same, and (2) in such a way that the person had control of the mental event or action and can in a certain way be held responsible for it or credited responsibility for it. Because of (1), origination is inconsistent with ordinary causation and determinism; because (2), an originated action is not merely an uncaused or chance event; origination is indeterminism plus something else. [Honderich, pp.158]

The "something else" Honderich refers to is nothing less than a power of intervention somehow in possession of a person exercising her free will. A power that is beyond causation and determinism and can in fact override any causal circumstance to the effect of achieving whatever is the choice of the person. The person is then the originator of the mental state or action she intended to bring about. It is obvious that this kind of origination is inconsistent with determinism. If we do not let go of determinism we must dismiss this idea of origination as incredible and simply false. But without the concept of origination there is not much ground to support free will as corresponding to anything real or effective in the world. Yet, not all is lost, there is perhaps a way to rescue the idea

of origination and reconcile it with determinism without entirely giving up its meaning. The key to doing just that lies in the epistemic gap.

Let us consider a mental process of deliberation that starts at time t1 with a setup of a causal circumstance C. This process is carried out consciously or subconsciously till the time t2, when a mental event or an action event E takes place. Within the time interval T between t1 and t2, the status of the resulting mental event or action is unknown because, as explained, it is within the epistemic gap. This is true in spite the fact that the determining setup (C, P) is already set at time t1[6], and therefore it can be said that E is already determinable at t1. Before time t2, however, there can be no knowledge whether E or its opposite or any other event in <E> would be the actual outcome of the process.

While in the epistemic gap, the person indeed is going through a change, a computation of a deliberative process is taking place. But as the change unfolds, for all we know, for all anyone can possibly know, either E or otherwise can still happen at time t2 and in this sense the outcome *is yet to be determined*. The epistemic gap is a sort of a limbo state where the outcome E of the mental process is both determined (generally in the sense of determinability) and not determined (particularly and actually). Since the indeterminacy of E within the gap is very strong being such that no one and nothing can know E before t2, the person involved can be said to be the originator of E in the following sense:

1. The person involved is actually implementing and performing the computational process necessary to bring about E. The combination (C, P) at time t1 is unique to the person involved. As such, the process of bringing about E is a unique event irreproducible or repeatable elsewhere or even by the same person at a later time.

[6] This is a simplification of a real life situation where both the causal circumstance C and the mental procedure P are inherently dynamic.

2. The person involved is the first to know E[7]. No one and nothing can know E earlier than this person. If the mental state or action E does have any effect at all in the world, what brings about this effect is the person involved by means of carrying out the necessary computational process.
3. At any point prior to time t2, if the person or any other agency would have engaged in somehow guessing the outcome E, E could still be otherwise, and this statement, that E could still be otherwise, cannot be refuted. At any point later than t2, this is not the case anymore. E could not have been other than it is. The determinability of E and the actual determination of E are again aligned. As was said before, the knowledge of E obtained at t2 by the computational process P changed both the person and the world towards this effect. The gap does not exist anymore.

Let us further explore this new sense of origination and how it fares in comparison to the classical meaning of origination. In distinction to the classical meaning of origination the power of supra mundane intervention goes away and with it the discrepancy between origination and determinism. The demand that the mental event or action E could be otherwise in such a way that the person had remained in any other sense the same is nonsense since whatever choice is made, the person, being the author of the choice, changes accordingly and this change cannot be reduced to this or that aspect of the person's mind. No choice can possibly be singled out and isolated in such a manner that the person originating that choice stays unchanged in absolutely any other aspect. Apart from these differences however, we can still claim that while in the process of deliberation (i.e., within the epistemic gap) the outcome is unknown and therefore could be E or the opposite of E or any other outcome generally consistent with (C, P). Moreover, we can claim that the knowledge brought forth by the person at t2 be it a mental state or an action is unique and original. This uniqueness and originality are enough to lend substance to the authorship of the person and therefore to the origination

[7] There are important exceptions to this point and what follows from it as will be discussed later.

at the core of her choice. Also, at least in some sense, the author carrying out the process can be credited or held responsible to the mental state or action E, him being the agent without whom E could not be brought forth.

The kind of freedom argued for here is not rooted anymore in some mysterious supra mundane intervention that defies determinism, but rather in the very mundane process of bringing forth the genuine and unique knowledge inherent in E that was not available otherwise. It can be said that in any such act of freedom a person describes and defines herself anew. When making a choice, any choice, a person may become conscious to how the choice defines who he is at the moment it is made. He may become conscious to the fact that the knowledge of the choice irreversibly changed him. Clearly this moment of coming to know one's choice is indeed a moment of surprise and wonderment, because it could not be known beforehand what this choice might be. If it was, this wouldn't be a moment of choice at all and one could have looked backward and find when the actual choice had been made. At the very moment of coming to know the choice that was made, reflections such as 'I could have chosen otherwise' are not valid anymore. At that very moment the particular instance of freedom within the gap disappears and responsibility begins. This responsibility reflects the manner by which the person was changed by the choice made.

The main objection that can be brought up here is of course regarding the meaning of freedom. It might seem that the kind of freedom described here is a reduced kind of freedom in comparison to the more classic one that defies determinism. It could be argued this is not freedom at all, that it does not correspond to our genuine experience of choosing one path over another. Though it is a fundamentally different kind of freedom, I do not think it is reduced. On the contrary, it is a more mature and deeper kind of freedom and I will explain why.

First, it is coherent and consistent with the wider understanding we have about the world involving the concept of determinism. Second, it is consistent with our experience of freedom while we are in the process of deliberation. Something not trivial is taking place in the course of determination that corresponds to our sense of authorship. Third, we can now argue that our choices are effective in the world and not epiphenomenal. Furthermore, evolution in general and each person's unique experience and wisdom are critical factors in shaping the mental processes of deliberation. If every choice is the bringing forth of a genuine and unique knowledge and if the way this knowledge is computed involves in fact a sum total of all the knowledge accumulated by evolution and the unique life experience of the person, it clearly follows that deliberation and choice are instrumental in the evolutionary motion of the individual and therefore of the species[8]. As such, our sense of exercising our freedom is entirely coherent with evolution. Moreover, our sense of free will corresponds to our evolutionary uniqueness; the novelty and uniqueness of a person's choice do have an effective evolutionary sense[9]. With these three points, the basic riddles regarding free will described in the introduction are resolved.

There is more subtlety here worth mentioning: one would ask how, in a worldview reigned by determinism, is it possible to find any real sense to freedom? To answer this, a deeper understanding of the epistemic gap is needed. We know that within the epistemic gap things unfold from C towards E in a deterministic fashion. There is no mystery in the gap that somehow circumvents determinism. Within the gap, E simply cannot be known because lower-level processes whose combined and lawful conclusion is necessary to the production of E are yet to take place. It is only an illusion propagated within our conceptual minds for millennia or more that we can use the fact that the whole process is deterministic to the effect of knowing something meaningful about E prior to the point in

---

[8] The point demands further explanation which is beyond the scope of this article.

[9] Of course not all choices are equally effective or equally important, but every choice is potentially significant.

time that E actually happened. If we leave this illusion behind, what we are left with is not a deterministic continuum but rather something entirely different as the supposed continuum is punctuated by innumerable epistemic gaps. Within these gaps we do not know and therefore we are free from the claws of determinism.

When a person comes to know something, this knowledge becomes part of the causal circumstances that determine his future mental states and actions. The person's future state or action can be said to be constrained by this knowledge. When in a state of *not* knowing, within the epistemic gap, that is, the person's future mental states and actions are locally indeterminate. The person does not have control over her future states or actions (as it is expected in the classical understanding of free will), yet nothing other than the person herself, while in the gap, has a determining power of control over her future mental states or actions either. The process unfolding within the gap is bringing forth the future state or action. This process can be identified with the very sense of personhood at that moment. Indeed, the person is not independent from its causal circumstances, yet while in the gap his future is not *a priori* fixed by these causal circumstances.

The deeper and more mature sense of freedom described here has to do with these gaps of *not knowing*, in the true sense of mystery they invoke even in the very simple details of life[10]. Freedom of the will in its classic sense is a confusion arising from our deeply ingrained need for control. The classic problem of free will is the problem of whether we are inherently able to control a given life situation. Origination in the classic sense is the ultimate control status. The sense of freedom described here leaves behind the need for control. The meaning of being free has to do with (consciously observing) the unfolding of who we are while being in the gap, the transition from a state of not knowing into a

[10] There is an even more subtle dimension to this line of thinking having to do with the definition of an event. From the standpoint of an agent, events have a certain granularity. A mental event is not reducible to the cumulative set of neuronal firings. While determinism is intact at all scales down to the minutest, the granularity of events makes the epistemic gap an actual feature of the agent's mental processes.

state of knowing, that is. It can be said that it is not the choice being originated by me but rather it is I, through choice, who is being continuously originated as the person that I am. The meaning of such freedom is not centered around control but rather around the novelty and uniqueness as they arise within each and every choice as one's truthful expression of being.

The way we are free

In this section I will explore the consequences of the epistemic gap in real life situations and how these consequences might influence a person's freedom. While freedom in its classical sense is considered to be an attribute of a person and independent of anything else, the kind of freedom described here is not independent from the circumstances a person finds herself in. The freedom described here depends on certain conditions that allow the existence of the epistemic gap. As we will see shortly these conditions do not depend solely on the person, and as they are dynamic in nature. Importantly, therefore, the level of freedom a person enjoys is dynamic according to circumstances.

The easiest way to imagine this is by the following thought experiment: With the help of a futuristic technology, a person's mind is entirely simulated by a super computer. The sense perception data of the person is accurately simulated by the same super computer and is fed to the person's simulation in such a manner that the simulation accurately replicates the person's mental states[11]. The only discernible difference between the simulation and the person is that the simulated duplicate of the person is vastly faster. At time t1, some causal circumstances C appear to exist for both the person and its simulated duplicate. The computation time involved in computing E from C is vastly longer for the person compared to his simulated duplicate. While it takes the person time T to compute E, the duplicate computes E in a negligible fraction of that time. It turns out that the person's response at t2=t1+T is not genuine anymore (and therefore not free in the sense

[11] The experiment assumes that subjective metal states do arise in the simulation. The point of the experiment however does not rely on this assumption.

argued before) because shortly after t1, as soon as the simulated duplicate produced E, the epistemic gap had collapsed. The person's response at t2 was already determined by a faster version of himself.

Remarkably, the epistemic gap does not collapse entirely because the simulated duplicate is also performing a computation that involves physical processes. The gap is only vastly shortened. But this is not the most significant observation. More significant is the fact that the very location of the gap regarding the unfoldment of events, shifted from the location of the person to the location of his simulated duplicate. In the scale of time relevant for the simulated person, he enjoys the freedom arising from the gap. Sadly this is not the case for our flesh and blood person: the mental states or actions he produces are already determined by his simulated duplicate long before he had the time to produce them. His future states or actions are effectively determined prior to him realizing it. His mental states and actions are fully predictable by his faster simulation. One could effectively know with confidence what he is going to do.

As discussed before, for every given situation of C unfolding into E by means of a computational procedure P, the duration of the epistemic gap is defined by the fastest available implementation of P. The gap if so has in principle a single location[12]; the location where the fastest known implementation of P is executed to bring E as a result. Only at that location freedom is realized. Additionally, in our thought experiment, it is assumed for simplicity that C becomes available to the person and the simulation simultaneously. In actual life situations this might not be the case. One implementation of P might get the necessary knowledge of C earlier than other implementations and by that become the first to produce E. An interesting result regarding freedom follows: a person's choice is free if and only if she is the first to produce E. In practical situations this can be the result of either having C available the earliest, or having the fastest

[12] With the exception that equally fast implementations are concurrently executed in different locations, but this case is not essential to the subject matter.

available implementation of P, or a combination of both. Generally, it can be said that a person's freedom depends on conditions set by his circumstances, namely how information (availability of C) and computational means (fastest implementation of P) are distributed within the circumstances of the deliberating person[13].

As strange as it might sound at first, this situation is not unfamiliar to our everyday intuitions. An essential aspect of the social fabric we are part of is that we have a theory about the minds of others. This theory amounts to us trying to simulate how a mind other than us might respond to given circumstances C. In short, we continuously try to figure, at least approximately, what E a mind other than us would produce given C. The whole point about figuring E is figuring it *prior to the time* that the other person will produce it, because by predicting just that beforehand, one gains a significant advantage in social interactions. Here it becomes clear that the thought experiment is just an idealization of our intense interest in applying and improving our theories of mind as part of our social interactions. In everyday situations we are in a continuous race to stay ahead of others in manifesting various effects from given circumstances C. We try to be genuine in our responses, while trying to guess others' responses prior to them acting. This is why it is not an unfamiliar experience that when we are in contact with persons that are slower than us in reading the situation and computing proper responses, we experience an expansion of our freedom and genuineness, while when we are in contact with persons that are faster than us, we experience that our freedom diminishes. We feel as if the other guy often figures what we are going to do or how we are going to respond even before we know. How free are we seems to depend on the dynamic circumstances we are in. In certain situations, we are freer than in other situations and this is often reflected in our experience of freedom. Who is the agency of highest freedom at a given situation depends on where the epistemic gap is located and this depends on the distribution of

[13] Further complications arise here if we take into account relativistic considerations regarding time and distance.

information and computational means among agents within a given system of interactions[14].

Freedom can then be understood as a dynamic property closely related to computational means and distribution of information. A person cannot expect to be free in the same manner in different situations. When one's mental states and actions are often predicted in advance by others who naturally use these predictions while interacting with him, one's freedom is diminished to the point where no genuine unfolding of his being is possible at all. The person becomes a subject to *a priori determined* conditions imposed on him. He will probably experience himself being trapped in a situation that does not allow him any genuine expression. He loses the capacity to originate because somebody or something already knows what will happen. In everyday life, what rescues our freedom is that we are all more or less equally competent in predicting each other's future states and actions. Furthermore, the computational procedures that implement our theories of mind are far from accurate or complete. They are more like an elaborate guess work with some probability of producing accurate predictions. Within such circumstances, freedom is still often viable. But this may soon radically change by the advent of neural and cognitive technologies. In fact, it is already in a process of a profound change.

Further Implications

If we accept the proposition that personal freedom is relational and that it depends on the distribution of information and computational resources within the circumstances of the person, we arrive to an interesting yet concerning question regarding the interdependent relations between freedom and technology. It seems obvious that technology will profoundly influence the conditions that are necessary for the epistemic gap to arise. Here, in the concluding section these influences are explored in some more detail.

---

[14] For simplicity it is assumed that agents are consciously intentional and inclusive of all the information and computational means they possess. A much deeper understanding of the nature of knowing and its relation to temporality and consciousness is necessary to further generalize the arguments presented here.

The influence of technology on the epistemic gap and therefore on the degree of freedom a person may enjoy seem to be a combination of three major scientific and technological factors:

1. Advances in neuroscience, cognitive science, behavioral psychology, evolutionary psychology and other scientific disciplines provide better and increasingly more accurate theories of mind. Such theories allow us to progressively approximate the conditions of the thought experiment described above, namely, to advance towards building accurate simulations of persons.
2. Advances in computing and computer technology, together with the knowledge mentioned in 1, may enable the creation of actual computer simulations of persons or parts of their minds. Eventually, powerful mind simulations might be fast enough as to accurately predict a person's behavior long before that person acts.
3. Information technology will soon allow the acquisition and distribution over information networks of vital information about every person's life. Such information will become increasingly detailed down to a continuous sense-motor logging, and in a further future it might include details about mental and emotional states as well. The erosion of privacy is another ongoing trend that will make it more and more difficult to keep such sensitive information discreet. The effects of information technology are in gaining faster access to the causal circumstances that affect a person's future choices and actions, and also in gaining the information needed to accurately model a person's mental states.

In simple terms, the combination of all these factors will make persons much more predictable to others and will have the effect of overall diminishing the number of instances of operating within an epistemic gap and therefore the conditions favorable to personal freedom. The implications on freedom as described here are that in the future

people able to augment their mental processes to enjoy higher computing resources and more access to information will become freer than others who enjoy fewer computing resources and access to information. Persons who will succeed to keep sensitive information regarding their minute-to-minute life happenings and their mental states secured and private will be freer than those who are not. A future digital divide will be translated into a divide in one's freedom, that is, the freedom to self-determine as opposed to being manipulated into predictable behaviors.

As our understanding of mental states and their physical implementation in the brain grows together with advances in computing and information technology, personal freedom will become increasingly dependent on technology. The more advanced the technology available to a person the more she might gain opportunities for genuine expressions unpredictable by others. If this implication is to be taken at face value, protecting the freedom of persons in a democratic society will have a lot to do with the fair distribution of information and computing resources and with effective protection of privacy. On a more speculative note, as future artificial minds or highly augmented biological minds might gain a computational advantage so great that no biological mind would be able to compete with, the quest for freedom might become an imperative that drives futuristic technologies towards the convergence of man and machine.

Another less obvious implication of technology on personal freedom is the possibility of analyzing and modifying the mental processes involved in computing our choices towards an optimized state of freedom. In the future we might augment our minds towards increased freedom not only by means of increased computing resources and information access, but also by upgrading the very way we compute. If we gain enough understanding how mental processes arise, we might be able to devise general methods of ridding them from unnecessary redundancies and increasing their aesthetic and creative qualities. To some extent we already do that when we self reflect on how we engage in our own experiences, and consciously try to augment our views and improve our future

responses. In the future, the sciences of the mind could provide us with powerful insights and tools to help us optimize our freedom as well as other qualities.

Bibliography

Bennett, C. H., & Landhauer, R. (1985, July), 'The Fundamental Physical Limits of Computation', *Scientific American* , pp. 48-56.

Chalmers, D. (1995), '*A Computational Foundation for the Study of Cognition',* Retrieved March 19, 2009, from http://cogprints.org/319/0/computation.html

Honderich, T. (2002), *How Free Are You?* (2nd Edition ed.), New York: Oxford University Press.

Horst, S. (2009, March 21), '*The Computational Theory of Mind',* Retrieved March 19, 2009, from Stanford Encyclopedia of Philosophy: http://plato.stanford.edu/archives/spr2009/entries/computational-mind/

Immerman, N. (2008, September 21), '*Computability and Complexity',* Retrieved March 24, 2009, from The Stanford Encyclopedia of Philosophy: http://plato.stanford.edu/archives/fall2008/entries/computability/

Lowe, E. J. (2000), *An introduction to the philosophy of mind,* Cambridge University Press.

McKenna, M. (2008, September 21), '*Compatibilism',* Retrieved March 21, 2009, from The Stanford Encyclopedia of Philosophy: http://plato.stanford.edu/archives/fall2008/entries/compatibilism/

Putnam, H. (1975), *Mind, Language, and Reality,* Cambridge: Cambridge University Press.

Searle, J. R. (2004), *Mind, a brief introduction,* New York: Oxford University Press.

Searle, J. R. (2001), *Rationality in Action,* Cambridge: MIT Press.

Turing, A. (2004, November 5), 'On Computable Numbers, With an Application to the Entscheidungsproblem', *Proceedings of the London Mathematical Society* , pp. 230-265.