

On Some Arguments Concerning Freedom and Determinism

Davor Pećnjak

In his text *Human Beings as Autonomous Persons*, Professor R. C. Pradhan (1999) claims, among other things, that human beings are not a kind of machines and that machine description of human beings and their mentality could not be fully given. One of the claims in his article is also that human beings have creativity. That could mean a lot of things. Of course, what comes first to the mind is the concept of creativity connected with the arts, science, literature, innovations in technology and so on. But I shall not dwell on this and shall not try to explain how human beings can produce new things in these areas. I have something else on my mind. More generally, if human beings are indeed autonomous persons, they have to be so in a very fundamental sense. Namely, they have to be able to cause, at least some of their actions, on their own. It means that at the very fundamental level, human beings should have freedom of the will and freedom of the action. This notion of freedom, for genuine autonomy of human beings, should be, in my opinion, construed as the libertarian sort of freedom. Accepting this, I think, would fit to Pradhan's views. So, in this article I would like to explore just a few ways how we can argue for this position. Let me say first a few words about some possible libertarian construals of freedom.

Libertarians, whichever variant of libertarianism they accept, are incompatibilists regarding the concepts of freedom and determinism. Incompatibilists hold that either determinism is true or the thesis of freedom is true, but both cannot be true at the same time. Libertarians, of course, think and argue that the thesis of freedom is true. Other incompatibilists who think that the thesis of determinism is true are hard

determinists. The most persuasive arguments for incompatibilism were given by Ginet (1966), van Inwagen (1975, 1983) and Lamb (1977) and they are widely accepted by libertarians¹. Let me briefly state so-called van Inwagen's first and third argument: "If determinism is true, then our acts are the consequences of the laws of nature and the events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our present acts) are not up to us." (van Inwagen, 1983: 16). The third argument is more formal and complex (van Inwagen, 1983: 93-104); it includes some modal reasoning and inference. We shall have to introduce some symbolism. The conjunction of all the natural laws is abbreviated as „L“. The whole state of the universe at some moment in the past is abbreviated by „P₀“; „P“ stands for any true proposition. „Np“ means „p, and no one has, or ever had, any choice about whether p.“ „□“ is a standard symbol for „necessarily“. Then, symbolized in such a way „((P₀&L)→P)“ is the thesis of determinism. It means, obviously, that conjunction of statements that completely describe a state of universe at some particular moment (in the past) and all the laws of nature, entail everything that happens after (and before) that particular moment in time. Besides the usual rules of propositional logic, van Inwagen uses the following two rules of inference also:

Rule α: If □p, then Np.

Rule β²: If Np and N(p→q), then Nq

Now, we can proceed:

- | | |
|------------------------------|----------------------------------|
| 1. ((P ₀ &L)→P) | thesis of determinism |
| 2. □((P ₀ →(L→P)) | from 1 by the law of exportation |
| 3. N((P ₀ →(L→P)) | from 2 by rule α |
| 4. NP ₀ | premise |
| 5. N(L→P) | from 3, 4, and rule β |
| 6. NL | premise |

7. NP

from 5, 6 and rule β

So, if determinism is true, no one can do otherwise than in fact he did. Determinism is not compatible with freedom. This is a conclusion of these van Inwagen's incompatibilistic arguments. If determinism rules the universe and us as beings in it, we would be like the machines under the laws of nature, however complex our behaviour is. Libertarians accept these arguments for incompatibilism and hold that determinism and freedom cannot go together. But freedom in the fundamental sense is possible, say libertarians. We can say that freedom of the will and freedom of the action exist when an agent is the ultimate source of his action or when an agent can do otherwise than he in fact did, under the same conditions. If I am the ultimate source of my action³ then there are no factors beyond myself which are responsible for the production of that action of mine. There are no other factors outside me which are the main influences of making that action; there are no external factors that could fully explain the action which I have done. This action is in my control. But if it is so, then it seems that I could refrain from doing that action also. That would mean that, under the same conditions, I could have done something other than perform the action in question. So, if we start analyzing⁴ the notion of being an ultimate source of some action, we arrive that it amounts to the possibility of doing otherwise under the same antecedent conditions. Reversely, if we say that it is within my power to do otherwise than what I in fact did, it means that I am in such a position that finally no constraint is present in doing that action. Further, it amounts to me being the ultimate source of my action. So being the ultimate source of action or when an agent can do otherwise than he in fact did, under the same conditions, could amount to the same thing in libertarian view. Though I think this is so, not each and every libertarian thinks this is the case. Eleonore Stump (1996), for example is perplexed with the so-called Frankfurt-style cases and would like to defend libertarianism against such scenarios claiming that alternate possibilities are not required for freedom of the will and action but that only (ultimate) sourcehood is. In these scenarios, due to Frankfurt (1969), an agent is in a position with respect to an action A, that he could not do it otherwise, i.e. he cannot avoid doing A, but A can be done by the agent himself or under

coercion. I shall present slightly modified story which goes like this⁵. There are two agents, X and Y. Elections are next Saturday and there are two parties for which voters can vote: North party and South party. Y would like that X vote for the North party. X is still in doubt and yet does not know for which one he will vote for. He can decide both ways. But Y wants very much that X votes for North party and decides that he would not let X vote for the South party. For that reason, to ensure that X votes for the North, Y implanted, during the night when X was asleep, a device in X's brain which will fully take control over X, would he decide or show a sign that he would vote for the South party. Activation of that device will completely move and control X that he votes for the North party (if he decided or show a sign that he would vote for the South party). Of course, X can decide by himself to vote for the North party. This is a course of events that Y would indeed prefer to because he would like minimally to interfere and in this case he would not have to interfere at all (device would not have to be activated at all if X votes for the North by himself).

So, action A (voting for the North party) is for agent X unavoidable. Still, if the intervener does not intervene, we would say that X has done A by his own free will (and we would ascribe moral responsibility for action A to the agent X) even though he could not have done otherwise – in the case that X showed a sign not to do A, intervener would take steps and coerce X into doing A (by activating implanted device which fully takes over and controls X) In this latter case, we would not say that X has done A by his own free will – he was manipulated and coerced. He was just a puppet in the hands of an intervener. What X did, stemmed in fact from the will and intention of an intervener (and we would not ascribe moral responsibility for action A to the agent X). Action A proceeded in this latter case from the will and intention of an intervener. Since X had been fully manipulated, the will and intentions of X had been removed in fact, and non-existent. X had only been a tool for an intervener – like a remote control device. So, in this latter case, there is no question whether X could have done otherwise, because, I would say, he did not do anything. He is not doing anything because he was completely under the control of an intervener and became a device and not an agent. Since actions are in virtue of intentions (together

with beliefs and desires) that produce them – these two ways of doing A are different. In the first case A is produced by the intentions of X, in the second case they are produced by the intentions of an intervener (and the will and the intentions of X does not have any role).

So there is an alternative possibility for X: he does A on his own free will or he does nothing in fact. Alternative possibility consists in this, I would say: X can decide to do A (and he proceeds to do A and does A - intervener does not intervene and this course of events is like that intervener does not exist at all) *or* he can decide not to do A, and in this latter case, intervener intervenes and fully takes over A and in fully taking over, intervener in fact erases A as an agent. Human being A exists but not the agent A. A becomes just a tool for exercising intervener's will and intentions. Once again, there is alternative possibility for A: doing with A on his own free will or doing nothing.

Stump also claims that doing A on his own and doing A when an intervener activates the implanted device, so action A is performed under coercion, are in fact numerically the same action; so there is no alternative possibility for X. Now, I would like to show that these two are not numerically the same action.

Human actions are clearly events. If we use Davidson's (1985) criterion of individuating events then, clearly, doing A on his own and doing A under coercion are not the same actions, and moreover, they are not numerically the same actions. Let me cite Davidson (1985: 179): „... events are identical if and only if they have exactly the same causes and effects. Events have a unique position in the framework of causal relations between events in somewhat the way objects have a unique position in the spatial framework of objects. This criterion may seem to have an air of circularity about it, but if there is circularity it certainly is not formal. For the criterion is simply this:

where x and y are events, $(x = y \text{ if and only if } ((z)(z \text{ caused } x \leftrightarrow z \text{ caused } y) \text{ and } (z)(x \text{ caused } z \leftrightarrow y \text{ caused } z)))$).

Circularity can be nevertheless charged as such: we define some event in terms of its causes and effects, but causes and effects themselves are in most cases events. So, events are defined in terms of other events and we have immediate circularity. However, we may accept Le Pore's (1985) defence of Davidson's criterion: events are very basic and as such we cannot analyze it more, so it is not possible that we can have a non-circular criterion for events. If we accept this and so say that events are individuated by their causes and effects we can say the following.

Because X's doing A on his own and doing A under manipulation and coercion of an intervener have different causes (even if effects are the same), these are two different actions and events. There is an alternative possibility for A then: there is doing A on his own or doing nothing. Doing nothing is in fact making A by an intervener manipulating and completely controlling X.⁶

What also comes to mind when we think about possible machine description of human beings is Lucas's (1961, 1970) argument against mechanistic view of the mind, which is revived by Penrose (1994). Lucas uses the notion of Gödel's incompleteness theorem (see for example Smith 2007). In a nutshell, this theorem says the following: if the formal system which is strong enough to yield basic arithmetics is non-contradictory, i.e. consistent, then it is incomplete; but consistency cannot be proved within the system. This means that not every formula which is the formula of the system can be syntactically proved. In other words, there can be both formula A and its negation, formula non-A, that neither can be proved. But, what human beings can do in such a case is to see and understand that one of these two formulas is true and that the other is false. Since a machine could only arrive to a formula to establish its truth through a syntactic procedure, machine cannot establish this because there is no such a procedure. In Lucas's (1970: 136) own words: "...in so far as (computer's) procedure is in accordance with definite rules, the Gödel method can be used to produce a formula which the computer according to those rules, or the physical system according to its description, cannot assert as true, although we, standing outside of the system, can see it to be true." So, here

we have something that human beings can do what machines cannot. Of course, one instance of something that we can do and the machines or any physical system, under some physical description, cannot do, in principle is enough to establish a difference in kind (and not in degree only). If it is so, then human beings are different from machines: machine description or full computational description of human mentality (and perhaps action) cannot be done. On the other hand, even if it is so, it would not be much of a result for human beings if it can be applied when we think about some Gödelian formula only. We rarely, if ever, think about some Gödelian formula. Even most of the mathematicians rarely think about some Gödelian formula. They think most of the time about proving theorems or how to calculate what is calculable. Most members of the humankind never heard even about Gödel, let Gödelian formula alone! So if there is just one type of performing an intellectual task and which is, in fact, seldom performed by a small number of human beings – that is not much of a result! It would mean that all other things we do – whether intellectually or physically, are in fact describable mechanistically. So, if we are in all our daily life, and even in all other formidable intellectual actions, mechanistically or computationally describable – we are not, then, (much) different from the machines. This difference in kind, though real, would not mean very much. And, if really a complete mechanistic description is enough, then it seems that this description is a deterministic description. If real, then we would be determined beings in our thinking and doing and we would not have freedom of the will and action in the sense of having a possibility and ability to do otherwise in the very same circumstances or to genuinely start a new causal process, previously undetermined.

Much better would be, if our mental processes, at least higher mental processes, or at least a subset of higher mental processes, are such in structure that Gödelian argument can be applied to them to describe them. It would mean that very structure of such a kind of mental processes is such that they cannot be fully formalized. If some mental process is such that we need to describe it by a certain Gödelian formula, then that process would be a process which does not have mechanical structure that can be computed. If this would be so, then more of our thinking would be such

that it is not something that a machine can do. Though it is good, in my opinion, that some of our mental processes, and, especially reactions, are such that they are quick and mechanical, we would be different – significant part of our mentality would be different - from the machines in a more substantial way than if we are such when thinking about Gödelian formula only. This would be especially important for our deliberative processes. If this argument could be applied to our deliberative processes, or at least to parts of deliberative processes, then it would mean that they are, or parts of them, non-computable or non-computational processes (Penrose 1994: see chapters 2 and 3 for extensive discussion). If such, then there will be no deterministic algorithm for our deliberations. They would be mental actions freely done. If some action springs from the results of such deliberations, as they often do so, as for example, going to the cinema, then these actions would also be free actions deriving their freedom from freely made decisions in deliberation. So, for those who endorse such Lucas's Gödelian arguments, they can be useful in a twofold sense. First, they show that human beings are not machines; and second, in a slightly stronger reading, they can show that, at least in some aspects, we enjoy freedom of the will and freedom of the action. I think that this is congenial with how Penrose (1994: chapters 2 and 3) revived and refined Lucas's ideas.

Professor Pradhan (1999) also accepts Kant's notion of humanity which treat human beings to be ends in themselves and, as moral agents, there is also a requirement that they have free will and freedom of action. Regarding Kant, in some philosophical matters, Kant was a critic of Hume. I shall not present Kant's critique of Hume to argue for libertarian notion of free will, but I would like to do something else here.

In the remainder of this article, I would like to examine what possibly can follow if we accept some concepts and notions of Hume's or Humean idea of causation, regarding free will problem. First of all, I would like to make clear that I am not a Hume scholar so I do not pretend that this would be a thoroughgoing and deep analysis and also that this is some kind of a work in progress concerning my overall interest in the problem of free will. So, here I shall analyse only what Hume (1993) has to say about

causation and free will in his *An Enquiry Concerning Human Understanding*. Detailed comparisons between this and *A Treatise on Human Knowledge* I have to leave to Hume scholars.¹

Let me just briefly remind ourselves about some common notions which are very well known about Hume's position on free will problem. Regarding *An Enquiry Concerning Human Understanding*, Hume is a compatibilist in free will problem. It means that he thought, along with some other figures from philosophy like Hobbes or Mill, that the concept of freedom is compatible with a concept of determinism. According to such classical compatibilists, freedom would consist in an ability of an agent to do the desired action and that there is no any constraint or impediment for exercising what an agent wants to do. Even if it is fully determined by previous events which are out of control of an agent, if agent has an ability to do X and there is no any constraint or obstacle for doing X and an agent does X – he did it freely, according to classical compatibilists.

I would neither attack compatibilism in this talk (though I think it is untenable position, but arguing against compatibilism I leave for another occasion) nor give some partial arguments against it. I would like to show that Hume's ideas can be effectively used to support libertarianism in fact, and not his compatibilism.

The problem of causality is one of the fundamental philosophical problems and there is no any theory of causation which is such that it would be accepted by most philosophers – which is hardly surprising! Just to mention that before Hume, very similar theory of causation, but with rather different consequences thereof, had been given by Malebranche (1997) so perhaps this kind of theory should be called Malebranche-Humean. But I shall talk only what Hume says about the causation.

According to Hume, and Hume is a radical empiricist, we experience various sequences of events around us and we also experience regularities

concerning these and other (sequences of) events. Namely, one kind of events follow another kind of events, or we can say that some events come before some other events. That what is interesting in it is that we observe that always or almost always, one kind of events comes before certain another kind of events. After the event X, event Y follows (X is prior to Y). We experience that almost always this is the case. It is not that we, at some moment, suddenly observe that it is true that event X happened but that it is followed by entirely different sort of event Z instead of Y. After X, again, Y, follows. From this, according to Hume, our mind forms the habit by constant observing that any time it becomes conscious that the event X happens, we expect that the event Y will follow and not any other kind of event Z. So, we observe in experience regularities in sequences of events. Hume goes on to say that our mind forms the idea of necessity from constant observing the same patterns of sequences of events: it seems to the mind that there is a necessity which makes that an event of the type X *must* be followed by the event of the type Y and not something else. The event X which comes first we call a cause and the event Y which follows and comes after X we call an effect. We treat this that there is a certain necessity such that when an (cause) X happened that there must be an (effect) Y. In other words, we say that happening of a cause X „necessitate“ happening of an effect Y (and nothing else). But, Hume says also that, in fact, we do not observe and we do not experience necessity by itself. We do not observe something else, beside the regular sequence of events, in which necessity would consist further. We do not observe and experience something additional which would be that necessity over and above the regularity of events. We cannot say what else is observed in this sequence which would „necessitate“ that X was followed by Y. As it was already said, we observe only sequences of events and their regular repeating. „Necessity“ we do not observe. It seems that there is no any other thing, besides repeating of regularities. There is no *further observation that it must* be so!

Regarding that we constantly observe the same or very similar sequences of events (for example, that X is followed by Y), we form a firm habit that when we hear about or observe an event X, we immediately think about Y. We extrapolate from this habit and we form the concept of „necessity“ – X

is „necessarily“ followed by Y. It seems to us then, that it could not be otherwise than after X „necessarily“ follows Y: or we can say it must be that Y follows X. Nevertheless, says Hume, our experience does not lend us support for this conclusion, because, except the sequence, we do not observe anything else – we do not observe and we do not have an experience what would be what makes that sequence in question that it must be like that. We do not observe any necessitating factor. This is quite in accord with Hume's views as a radical empiricist. To claim that there is some kind of necessity, we should have an experience of it, and that would allow us to have a right to claim that there really is necessity at work beyond the regularity of events. So if we do not observe any additional factor in which necessity would consist that would tell us that it *must* be that X is followed by Y, then, according to radical empirical philosophy, we should say that there is not the case that X must be followed by Y. We are lucky that regularities repeat because that makes our living a lot easier but it, metaphysically speaking, does not have to be that X is followed by Y. So, perhaps it is possible that tomorrow we shall observe that X is followed by Z instead of Y. If there really is such a possibility, then we can infer that possibly, from the same or similar sequence of causes, another, different sequents, which would be different effects, can follow. If we apply this on human decisions, will, and actions we can say that from the same initial antecedent situation (mental causes), some different effects can follow. If it is so, that would mean that from the same antecedent situation, it would be possible for an agent to do otherwise then he in fact did. If it is possible for an agent to do otherwise in the same situation, then it is enough for libertarianism.

So, Hume as a radical empiricist, has to allow this kind of possibility when he says that „necessity“, like something additional and over and above the sequence of events, what would make that this sequence must be such-and-such, is not something that is ever observed. If it is not observed ever, then there is no such thing according to the viewpoint of the Humean or radical empiricism.

It is true that Hume nowhere says this explicitly, but regarding radical empirical standpoint, to claim that there exists „necessity“ in empirical matters, we should have to notice and observe that necessity (as something over and above regular sequence of events). If we do not observe necessity as itself, then there is a possibility (or some probability) that a certain sequence which begins in the same way as always, at some moment t starts to unfold differently than before, differently than it was usual and that we, according to our habit, expected.

If it is so, that would be more than a welcomed option for a libertarian: under the same conditions until moment t , it could be otherwise. In this way, fundamental requirement of libertarianism could be satisfied. It is because libertarianism requires that an agent should be able to do otherwise than in fact he did.

On the other hand, Hume also says that similar things happen with our ideas and thoughts. He claims: it is so that one idea or thought follows another one, and it happens also in a regular series or sequence. When one idea or thought happens, let's say a thought A , then this idea is followed exactly by an idea or thought B . When there is a thought A , then there is thought B . We form a habit to have this.

What I would like to say is that it is often true, but generally it is not close to truth. Let me illustrate this by the example from everyday life.

For example, when I happen to have a thought or happen to think about a race car Tyrrell 005, thoughts that follow this thought almost never are the same. Once it is a thought (which immediately follows a thought about Tyrrell 005) about a driver named Jackie Stewart who was a world champion in formula one with this car in 1973. Next time it is a thought that racing team Tyrrell was sponsored by the oil products factory Elf. Third time it is a thought that that car was nicely blue coloured. Fourth time it is a thought that it would be nice that even I could participate in car races with this car. Fifth time I remember my childhood. It seems to me that other people also have such experiences that one and the same thought is immediately followed by completely different other thoughts. Even when

we think about billiard balls, it does not have to be case that when I think that one ball struck another, that other one starts to move. The thought that immediately follows the thought about the collision of billiard balls may be so that I think that both of them stop, that one of them or both go to air, that one of them is turned into an electric organ. It is not true that one and the same thought is regularly followed by another and the same thought.

Here we can come to similar conclusion as before: an agent could do otherwise in his thinking: his deliberation is open and so the fundamental requirement of libertarianism could be satisfied: an agent could do otherwise then what he in fact did.

In this article, because I am prepared to defend libertarianism, I examined a few possible libertarian interpretations and I hope that they are congenial with Pradhan's view that human persons are not the machines and that they are autonomous beings in a very fundamental sense.

Acknowledgment:

References

- Davidson, D. (1985). "The Individuation of Events", *Essays on Actions and Events*, D. Davidson, (Oxford: Clarendon Press), 163-180.
- Finch, A. & Warfield, T. (1998). The "Mind Argument and Libertarianism," *Mind*, Vol. CVII, No. 427, 515-528.
- Frankfurt, H. (1969). "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy*, Vol. LXVI, No. 23, 829-839.
- Ginet, C. (1966). "Might We Have No Choice", *Freedom and Determinism*, ed. K.Lehrer, (New York: Random House), 87-104.
- Hume, D. (1993). An Enquiry Concerning Human Understanding (<http://18th.eserver.org/hume-enquiry.html>)
- Huoranszki, F. (2011). *Freedom of the Will: A Conditional Analysis*, (New York and Oxon: Routledge)

- Kane, R. (ed.) (2002). *The Oxford Handbook of Free Will*, (Oxford: Oxford University Press)
- Lamb, J. (1977). "On a Proof of Incomptibilism", *Philosophical Review*, Vol. LXXXVI, No. 1, 20-35.
- Le Pore, E. (1985). The Semantics of Action, Event, and Singular Causal Sentences", *Actions and Events*, ed. E. Le Pore and B. McLaughlin (New York: Blackwell), 151-161.
- Lucas, J. R. (1961). Minds, Machines and Gödel", *Philosophy*, Vol. XXXVI, No. 137, 112-127.
- Lucas, J. R. (1970). *The Freedom of the Will*, (Oxford: Clarendon Press)
- Malebranche, N. (1997). *Dialogues on Metaphysics and on Religion*, ed. N. Jolley, translated by D. Scott, (Cambridge: Cambridge University Press)
- McKay, T. and Johnson, D. (1996). "A Reconsideration of an Argument Against Compatibilism", *Philosophical Topics*, Vol. XXIV, 113-122.
- Penrose, R. (1994) *Shadows of the Mind*, (Oxford: Clarendon Press)
- Pink, T. (1996). *The Psychology of Freedom*, (Cambridge: Cambridge University Press)
- Pink, T. (2004). *Free Will: A Very Short Introduction*, (Oxford: Oxford University Press)
- Pradhan, R. C. (1999). Human Beings as Autonomous Persons", *Persons, Mind and Value*, ed. S. K. Mohanty, (New Delhi: Decent Books), 29-41.
- Smith, P. (2007). *An Introduction to Gödel's Theorems*, (Cambridge: Cambridge University Press)
- Stump, E. (1996). "Libertarian Freedom and the Principle of Alternative Possibilities", *Faith, Freedom and Rationality*, eds. M. Beaty and J. Jordan, (Lanham: Rowman and Littlefield), 73-88.
- Timpe, K. (2008). *Free Will: Sourcehood and Its Alternatives*, (London and New York: Continuum)
- Van Inwagen, P. (1983). *An Essay on Free Will*, (Oxford: Clarendon Press)
- Van Inwagen, P. (2002). Free Will Remains a Mystery, *The Oxford Handbook of Free Will*, ed. R. Kane, (Oxford: Oxford University Press), 158-177.

End Notes

- ¹ There is, of course, controversy over these arguments, for compatibilists in the first place. Literature is vast; see for example Huoranszki (2011).
- ² For controversy over Rule β see for example McKay and Johnson (1996), Finch and Warfield (1998) and van Inwagen (2002)
- ³ By action I mean also mental action, following Pink (1996, 2004). Mental processes as thinking, deliberating, deciding I take to be a species of action.
- ⁴ Of course, here I maximally simplified this analysis.
- ⁵ Literature which discusses Frankfurt-style examples is also vast. See for example Kane (2002, part V).
- ⁶ For another, more elaborate, defence of the position that sourcehood and alternative possibilities go together, see Timpe (2008: chapter 7).