#  Department of Philosophy

 The University of Notre Dame

 Notre Dame, Indiana 46556-4619

 574-631-5910 dept/ 574-277-7427 home

 Fax: 574-631-8209

 Email: peter.vaninwagen.1@nd.edu

**A Promising Argument**

## Peter van Inwagen

**0.** Let us say that it is at a certain moment *up to* one *whether* one will do A or doB if one is then faced with a choice between doing A and doing B and one is then able to do A and is then able to do B.[[1]](#endnote--1) And let us say that it is at a certain moment *undetermined whether* one will do A or do B if there is a possible world in which the laws of nature are the same as those of the actual world and whose state at that moment is identical with the state of the actual world at that moment and in which one will do A *and* a world satisfying those same two conditions in which one will do B.[[2]](#endnote-0)

 *Libertarianism* is the conjunction of the following two theses:

*The Free-will Thesis*: Various human agents at various times and on various occasions have been in the following situation: they were faced with a choice between alternative courses of action (e.g., between lying and telling the truth, between becoming a physician and becoming a concert pianist, between meeting Jill in Phoenix on Thursday and not meeting Jill in Phoenix on Thursday) and it was then up to them which of these courses of the action they would pursue.

*Incompatibilism*: If one is at a certain moment faced with a choice between doing A and doing B, it is then up to one whether one will do A or do B *only if* it is then undetermined whether one will do A or do B—and *necessarily* so.

**1.** Many philosophers have contended that libertarianism is necessarily false—that it is a metaphysically impossible or even a logically incoherent thesis. Of the various arguments that have been given for the impossibility of libertarianism, the simplest is that the impossibility of libertarianism is an immediate consequence of the following principle, which (it is contended) is a necessary—perhaps a conceptual—truth:

*The* *Indetermination-inability Principle*: If one is at a certain moment faced with a choice between doing A and doing B, and if it is then undetermined whether one will do A or do B, it is not then up to one whether one will do A or do B; in fact, one is not then able to do A *and* not then able to do B.[[3]](#endnote-1)

(The impossibility of libertarianism follows from the Indetermination-inability Principle because that principle and incompatibilism jointly entail that human agents are never in the following situation: they are faced with a choice between alternative courses of action and it up to them which course they will pursue; and the Free-will Thesis entails that human agents are sometimes in that situation.)

 But is the Indetermination-inability Principle true? The principle can seem very plausible if one considers only cases of this sort:

Jack is deliberating about whether to hit the right-hand or the left-hand side of the dartboard. It is now undetermined whether he will hit the right-hand side or the left-hand side owing the fact that it will be undetermined at the moment the dart leaves his hand which side of the board it will hit.

It is at least very plausible to suppose that Jack is not, during the course of his deliberations, able to hit the right-hand side and is not able to hit the left-hand side. But such cases are not decisive, since they involve the concept of *success* or at least the concept of *result*: they are cases in which in which an agent is now faced with a choice between doing A and doing B, and in which, if the agent should endeavor to do A or should endeavor to do B, whether the agent would succeed in either endeavor is now undetermined. But what about cases in which an agent is now faced with a choice between doing A and doing B, and in which it is now undetermined whether the agent will do A or do B—but in which it is now *determined* that, once the agent has decided one way or the other, the agent will succeed in doing the thing decided on?[[4]](#endnote-2) Should we say that such an agent is not now able to do A and not now able to do B? That is the real question, for the Indetermination-inability Principle implies that if one is now faced with a choice between doing A and doing B, and if it is now undetermined whether one will do A or do B, then (no matter what *else* may be determined or undetermined) one is not now able to do A and not now able to do B.

**2.** In “Free Will Remains a Mystery,”[[5]](#endnote-3) I presented an argument for the conclusion that if one is now faced with a choice between speaking and keeping silent, and if it is now undetermined whether one will speak or keep silent, then one is now unable to keep silent. (An exactly parallel argument could be used to defend the conclusion that someone in those circumstances is now unable to speak.) I quote the argument in full:

You are a candidate for public office, and I, your best friend, know some discreditable fact about your past that, if made public, would—and should—cost you the election. I am pulled two ways, one way by the claims of citizenship and the other by the claims of friendship. You know about my situation and beg me not to “tell.” I know (perhaps God has told me this) that there exist exactly two possible continuations of the present—the actual present, which includes your begging me not to tell and the emotional effect your appeal has had on me—, in one of which I tell all to the press and in the other of which I keep silent; and I know that the objective, “ground-floor” probability of my “telling” is 0.43 and that the objective, “ground-floor” probability of my keeping silent is 0.57. Am I in a position to promise you that I will keep silent?—knowing, as I do, that if there were a million perfect duplicates of me, each placed in a perfect duplicate of my present situation, forty-three percent of them would tell all and fifty-seven percent of them would hold their tongues? [Here I should have added the qualification ‘to a very high probability’.] I do not see how, in good conscience, I could make this promise. I do not see how I could be in a position to make it. But if I believe that I am able to keep silent, I should, it would seem, regard myself as being in a position to make this promise. What more do I need to regard myself as being in a position to promise to do X than a belief that I am *able* to do X? Therefore, in this situation, I should not regard myself as being able to keep silent. (And I cannot see on what grounds third-person observers of my situation could dispute this first-person judgment.)[[6]](#endnote-4)

I will call this argument the Promising Argument.

 Shortly before “Free Will Remains a Mystery” appeared in print, and too late for me to do anything about it, Michael Bratman convinced me (in conversation) that the Promising Argument was invalid. Here is my own statement of his diagnosis of the flaw in the argument:

You set the case up this way. Your friend wants you to promise not to reveal his misconduct, and, while you are deliberating about whether to make such a promise, you believe that that it is, at that moment, undetermined whether you will tell all or keep silent. Now you might well have that belief in those circumstances and it might even be true—but, if it is true, it doesn’t follow that, if you were to make the desired promise, it would *then* be undetermined whether you would tell all or keep silent. Perhaps if you were to promise to keep silent, it *would* then be determined that you would keep silent (it may for example be psychologically impossible for you to do something that you have promised not to do). If that would be the result of your promising to keep silent, then you *are* in a position—you are *now* in a position—to promise your friend to keep silent. Your argument was essentially this:

If I am able to keep silent, then I am in a position to promise to keep silent

If it is undetermined whether I shall keep silent, then I am not in a position to promise to keep silent

*Hence*, if it is undetermined whether I shall keep silent, I am not able to keep silent.

As I have shown, the second premise of this argument is false, or at any rate, might well be false.

The flaw is real. But can the basic idea behind the argument be saved? Can the argument be revised so as to eliminate the flaw? The following section contains my best attempt at such a revision.

**3.** Let us consider a more elaborate version of the indetermination-inability principle (the *New Indetermination-inability Principle*):

Where *t*2 is a future moment and *t*1 is a future moment earlier than *t*2:

If one is now faced with a choice between doing A at *t*2 and doing B at *t*2, and if, at *t*1, it will be undetermined whether one will do A at *t*2 or do B at *t*2 (and if this *would* then be undetermined whatever one might do between now and *t*1), then one is not now able to do A at *t*2 and one is not now able to do B at *t*2.

The New Indetermination-inability Principle may be defended by an elaboration of the Promising Argument—the New Promising Argument[[7]](#endnote-5):

My good friend Jake Higgins is a candidate for re-election to the Senate. He has been having a sleazy affair with Mary O’Brien, a married member of his staff, and I know all about it. Jake has learned that at a press conference tomorrow, a reporter well known for her muckraking skills plans to ask me the following question: “Are the rumors true—is Senator Higgins intimately involved with Mrs. O’Brien?” Jake knows that there is no hope of his being re-elected if I do not explicitly say that the rumors are unfounded (an evasive response to the question will be taken as confirmation of the rumors). He begs me to promise to deny everything; that is, to promise to lie to the press. How shall I respond to this request if I have the following belief?

At the moment the question is asked, it will be undetermined whether I shall respond with a lie or with the truth (and, therefore, it is *now* undetermined which I shall do): if there were a large number of perfect duplicates of me (in identical environments) at the moment the reporter asked her question, some (to a near certainty) would lie and some would tell the truth. And, moreover, it would at the moment the question was asked be undetermined how I should respond to it even if, at some moment between the present moment and that moment, I promised to lie.

(Perhaps I believe this because I believe that whichever decision I make—to lie or to tell the truth—will be a free decision and believe on philosophical grounds that free decisions must be undetermined events.) If I have this belief, it seems, I should now regard myself as *not in a position* to promise Jake to lie to the press tomorrow. (The case is not all that different from this one: Suppose I believe that when I try to start my car, it will at that moment be undetermined whether it will start—and that this will then be undetermined no matter what I do between the present moment and the moment at which I try to start the car; then I should regard myself as not being in a position to promise someone a ride.) And that means that I believe that I am not now able to lie to the press tomorrow—for if I were now able to lie to the press tomorrow, I *should* be in a position to promise Jake to lie to the press tomorrow.[[8]](#endnote-6) And the same argument (mutatis mutandis) will continue to apply as long as there is a future moment at which it will be undetermined whether I shall lie (in response to the reporter’s question) or answer honestly.[[9]](#endnote-7) If, for example, it is now 11:58, and if the reporter is going to ask her question at noon, and if it will then be undetermined whether I shall respond to her question with a lie or the truth—and if this would then be undetermined even if at some point in the next two minutes I promised to respond with a lie—, then I am not now, at 11:58, able to respond with a lie.[[10]](#endnote-8)

I think it’s fair say that if the New Promising Argument proves that in the case I have imagined I am not now able to lie tomorrow, then an easy generalization of the reasoning it embodies establishes the New Indetermination-inability Principle.

 Well and good. But does the New Indetermination-inability Principle—like the (original) Indetermination-inability Principle—imply the falsity of libertarianism?

**4**. One can deduce the falsity of libertarianism from the New Indetermination-inability Principle only if one can deduce from the New Indetermination-inability Principle that no human agent *x* is ever in the following situation:

—*x* is at some given moment faced with a choice between doing A and doing B

—it is at that moment up to *x* whether *x* will do A or do B

—it is at that moment undetermined whether *x* will do A or do B.

And one cannot deduce from the New Indetermination-inability Principle that no agent is ever in that situation. There are two reasons why the New Indetermination-inability Principle does not have that consequence. These reasons can be most easily displayed in a concrete example.

 Suppose that it is now 11:58 and that I am faced with a choice between lying at noon and telling the truth at noon. Suppose it is now undetermined whether I shall lie at noon or tell the truth at noon. Can we deduce from these suppositions and the New Indetermination-inability Principle that it is not now up to me whether I shall lie at noon? (That is, can we deduce that I am not both able to lie at noon and able to tell the truth at noon?) We cannot—and, as I said, for two reasons.

 First, it may be that 11:58, the present moment, is the *last* moment at which it is or will be undetermined whether I shall lie at noon or tell the truth at noon; it may be that at every moment between 11:58 and noon it will either be determined that I shall lie at noon or else determined that I shall tell the truth at noon—although now, at 11:58, it is—as it was at all earlier moments—undetermined *which* of these alternatives is the one that will be determined after 11:58.[[11]](#endnote-9) Note that the New Indetermination-inability Principle does not tell us that if it is now the *last* moment at which it is undetermined whether I shall lie or tell the truth at noon, then I am now unable to lie (or unable to tell the truth) at noon; it allows us to deduce that I am unable now, at 11:58, to do something at noon only given that it will be undetermined at some *later* moment, some moment between 11:58 and noon, whether I shall do that thing at noon.

 Secondly, suppose that there *is* a moment between the present moment and noon—11:59, say—at which it will be undetermined whether I shall lie or tell the truth at noon. We cannot deduce from that supposition and the New Indetermination-inability Principle the conclusion that I am at present unable to lie (tell the truth) at noon; to reach that conclusion, we should need a further premise: that it *would* be undetermined at 11:59 whether I should lie at noon or tell the truth at noon *no matter what I might do* between the present moment and 11:59. And why should we suppose *that*? How do we know, for example, that the following statement is false: if I *were* to promise thirty seconds from now to lie at noon, then it *would* be determined at 11:59 that I should lie at noon? (That counterfactual is certainly consistent with the proposition that that is now—at 11:58—undetermined whether I shall lie at noon: suppose it is now undetermined whether I shall make that promise.)

 I conclude that the New Promising Argument, for all its promise, is in one important respect a failure. True, it succeeds in establishing the New Indetermination-inability Principle, but it fails of its larger purpose, for it does not imply the falsity of libertarianism. The revisions of the Promising Argument that were needed to evade Bratman’s insightful criticism of that argument have weakened the conclusion of the revised argument to the point at which its conclusion does not imply the falsity of libertarianism.

 I would point out that this “failure” of the New Promising Argument by no means establishes the possibility of libertarianism. After all, an argument can be the most abject failure imaginable and nevertheless have a true conclusion. For all the failure of the New Promising Argument shows, the (original) Indetermination-Inability Principle may be true, and that principle entails the falsity of libertarianism. The burden of the present section is that reasoning that turns on the relation between one’s being able to do a certain thing and one’s being in a position to promise to do that thing cannot establish the Indetermination-inability Principle; such reasoning can establish only the weaker New Indetermination-inability Principle.

**5.** If the New Indetermination-inability Principle does not imply the falsity of libertarianism, it does imply the falsity of a thesis that most libertarians would be strongly inclined to accept, even if that thesis is not a strict logical consequence of libertarianism. I will call this thesis the *Deliberation-freedom Thesis*[[12]](#endnote-10):

At various times and on various occasions, there has been a human agent *x* who was in the following situation:

—*x* was faced with a choice between doing A and doing B

—*x* chose to do A after an extended period (several minutes, at least) of serious deliberation

—At every moment during that period of deliberation, it was up to *x* whether *x* would do A or do B (at every moment during that period of deliberation, *x* could say truly, “It is now up to me whether I shall do A or do B: I am now able to do A and I am now able to do B”[[13]](#endnote-11))

—During that period of deliberation, *x* was “condemned to freedom”: from the moment at which *x* first began to try to decide whether to do A or do B till the moment at which *x*’s choice to do A or *x*’s choice to do B was (irrevocably) made, there was nothing *x* could do to change the fact that, till the choice was made, it would up to *x* whether *x* would do A or do B (nothing, that is, short of breaking off deliberating about whether to do A or do B).

The New Indetermination-inability Principle and incompatibilism jointly entail the falsity of the Deliberation-freedom Thesis[[14]](#endnote-12). Libertarianism, as I have defined libertarianism, does not entail the Deliberation-freedom thesis, but it does seem to me to be a thesis that most libertarians would be very strongly inclined to accept. Some libertarians (and perhaps some opponents of libertarianism) might even want to say that the definition of libertarianism in Section 0 of this paper was unsatisfactory, unsatisfactory because “real” libertarians mean something more by ‘libertarianism’ than the rather weak thesis I gave that name to in Section 0.

 Perhaps there are philosophers who would say this and perhaps they are right. And perhaps they are right because “real” libertarians accept the Deliberation-freedom Thesis. Let us take that possibility into account by renaming the thesis that was called ‘libertarianism’ in Section 0: let us now call that thesis *weak* or *minimal* libertarianism. And let us call the conjunction of incompatibilism and the Deliberation-freedom Thesis *strong* libertarianism. (Strong libertarianism obviously entails weak or minimal libertarianism.) The New Indetermination-inability Principle argument implies the falsity of strong libertarianism but does not imply the falsity of weak or minimal libertarianism.

 The New Promising Argument refutes strong libertarianism (its conclusion implies the falsity of strong libertarianism). But how important an accomplishment that is is debatable.

 The following two questions are obviously relevant to the debate: *Are* there in fact any strong libertarians?; How committed are such strong libertarians as there may be to strong libertarianism? I am fairly sure that there are some strong libertarians, but am not at all sure how committed to strong libertarianism they may be. For all I know, most strong libertarians (always assuming that there are any) would be willing to say something along the following lines: “I *did* accept ‘strong libertarianism’ but the New Promising Argument has convinced me that that thesis is false. Well, no matter. Strong libertarianism is not really essential to my views on free will. I’m content to make a strategic withdrawal to weak libertarianism and defend that position.” If the friends of strong libertarianism are fair-weather friends of that sort, then the New Promising Argument is not as important as it would be if strong libertarianism had committed advocates.

 Another sort of consideration that is relevant to the question of the importance of the New Promising Argument has to do with the implications of the argument for the construction of libertarian models or theories of free will. By this I mean models of, or theories that attempt to provide an account of, what *goes on* within an agent when that agent performs an act of free will[[15]](#endnote-13)—models or theories of free will that in some way incorporate incompatibilism. (Many libertarians—I am not one of them—have attempted to construct such models or theories.) Let us say that a theory of free will is a *strong* libertarian theory if it entails both incompatibilism and the Deliberation-freedom Thesis. And let us say that a theory of free will is a *weak* libertarian theory if it is not a strong libertarian theory and it entails both incompatibilism and the free-will thesis. (Strong libertarianism is not a strong libertarian theory because it is not a theory at all; it is only a thesis or position. By the same token, weak libertarianism is not a weak libertarian theory.) The New Promising Argument may simplify the search for a libertarian theory of free will in virtue of the fact that it implies that no strong libertarian theory can be correct.

1. The dummy phrases ‘do A’ and ‘do B’, when they occur in the same sentence-schema are to be replaced by “incompatible” action-phrases—‘tell a lie’ and ‘tell tell the truth’, for example, or ‘be in Phoenix at noon tomorrow’ and ‘be in Chicago at noon tomorrow’. [↑](#endnote-ref--1)
2. None of the three propositions

Sally is now faced with a choice between doing A and doing B

It is now up to Sally whether she will do A or do B

It is now undetermined whether Sally will do A or do B

entails that Sally will either do A or do B. It might be, for example, that at a certain point in her life Sally was faced with a choice between becoming a physician and becoming a concert pianist, that it was then up to her whether she would become a physician or a pianist, that it was then undetermined whether she would become a physician or a pianist—*and* that she eventually decided to pursue neither of those vocations and became a journalist. [↑](#endnote-ref-0)
3. It is tempting to suppose that if one is *not able* to perform a certain contemplated action, it follows logically that one *will not* perform that action. This tempting thesis obviously entails there could be no agent who was going to do either A or B and who was neither able to do A nor able to do B. Anyone, therefore, who endorses, e.g., the conditional, ‘If it is now undetermined whether Sally (who in a moment will either lie or tell the truth) will lie or will tell the truth, then Sally is not now able to lie and is not now able to tell the truth’, must purport to be using ‘not able’ in a sense that does not have the tempting thesis as a consequence: that person must admit that there are possible (and no doubt actual) cases in which someone does something that he or she had not been able to do. But it does not seem implausible to suppose that ‘not able’ has a sense that permits this. It does not seem implausible to suppose that there is a sense of ‘not able’ in which someone who had never before touched a dart and who casually tossed a dart at a wall was, before the dart was tossed, *not able* to hit the particular square millimeter of wall in which the dart fortuitously came to rest. And it does seem plausible to suppose that if one is not able (in this sense) to do A, and not able (in this sense) to do B, then it is not up to one whether one will do A or do B. (In fact, it seems plausible to suppose that if one not able to do A, then it is not up to one whether one will do A or do B.) [↑](#endnote-ref-1)
4. That is to say, there are possible futures in which the agent decides to do A and does A and possible futures in which the agent decides to do B and does B; and there are no possible futures in which the agent decides to do A and does not do A, and no possible futures in which the agent decides to do B and does not do B. (A “possible future” is a possible world in which the past, present, and laws of nature are the same as those of the actual world.) [↑](#endnote-ref-2)
5. “Free Will Remains a Mystery” was originally printed in James E. Tomberlin, ed., *Philosophical Perspectives 14: Action and Freedom* (Oxford: Blackwell, 2000), pp. 1-19. It was reprinted in the first edition of this Handbook: Robert Kane, ed., *The Oxford Handbook of Free Will* (Oxford: Oxford University Press, 2002), pp. 158-177. [↑](#endnote-ref-3)
6. *The Oxford Handbook of Free Will* (1st edn.), p. 175. [↑](#endnote-ref-4)
7. I have taken the opportunity afforded by this revision to introduce a few “improvements” into the example on which the argument rests, improvements that are not strictly needed to meet Bratman’s criticism. [↑](#endnote-ref-5)
8. It may sound a little odd to say, “I am not *now* able to lie *tomorrow*.” For that matter, it may sound odd to say, “I *am* not able to lie *tomorrow*.” The oddness of both sentences is due entirely to the fact that lying is not—at least in any normal case—an action that requires some longish sequence of preliminary actions. We can certainly say, “I am not able to meet you in Phoenix tomorrow” (meeting you in Phoenix tomorrow might well require a longish sequence of preliminary actions), and this is essentially the same assertion as “I am not now able to meet you in Phoenix tomorrow” (that is: it’s now too late for me to do all the things needed to accomplish this end)—although in the latter case the hearer might not be sure what pragmatic function the semantically redundant adverb ‘now’ was supposed to be performing. In any case, there is nothing puzzling about the function of ‘now’ in the assertion, “Yesterday, on Monday, I was still able to meet you in Phoenix tomorrow (Wednesday), but I am not now, on Tuesday, able to do that.” Or if someone said to me, “Lie for me tomorrow,” the reply “If I were now to promise to lie for you tomorrow, I’d be making a promise I was not able to keep” would not sound at all odd. [↑](#endnote-ref-6)
9. More exactly: Let *t*2 be the moment at which the question will be asked, and *t*1 be an earlier future moment; a similar argument will show that I am not now able to lie at *t*2, provided that (a) at *t*1 it will be undetermined whether I shall lie at *t*2, and (b) if between the present moment and *t*1 I were to promise to lie at *t*2, it would (still) be undetermined at *t*1 whether I should lie at *t*2. [↑](#endnote-ref-7)
10. I contend that the New Promising Argument can be generalized so as to apply to *all* pairs of moments that satisfy the conditions specified in the previous note. I concede that this contention faces an obvious “practical” difficulty: it takes a certain amount of time to make a promise. [↑](#endnote-ref-8)
11. It is easy to see that such “indetermination of determination” cases are possible. Imagine a deterministic mechanism that displays a red bulb and a green bulb. Imagine that there is inserted into the works of this mechanism a smaller, indeterministic mechanism that will, in its indeterministic way, at 11:58, turn a switch in the deterministic mechanism either to the right or to the left. And imagine that if the switch is turned to the right, the red bulb must light two minutes later and that if it is turned to the left, the green bulb must light two minutes later. Then at every moment between 11:58 and noon, it will either be determined that the red bulb will light at noon or else determined that the green bulb will light at noon—although it will be at 11:58 (as at all earlier moments) undetermined *which* of these alternatives will be the one that is determined after 11:58. [↑](#endnote-ref-9)
12. Strictly speaking, this “thesis” is not a thesis but a thesis-schema. Instances of the schema can be obtained by replacing the dummy phrases ‘do A’, ‘do B’, ‘doing A’, and ‘doing B’ with appropriate English phrases. To say, “I accept the Deliberation-freedom Thesis” is to commit oneself to the schema’s having some true instances. [↑](#endnote-ref-10)
13. It is a firm conviction of mine that if one is at a certain moment deliberating about whether to do A or to do B, it *follows* that one believes at that moment that one is then able to do A *and* then able to do B. But, of course, if one necessarily believes that *p* whenever one is F, it doesn’t follow that it’s *true* that *p* whenever one is F and believes that *p*. It follows from my “firm conviction” that philosophers who believe that the free-will thesis is false have contradictory beliefs whenever they engage in deliberation. It does not follow that the free-will thesis is true. [↑](#endnote-ref-11)
14. That is, no instance of the schema is consistent with conjunction of the New Indetermination-inability Principle and incompatibilism. [↑](#endnote-ref-12)
15. In the language of the present paper, an act of free will may be defined in this way: if an agent was at some point faced with a choice between doing A and doing B, and if the agent did A, and if at some moment before the agent did A, it was up to that agent whether he or she would do A or do B, then the agent’s doing A was an act of free will. [↑](#endnote-ref-13)