

Self-Evidence Derailed: Descartes's Cogito and its Anticipations

Aleš Bunta

Centuries after Descartes formulated his first principle of philosophy—*cogito, ergo sum*—the Cartesian cogito remains, for better or worse, recognized almost universally as one of the most important turning points in the history of philosophy. However, what is almost as famous as the cogito itself is the fact that strikingly similar (if not identical) arguments existed long before Descartes introduced his dictum.

Of the so-called “anticipations” of the cogito argument, the most relevant appeared in the respective works of St. Augustine and St. Thomas Aquinas, the two philosopher-saints. The fact that the cogito had been anticipated by these two men whose legacies were still immensely influential during Descartes's period suffices to give rise to several significant questions: Was Descartes aware of the similarities? How could he not have been, given his formal Jesuit education at *La Flèche*? How is it that other scholars of the period, who must have been aware of these similarities, did not speak more openly against the alleged novelty of Descartes's insight? All in all, how did Descartes succeed in asserting that his first principle was new when so many must have thought that it was not?

It may come as a surprise that the first serious allegations of plagiarism surfaced as late as 1689, more than half a century after Descartes introduced his “first principle,” when Bishop Huet published his *Censura Philosophiae Cartesianae* (cf. Blanchet 1920, p. 19).

Forty-seven years earlier, Arnauld opened his notorious comment on the *Meditations* with the following remark: “The first thing that I find remarkable is that our distinguished author has laid down as the basis for his entire philosophy exactly the same principle as that laid down by St. Augustine—a man of the sharpest intellect and a remarkable thinker, not only on theological topics but also on philosophical ones” (Descartes 1985b, p. 139). It is certainly worth noting that Arnauld, who was evidently convinced that the cogito argument was not Descartes’s original invention, found a way of expressing his opinion without explicitly stating it. And yet, the hint left little room for misunderstanding.

Descartes’s short response to this comment is perplexing, to say the least: “I shall not waste time here by thanking my distinguished critic for bringing in the authority of St. Augustine to support me” (ibid., p. 154). Is this blatant cynicism? An indication that Descartes had no answer to Arnauld’s remark? Or is it possible that Descartes felt confident that the problem of the Augustinian anticipation presented no serious problem after all? And could that be the reason why he decided to respond with a mere pleasantry? And yet, the comparison Arnauld proposed in the course of his argument still appears very convincing.

Even though Aquinas’s anticipation was only a reformulation of Augustine’s cogito-like argument, the two arguments, when compared, turn out to not be the same. At a certain level, we can claim that the difference between the two anticipations is even greater, or at least clearer, than the difference between the respective anticipations and Descartes’s cogito. Thanks to Jaakko Hintikka’s famous interpretation, we now know that even though both anticipatory arguments prove or demonstrate the same thing, namely the irrefutability of one’s own existence, and that they do so in similar terms, their internal logics are in fact surprisingly diverse. While St. Augustine’s argument proceeds as a logical deduction of the irrefutability of one’s own existence, St. Thomas’s argument appears much closer to Hintikka’s own performative

account of the cogito, according to which the certainty of *sum* is not logically deduced from the fact of thinking. Rather, it emerges as a consequence of the attempt to consider one's nonexistence that instantly proves to be self-defeating, serving as an immediate manifestation of the certitude of the exactly opposite claim (Hintikka 1962, pp. 15, 16).

The obvious question is: Which of these internal logics is at work in Descartes's thought? To no great surprise, the answer is that Descartes appears to have borrowed from both argumentative principles. His reflections on the cogito from his *Replies* and *Conversation with Burman* are in fact not completely homogeneous and allow for two different lines of interpretation, each of them associable with either of the two saints. And this may well be one of the answers to the question of how Descartes managed to unintentionally defy all odds in asserting the novelty of his argument. Descartes appears to have repeatedly made use of one saint to keep him in the shadow of another saint's heavenly shine, so to speak, for whenever one tries to pin his argument to one of the two argumentative principles, a sort of remainder of the other principle appears indispensable. We nonetheless have to keep in mind that, apart from his peculiar reply to Arnauld, in his published work, at least, Descartes never so much as mentioned the problem of anticipations. The only source that we have on Descartes's own view of the issue is his private letter to the Dutch theologian Colvius, in which he claims that up until that very day, November 14, 1640, i.e., almost four years after he first introduced his first principle, he was unaware of any similarities.

Even to this day, Descartes's apparent dilemma regarding his central insight remains one of the most disputed problems within Cartesian studies. Thus, perhaps we should try to define it more precisely from this specific aspect. The reason why Hintikka can legitimately claim that the cogito argument should not be regarded as a logical inference relies heavily (although not exclusively) on the fact that on a few occasions Descartes himself appears to have

said so. In his *Replies to the Second Set of Objections*, for instance, he says the following:

When someone says, “I am thinking, therefore I am, or I exist”, he does not deduce existence from thought by means of a syllogism, but recognizes it as something self-evident by a simple intuition of the mind. (Descartes 1985b, p. 100)

Harry Frankfurt fiercely contested Hintikka’s conclusion that this sentence should be regarded as Descartes’s general rejection of the inferential explanation of the cogito (Frankfurt 1999, p. 9). Instead, Frankfurt argued that the sentence merely shows that Descartes rejected the possibility of explaining cogito in terms of a simple syllogism based on the major premise “everything that thinks is or exists.” And such a rejection, according to Frankfurt, does not preclude the possibility that Descartes thought of his argument in terms of another type of logical deduction. However, even though Frankfurt was possibly right in saying so, this still does not explain what exactly Descartes had in mind in stating that the cogito *emerges as self-evident*.

The answer to this question can perhaps be found in our other philosopher-saint. While Descartes himself was usually not a keen reader of other philosophers, he did study St. Thomas Aquinas rather extensively. Aquinas was in fact very much concerned with the problem of self-evidence, and for a very specific reason. He opposed the idea that God’s existence was self-evident. Therefore, he compiled a lengthy list of arguments that supported the thesis of God’s self-evidence, which included some of the most sophisticated argumentative principles of the Middle Ages, such as the notorious ontological argument elaborated by St. Anselm. So, the arguments in favour of God’s self-evidence were themselves anything but self-evident. This “black list” also included a specific type of argument, which was, in both its structure and content, related to the earlier cogito-like argument of Augustine.

While explaining one such argument, Aquinas writes in *De Veritate* (Q10, a12, 7): “God has existence more truly than the human soul has. But the soul cannot think that it does not exist. Therefore, much less can it think that God does not exist” (Aquinas 2008, p. 66). The argument appears solid enough. However, Aquinas immediately adds the following antithesis: “For something is immediately evident in two ways: in itself and to us. That God exists, therefore, is immediately evident in itself, but not to us. Therefore, to know this it is necessary in our case to have demonstrations proceeding from effects” (ibid., p. 68).

Aquinas’s solution is brilliant: it is probably even true that God’s existence is self-evident. As a matter of fact, as he explains, one can even deduce the self-evidence of God’s existence by means of a legitimate argument. For Aquinas does not say that the argument itself was false: if our own existence is *stricto sensu* self-evident, for we cannot even think of ourselves as nonexistent, then the existence of God, which is truer than ours, must by analogy be even more self-evident than our own. However, even though God’s existence may be *considered* as veritably self-evident, it is *not self-evident to us*. The self-evidence of God, even though it cannot be truly opposed, is floating somewhere out there without reaching its proper recipient within us. The very fact that it *can* be deduced, or even merely *thought of*, probably already speaks against its presupposed ability to directly appear in our thoughts in an uninvited way.

Let us now turn to the “anticipation” itself. This appears later in *De Veritate* (Q10, a12, ad7), mostly as an additional explanation of the initial premise with which Aquinas began; its primary purpose is to specify precisely why the “human soul cannot think that it does not exist.” The anticipatory argument is, in fact, an exemplary argument meant to explain the structure of this particular self-evidence. In other words, Aquinas’s anticipation of the cogito is an argument that serves as a *demonstration* of self-evidence in general and does not focus on establishing the indubitability of

one's own existence in particular. So, in contrast to Descartes, who appears to rely on some sort of intuitive self-evidence to prove the indubitability of one's own existence, Aquinas actually feeds one's own existence to self-evidence in order to show how this *essentially invasive* mechanism really works. The argument itself runs as follows: "No one can assent to the thought that he does not exist. For in thinking something, he perceives that he exists" (ibid., p. 70). Since the attempt itself of thinking of oneself as non-existent immediately reveals itself as contradictory, this shows that any attempt at denying one's own existence serves as its immediate proof. And obviously, in contrast to the alleged self-evidence of God's existence, this argument serves as an example of *true self-evidence*. For the certainty of our existence does not really appear as the result of being *thought of* (such as a solution to a mathematical problem emerges as a result of one's thinking of this problem); it rather emerges from within our thought—in principle—regardless of what we are thinking. But it proves itself most convincingly if we try to deny it. And this uninvited *appearing* within our thought seems to represent precisely what this particular type of self-evidence is all about.

The deep connection between Hintikka's performative interpretation and Aquinas's argument becomes immediately clear. The performative argument is in fact a derivation of Aquinas's formula of self-evidence. We cannot know whether Descartes actually read this anticipatory argument in the *De Veritate*. However, it seems highly likely that Descartes must have been at least aware of some other arguments that Aquinas treated as the arguments of God's alleged self-evidence, such as the argument that Aquinas introduces right at the beginning of the second chapter of his *Summa Theologica* (Q2, a1):

Again, it is self-evident that truth exists, for even denying so would amount to admitting it. If there were no such thing as truth, it would be true that there is no truth. So, something is true and, therefore,

there is truth. But God is truth itself: "I am the way, the truth, and the life." So, it is self-evident that God exists. (Aquinas 2006, p. 21)

Even though Aquinas does not show in this argument how one's own existence can be perceived as self-evident, he nonetheless uses the same *performative structure* to make his argument: If I say that the truth does not exist, I clearly think that it is true that there is no truth, which ultimately means that by thinking that the truth does not exist, I actually directly confirm its existence *by stepping in its place myself*. It is evident that even if Descartes knew only this passage, this could still have been a major influence on how he devised his cogito argument in *Discourse on Method*.

Hintikka's performative theory, which implicitly emphasizes Descartes's relation to Aquinas, has several advantages: first, the fact is that Descartes's "original" argument in *Discourse* can be explained as a modified, "analytical" version of the performatory argument such as we observed in Aquinas. Second, Descartes himself said that the cogito is not to be perceived in terms of deducing "existence from thought by means of a syllogism." Third, in contrast to his knowledge of Augustine, which was surprisingly shallow, Descartes was well educated in Aquinas. And fourth, Descartes says that cogito must be perceived as a specific type of self-evidence, and we can now see that Aquinas not only resorts to such a structure of self-evidence, but he also relates it internally to a cogito-like argument. The fact that his own approach is diametrical, as he introduces the cogito-like argument as a demonstration of how the intrusive mechanism of self-evidence overwhelms us, and not as an argument of existence, makes the whole comparison even more compelling.

However, Hintikka's interpretation also contains at least one fundamental flaw: his performative theory of the cogito argument cannot be convincingly applied to Descartes's argumentation leading up to the point of certainty in the *Meditations*, which is Descartes's most philosophically elaborated and sophisticated

work. In my opinion, Frankfurt is right in saying that Descartes's argumentation in the *Meditations* must be perceived as a series of deductive arguments after all.

Frankfurt's argument elaborates what I will subsequently call the *argumentative substructure* of the cogito in the *Meditations*:

Instead of showing that *sum* can be deduced from a premise that is certain in its own right, Descartes in effect points out that a premise from which *sum* can be elicited is an essential and inescapable element of every context in which the need for assurance concerning *sum* arises. (Frankfurt 2008, p. 152)

However, this fundamentally correct and concise interpretation of Descartes's argumentation in the *Meditations* raises one serious difficulty. If we accept Frankfurt's interpretative thesis as the ultimate answer to Descartes's argumentation in the *Meditations*, then we must also effectively recognize *as a fact* that Descartes's treatment of the problem of the indubitability of one's own existence was no different from St. Augustine's approach to the same problem. This becomes immediately apparent once we put Augustine's various arguments together in a slightly more systematic way.

*

Augustine's dictum "*Si fallor, sum*" consists of the following argument: "If I am mistaken, I am. For if one does not exist, he can by no means be mistaken. Therefore, I am, if I am mistaken" (*City of God*, XI. 26; cf. Matthews 1992, p. 29). Now, let us compare this passage from *City of God*, a passage that Augustine himself described as one of those "truths" in the face of which the "quibbles of the sceptics lose their force," to Descartes's introduction to his first principle in the *Discourse on Method*.

I noticed that while I was trying to think everything false, it was necessary that I, who was thinking this, was something. And observing

that this truth “I think, therefore I am” was so firm and sure that all the most extravagant suppositions of the sceptics were incapable of shaking it, I decided that I should accept it without scruple as the first principle of philosophy I was seeking. (Descartes 1985a, p. 127)

Putting these two arguments side by side and noting their common references to skepticism, it instantly becomes clear that the true difficulty we are facing is not one of finding several obvious similarities that make one wonder how Descartes could have been ignorant of Augustine's passage while writing his own. Instead, the true task is to find any essential difference between the two arguments at all, save, of course, the first impression that Descartes's passage is more elaborate. Aquinas's passage, on the other hand, is a different story. As noted above, Aquinas introduced his own version of the argument in a context that could hardly have been more diverse in all possible respects.

Similar passages are spread throughout Augustine's works. To take another example, this time from the dialogue *On Free Choice of the Will*, which is especially important because it was used by Arnauld to turn Descartes's attention to the problem of anticipations of the cogito argument. This time we are faced with a dialogue between the saint himself and Bishop Evodius.

So, to start off with what is clearest, I ask first whether you yourself exist. Are you perhaps afraid that you might be deceived in this line of questioning? Surely if you did not exist, you could not be deceived at all. (*On Free Choice of the Will* II. 3; cf. Matthews 1992, p. 11)

Evodius's reply is rather short: “Go on.”¹

The first thing one may notice (also due to Evodius's short answer), is that Augustine's arguments are in effect so simple,

¹ While quoting this passage, Arnauld was pointing to a parallel with the following sentence from *Meditations*: “But there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me. In that case, I too undoubtedly exist, if he is deceiving me” (Descartes 1985b, p. 17).

and so unpretentious in their own way, that they appear to be self-evident. And this is important to observe because it shows that the entire history of the cogito argument was marked by an affinity for self-evidence.

From a slightly different angle, in Augustine's dialogue with Evodius this tendency towards self-evidence can be observed in Evodius's relative silence. What we see in this dialogue is that the subject who ultimately meets the truth remains silent throughout the argument. Quite literally, it is as if the process of his own reasoning proceeded without him. This can be said of Evodius (for it is the irrefutable truth of his own existence that is at stake). It is as if his own thinking addressed Evodius from an external position, a position represented by Augustine, who in turn asks to confirm the certainty of Evodius's existence. Evodius is clearly the key player in this game, for it is he and he alone who attains the truth of the irrefutability of his own existence. However, the entire process of thinking involved in the argument seems to serve no other purpose but to reveal itself to Evodius as something self-evident and almost redundant.

However, no matter how self-evident the result of Augustine's argument may appear to be, or how scarce the words of the subject intended to receive its meaning are, the argument that lies beneath this effect is neither simple nor does it lack strategic thinking. So, according to Augustine, why is one's own existence irrefutable? And why is this simple truth, which he himself treats as self-evident and of which he insists one only needs to be reminded, nonetheless extraordinarily important?

The argument's surface is, in fact, perfectly simple: if I presume that I exist—which is what I normally do—I cannot be making a mistake (by assuming so), for I cannot make a mistake without existing. Since my own existence is the necessary *pre-condition* of any possible mistake I can make, I can conclude that my existence itself, my assumption that I exist, cannot be counted among my mistakes. In other words, I cannot be wrong

in presuming my existence because *if I were* wrong in presuming my existence, I would not be here making presumptions of any kind whatsoever. Nor, for that matter, would I be asking any questions regarding whether they were right or wrong. Yet, I am clearly presuming something. And it is precisely this, the most superfluous of the possible pieces of evidence, that lies at the heart of any cogito argument.

Therefore, even though the skeptics may be perfectly right in saying that there is no way of telling whether our senses are not constantly deceiving us, and even though it is possible that we are thus kept in a state of constant delusion without knowing it—I can nonetheless remain reassured that I can attain truth *from within this falsehood itself* that possibly surrounds me. While I presume that I exist, I also spontaneously reflect (my own existence as) *the precondition of falsity's own appearing*. So, even if the delusion is all-encompassing, and I am consequently wrong in all that I think, I must still necessarily exist as the precondition of my own being deceived. Therefore, my spontaneous presumption that I exist cannot be wrong, regardless of the circumstances.

Not only is the argument immune to falsity, as it is anchored in that minimum of being (represented by “my” existence) that falsity itself requires to appear, but even more importantly, the *type of truth* that Augustine discovered is only unilaterally opposed to falsity. My existence, since it cannot be perceived as false, remains the opposite of falsity; however, falsity itself does not really oppose this truth (embodied by my existence), rather it serves as its ultimate proof. Any possible mistake that I make only confirms the irrefutable “truth” of my existence, it does not contradict it. Rather, *truth itself exists* as this internal contradiction of falsity, which *cannot befall my existence without contradicting even its own possibility*. Therefore, what Augustine really discovered is a specific *dimension of truth*. And it is important to stress this because he himself appears to have regarded his insight in precisely this way. This also explains why he treated the argument itself

as something almost laughable: for it is the specific nature of the truth he discovered that matters. In other words, by revealing the irrefutability of one's own existence, Augustine discovered a specific type or dimension of truth, which cannot be regarded as the pure opposite of falsity. This truth must be perceived as the internal contradiction of falsity as such, which nonetheless makes of this truth an even more effective weapon, for it ultimately turns falsity into its own affirmation.

A dialogue like the one between Evodius and Augustine represents a perfect form of expressing Augustine's insight: the Augustinian argument should fundamentally be regarded as a simple *mental experiment designed to fail*. Try to deny your own existence and you will be instantly proven wrong. And obviously, because Augustine's argument is the first historical example of an insight that is essentially embedded within what we regard today as "the subjective perspective," the repetition of such micro-experiments is also the only way to transmit the insight from one subject to the next. Even though truth itself is almost self-evident, the argument is only rendered true by means of its repetition. That is to say, by enacting the experiment's permanent failure.

Regardless of its apparent simplicity, Augustine's "*Si fallor, sum*" is in fact a formidable and complex argument which—on the basis of thinking alone—turns falsity itself into an efficient instrument of truth. More precisely, Augustine's argument can be regarded as a kind of strategic pincer movement that captures falsity from both sides: on one hand, the argument reaches beneath falsity by revealing that the presumption of one's own existence must be considered to be a reflection of falsity's own precondition, while on the other hand, this encirclement of falsity is completed by the fact that the evidence of the certainty is in fact *so shallow and superfluous that it cannot even be considered wrong*—for it is simply one's awareness that one is presuming something, regardless of whether this presumption itself is right or wrong. Ultimately, Augustine's cogito truly demonstrates the

self-evidence of which one merely needs to be reminded, but this is precisely where the argument's greatness lies.

No matter how remarkable Augustine's argument might be, what interests us here is another question: Is the argument that is developed by St. Augustine really the same argument that Descartes later introduced as his "first principle of philosophy"? Regardless of the similarities, which are clear and compelling, the two passages we just compared are in fact dissimilar in one specific way, which can make all the difference. Or at least Descartes thought so. So, let us examine the two arguments once again.

First Augustine: "If I am mistaken, I am. For if one does not exist, he can by no means be mistaken. Therefore, I am, if I am mistaken."

And then Descartes: "I noticed that while I was trying to think everything false, it was necessary that I, who was thinking this, was something."

The first tangible distinction (at the level of the passages themselves) lies in the fact that Descartes not only proves the subject's existence, but he also posits or allocates this existence exclusively within the order of thinking. In contrast to Augustine, who "only" proves existence by showing that I cannot be wrong in presuming it, Descartes attributes subject's existence to doubt (the attempt at thinking everything false) directly. "I noticed that while I was trying to think everything false, it was necessary that I, *who was thinking this, was something*." So, in other words, Descartes not only proves this existence, but he also simultaneously begins to answer the question: "What is it that exists with such irrefutable certainty?" And his answer is: doubt.

Descartes himself was firmly convinced that this distinction resolves the entire problem of Augustine's anticipation. In *Letter to Colvius*, surprisingly enough, he instantly admits that he and Augustine rely on the same argument to prove our existence. But the real issue within his own argument, he says, is not just the mere proof of existence, but precisely demonstrating "that this I which

is thinking is an immaterial substance with no bodily elements” (AT III, p. 247; cf. Matthews 1992, pp. 12, 13). It is a well-known fact that this argument of Descartes spectacularly fails. Descartes did not consider the fact—or was even truly unaware of it—that arguably Augustine’s main philosophical goal was, no less than his own, precisely to show that body and mind were two diverse substances.² And one of the main functions of his cogito-like argument was to support this ontological view. So, Descartes’s attempt to move away from Augustine ends up stumbling into the holy man coming around the corner. However, what interests us here, at least for now, are the structures of the two arguments. And, in this sense, Descartes’s opinion remains an important guideline. For it shows that his own argument had a double, almost paradoxical purpose: He not only uses it to prove one’s own existence, but at the same time he also uses the argument to derail this existence from the order of being. He thus shows that existence only exists as thinking (and nothing else). So, how does Descartes reach this double effect at the level of the argument itself?

Even if my thinking that nothing exists were true and nothing is out there, the mere fact that this thought itself is being considered proves that the thought, contrary to other things, exists. The argument thus already indicates that the thought must be regarded as ontologically independent of the rest of the world. For the argument shows that even if nothing else exists, the thought itself *exists necessarily*. Even if it tries as hard as it can, it cannot escape the grip of existence, while the fate of the rest of the world remains obscure. Clearly, moving forward or outward in this direction leads us back to Augustine’s own dualist ontology. So, we must refocus on the argument itself, and ask ourselves precisely what this certainty hinges on. And the answer is, on thought’s own contradiction. What stands at the

² For a more elaborate study on this problem, see, for example: Matthews 1992, pp. 14-18, or Menn 1998, pp. 210-17.

centre of Descartes's cogito is precisely this consideration of pure nothing, which is instantly proven to be self-refuting *by the very fact that it is being considered*. For its immediate result is that by the effort to prove itself wrong this thought demonstrates that it necessarily exists even if nothing else does. Contrary to its own effort, it is cast out of the nothing it tries to imagine while the existence of all other things goes down the drain. Descartes pins the subject's existence precisely to this self-refuting thought and not to its subsequent reflection. Clearly, this thought to which my existence is pinned *is a mistake*. It cannot be regarded otherwise because it is proven wrong by the mere fact that it is being considered. However, it is precisely the contradiction that emerges as a *result* of the argument that serves as the point of certainty on which the whole of Descartes's argument hinges. So, ironically enough, Descartes's argument is the veritable *si fallor, sum*, for while Augustine merely anticipates the conclusion that I cannot be making a mistake when I think that I exist, Descartes actually pins this existence directly to a thought that is proven true (necessarily existent) by proving itself wrong.

However, does not Augustine say the same? He clearly says that even if everything were false, I would then exist with even greater certainty. For falsity affirms and does not contradict my existence. The arguments clearly meet in the claim: "Even so, if everything is false, I exist." Moreover, existence in both cases hinges on a mistake. But Augustine's dictum neither directly indicates that existence is inherent to thinking, nor does it show that I could exist even if nothing else existed. As a matter of fact, his strong point is precisely that my existence is *necessarily brought along* by something else existing within the order of being and this is falsity, or more precisely, *what appears false* (to me). Obviously, this does not mean that "my" existence is not absolutely certain. Augustine's certainty is still rooted within the order of being. There must be something else out there for something to appear false to me. Since my existence is the precondition of this

false appearing, this leads to the conclusion that my existence, as *the being* of falsity's *appearing*, thus reveals itself beyond falsity through the appearing of falsity itself. On the other hand, Descartes's argument in *Discourse* would almost "prefer" there to really be nothing behind the curtain of appearance. For this scene of Descartes floating in the void like a lost astronaut in space is precisely what enables him to say that only thought itself exists unquestionably.

So, it seems that we can conclude that the internal logics that drive the two arguments are in fact surprisingly diverse, and that Hintikka's general assessment that Augustine's cogito can be regarded as a logical deduction, while Descartes's argument was best perceived as a performatory argument, was correct. In fact, if only *Cogito* and *Si fallor* are put in question, such an assessment would indeed be sustainable. However, Augustine also developed several other types of cogito-like arguments that were not based on the assumption of one's existence, but upon *doubting this same assumption*.

"Who would doubt that he lives?" asks Augustine, "for even if he doubts, he lives; if he doubts, he remembers why he doubts; if he doubts, he understands that he doubts; if he doubts, he wishes to be certain; if he doubts, he thinks; if he doubts, he knows that he does not know; if he doubts, he judges he ought not to consent rashly" (*Trinity*, X.10.14; cf. Matthews 1992, p. 14). Quite a few certainties arise within Augustine's insight that are deducible from doubt directly, the first of which is precisely this implicit or negative awareness of existence harboured by doubt itself.

At this point, we must observe the main problem of Hintikka's rejection of Augustine's anticipation. On the one hand, Hintikka was arguably right in saying that Augustine's arguments can be regarded as just local applications of the general ontological principle "whatever thinks must exist," which can be traced back to Parmenides. However, this can only be ascertained regarding the mistake argument, where the point of certainty is indeed

logically deduced. And obviously, even the doubt argument *can* be perceived in this way: I cannot doubt my existence because in order to doubt I need to exist. However, Augustine's insistence on the negative awareness harbored by doubt itself, which seems to indicate that doubt itself represents a specific form of knowledge, strongly suggests another solution. Namely, that Augustine's arguments of doubt need to be perceived as performatory arguments, after all.

The situation gets even more complicated. For in the *Meditations*—and that means *after he had already been warned* of the similarities by Mersenne—Descartes approached even closer to Augustine's line of argumentation. The key to his deduction of the point of certainty is closer to Augustine's dialogue with Evodius than to his own, arguably performative, argument in the *Discourse*, at least at the level of its substructure.

**

We have already said that Descartes responded to the problem of Augustine's anticipation just once, in his private letter to Colvius. However, Stephen Menn recently noted another interesting source (Menn 1998), namely one of Descartes's earliest fragments, which Baillet later published under the title "The Appendix" to *Studium Bonae Mentis*. The passage is relevant, for it can be regarded as Descartes's own miniature theory of originality.

In this fragment, Descartes attempts to answer the following (quite familiar) question: How can I say *anything at all* without merely repeating already known philosophical ideas? This partially explains why it is not so relevant that Descartes does not use the expression "originality" in the passage. Descartes tries to determine a very specific type of originality; let us name it the *intrinsic minimum of originality* (or even zero-degree originality). Let us first take an example:

The impression is that the sentence "how can I say *anything at all*, without merely repeating the already existing ideas?" might

at first seem to be another version of the question: “How can I assert *something new*?” However, for Descartes, there is an essential difference between these two questions.

The full version of Descartes’s question could be rephrased thusly: What is it that separates thinking from merely repeating the ideas of other people given that *in philosophy, thinking can only proceed by repeating* the philosophical ideas that have been previously conceived by others?

What constitutes the originality of thinking cannot be considered as *a property of any particular philosophical sentence or opinion* (for these are, in fact, in constant fluctuation). Consequently, this also means that the second sentence, “How can I assert *something new* (a new opinion)?”, cannot pass this intrinsic criterion, while there still may be some hope for the first sentence: “How can I say *anything at all*?”

As we can write no words in which there are letters other than those of the alphabet, nor complete a sentence unless it consists of the words that are in the dictionary, so neither [can we compose] a book except out of the sentences [or opinions, *sententiae*] that are found in others. But if the things I say are so coherent among themselves and so connected that they follow from each other, then it will not follow from this argument that I have borrowed my opinions from others any more than I have taken the words themselves from the dictionary. (AT X, p. 204; cf. Menn 1998, pp. 12, 13)

The passage clearly addresses the concept of originality even though the word itself is never used. Originality is addressed in its *negative form* insofar as Descartes posits a hypothesis that represents the antithesis of originality, without explicitly referring to the notion of originality. In principle, there is only one way of efficiently eradicating the possibility of originality and that is to assert that thinking as such can only consist of repeating already existing thoughts. If thinking is essentially nothing other than repeating what has already been said, then there is no such thing

as “original thought.” Descartes introduces this hypothesis by saying that a book cannot be composed in any other way “except out of the sentences (or philosophical opinions) that are found in others.” So, the first sentence of the passage is a negative hypothesis: writing a new book relies on opinions that are found in others, no less than writing a single sentence is limited by the number of words that actually exist in the dictionary, or as writing a single word is determined by the choice of letters that can be found in the alphabet. Thus, the possibility of any originality is reduced to zero.

However, another observation is equally important. The mere repetition of already existing thoughts is not only *the existential negation of originality*; it also embodies the inner *qualitative contradiction within thinking* itself. So, it is very important to note that Descartes does two things at once: he introduces the definition of the nonexistence of originality (that is to say, he determines its negative limit designated by repetition), but at the same time he introduces this definition in such a precise way that the definition implies a contradiction within the very possibility of thought.

If there is such a thing as the “auto-negation” of thinking, a “cognitive process” that reflects its own nonexistence while still belonging to the domain of thinking, then the answer is clear: such inner contradiction of thinking is represented by mere repetition of the thoughts of other people. When we merely repeat what other people have already said, we clearly do *not* think (in any qualitative sense of the word), but the process of repetition in which we thus partake is nonetheless still to be perceived as thinking. Let us put it this way: repeating is hollow thinking. It radiates nonbeing from within thinking itself. As a matter of fact, the hypothesis that thinking is simply a version of repeating seems to negate not only the possibility of (my) original thought, but also the possibility of my existence at all. For if my thinking consists entirely of repeating, then I am nothing but a copy of *something*

else repeating itself in the shape of “my” alleged thought. “I” thus still think, but my existential status is then reduced to that of a mere representation of some other thing’s *past* existence. So, I am myself quite literally nothing. *Cogito, ergo non sum*. Obviously, this is a far-fetched speculation, but let us just remember, after all, that Descartes was the philosopher renowned for questioning his own existence. Furthermore, since so many commentators report his “excessive passion for originality” (Blanchet 1920, p. 57), it is at least conceivable to imagine that, for Descartes, original thought was almost a matter of existence itself.

By treating repetition as the intrinsic negation of originality, Descartes achieved a double effect: the negation of originality incorporates the qualitative negation of thinking as such. Why is this so important? Because in doing so, Descartes exposed that there is *a certain minimum of originality that thinking must respond to in order to qualify as such*. This is the primary meaning of what I have called the minimum of originality (or zero-degree originality). The minimum of originality is therefore to be imagined as some sort of invisible threshold that thoughts must surpass to legitimize themselves as thinking. And the minimum itself is obviously defined by thought’s intrinsic capacity to exceed the level of mere repetition (of the already existing thoughts)—*at least to the extent of reaching zero-degree originality*.

So, the hidden dichotomy of the passage runs like this: on the one hand, thinking as such is defined by an intrinsic capacity to exceed the level of mere repetition; but, on the other hand, philosophy, which is allegedly the highest form of thinking, due to the necessity of repetition does not meet this minimum criterion of thinking as such. How did Descartes resolve this dichotomy? He resolved it so quickly that the dichotomy itself is not at all easily observable.

Descartes both begins and ends with the comparison to letters in the alphabet and to words in the dictionary. What Descartes first needs is namely to get rid of the *logic of origins*. For what

effectively shackles us to the past is not some sort of complete impossibility of originality, but instead the appearance that there actually may be *some original ideas*, which continue to be repeated. So, what Descartes does first with this comparison is to admit that no entirely new philosophical opinion was possible in a strong sense. Precisely because the comparison makes such a strong statement, it must convince even the most conservative observer (let us not forget that, upon writing this, Descartes had just finished his studies). However, once the comparison is accepted as the renunciation of any higher aspirations to originality, this comes with a price (for the Other), so to speak. As soon as one accepts the comparison to the letters in the alphabet and words in the dictionary, the status of the original itself is also called into question. For it is in fact no different than the status of simple words or letters.

Another way of putting this is to say that Descartes severed the relation between origin and originality, thus appropriating the past to the benefit of the present. Thought's originality has nothing to do with whether or not the particular thought had been used before. Since the number of thoughts, according to Descartes's argument, appears to be finite, we would have to conclude that all philosophical opinions have always already appeared at some previous stage. What truly legitimates thought's intrinsic minimal originality depends instead on the question of whether or not a philosophical thought can be considered as the *spontaneous result* of an author's current thinking. A thought is legitimately used "if the things I say are so coherent among themselves and so connected that they follow from each other."

On one hand, this claim exposes Descartes's view on the appropriation of other philosophers' thought: indeed, it is clear that he regarded the entire history of philosophy as a junkyard full of recyclable materials that should not go to waste. On the other hand, it is no less clear that his standards of originality were *set much higher* than our own. We can use other people's

philosophical opinions, says Descartes; however, we can only do so if we manage to mould them into an *entirely new philosophy*.

This is precisely what we need to keep in mind when reading Descartes's argument in relation to the Augustinian anticipation as he presented it in his *Letter to Colvius* approximately twenty years later:

I am obliged to you for drawing my attention to the passage of St. Augustine relevant to my [I think therefore I am]. I went today to the library of this town to read it, and I find that he does really use it to prove the certainty of our existence. He goes on to show that there is a certain likeness of the Trinity in us, in that we exist, we know that we exist, and we love the existence and the knowledge we have. I, on the other hand, use the argument to show that this I which is thinking is an immaterial substance with no bodily elements. These are two very different things. In itself it is such a simple and natural thing to infer that one exists from the fact that one is doubting that it could have occurred to any writer. But I am very glad to find myself in agreement with St. Augustine, if only to hush the little minds who tried to find fault with the principle. (AT III, p. 247; cf. Matthews 1992, pp. 12, 13)

In his letter to Descartes of July 3, 1648, Arnauld wrote the following: "I read your treatise on the separateness of mind and body with great passion. I found it clear, reasonable and superb; it does not matter that, for the sake of truth, this topic has been almost equally superbly treated by St. Augustine almost throughout his book on the Trinity, especially so in chapter 10" (AT V, p. 186). This shows that at the level of facts Descartes's argument seems to be a genuine blunder. Augustine did in fact use several cogito-like arguments in relation to his own dualist ontology. More specifically, he used these arguments to prove that the mind could attain certainty of itself regardless of the senses, showing that the body and the mind were two diverse substances. In this sense, Descartes's general argument, which emphasizes his own dualism, fails to distinguish itself from Augustine's argument.

What makes the problem of the anticipations so difficult is precisely the fact that, once we reach the final destination, we lose. The only way to find distinctions at the level of the cogito itself is to stick as long as possible to the structures of the arguments themselves, while awaiting the inevitable return of the same.

However, beneath the general failure of Descartes's argument, two other significant emphases can be found. As noted above, Descartes does not really defend the novelty of his argument at the level of the argumentative substructure, for he immediately admits that Augustine used precisely the same argument "to prove the certainty of our existence." In fact, at this basic level, Descartes defends his own authorship of the cogito not by drawing a clear distinction between himself and Augustine, but rather *by the complete abolition of authorship as such* in relation to the cogito argument in general. And this obviously goes hand-in-hand with his view that the cogito is to be considered conditionally self-evident. Descartes's point is that the ideas that illustrate a certain self-evidence cannot be claimed by any single author. In this regard, it makes perfect sense to say: "In itself it is such a simple and natural thing to infer that one exists from the fact that one is doubting that it could have occurred to any writer." Descartes's argument is in complete agreement with his own general views on the problem of philosophical originality discussed above.

Therefore, at the level of argumentative substructure, there is no difference. This does not mean, however, that the Cartesian and Augustinian arguments were essentially the same. According to Descartes, the true difference between these arguments lies within the intended purpose of their use, which retroactively changes the *essence* of the argument itself. This *formal aspect* of Descartes's argument is crucial. It is not his answer to the challenge itself that is important, because his answer is ultimately misleading. What is truly important is the fact that Descartes himself clearly distinguished between two levels of the argument. There is the argumentative substructure, which he admits is shared with

the argument offered by Augustine. This nonetheless does not present a problem for Descartes because he is convinced that the argumentative substructure of the cogito relies on self-evidence and due to this feature it is therefore authorless. But apart from this argumentative substructure, there is also some sort of meta-layer of the argument that retroactively changes the quality of the initial argument. Descartes's own answer hinges entirely on this meta-level, which is where his own answer ultimately fails. However, in what follows it will become apparent that this distinction between the two levels of argumentation is in fact the key to the solution.

With regard to Descartes's demonstration of certainty in the *Meditations*, I would like to briefly propose three connected hypothetical observations:

a) At the beginning of the *Second Meditation*, Descartes actually develops *two deductions* of the cogito argument and not just one. While both of them are fundamentally Augustinian, the first argument might also be perceived as a slightly modified version of the cogito in the *Discourse*. However, Descartes's argument in the *Meditations* does allow for the point of certainty to be deduced in the same way as it is logically deduced in Augustine's "*Si fallor*." The same goes for Descartes's second cogito argument in the *Meditations*, the argument of the deceiving demon, which is based on practically the same argumentative principle as Augustine's dialogue with Evodius, at least at the level of its argumentative substructure. Taken individually, neither of the two arguments display any specifically Cartesian feature that could serve to distinguish Descartes's argument from that of St. Augustine.

b) At the same time, however, one must observe that Descartes eventually anchors the point of certainty in the second of the two deductions. His final mastery of doubt depends on some sort of *public demonstration*, which is based on the famous proof of the deceiver. This demonstration itself, however, does not really draw

its entire effect from the Augustinian ground of the argument, but rather depends on the continuity of both deductions. More precisely, it depends on the difference that emerges between the two arguments *only after* we have posited them in continuity.

c) The two arguments clearly work in support of each other. However, they work for each other by working *against* each other. According to my hypothesis, the second of the two arguments retroactively shows that the first cogito argument needs to be reconsidered as the last of the *arguments of doubt*. In other words, the first of the two cogito arguments in the *Meditations* does not resolve the problem of doubting our own existence. Rather, it can be perceived as the argument that makes it possible for such doubt *to emerge as a rationally legitimate possibility* in the first place.

Descartes proceeds in the following order. First, he presents one version of the cogito argument:

But I have convinced myself that there is absolutely nothing in the world, no sky, no earth, no minds, no bodies. Does it now follow that I too do not exist? No: if I convinced myself of something then I certainly existed. (Descartes 1985b, pp. 16, 17)

Then he continues with the second argument:

But there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me. In that case I too undoubtedly exist, if he is deceiving me; and let him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something. (Ibid., p. 17)

The key to the whole story is probably this appearance of the deceiving demon. Does Descartes only want to say that the point of certainty that he reaches in the first argument is so solid that it cannot be shaken, not even by the omnipotent demon? Is its entire purpose purely epistemological? Or is there another, more ontological purpose behind this argument?

The first thing we must note is that the argumentative basis of the two arguments is the same. In both cases, the argument is fundamentally built upon a demonstration of the meditator's existence, demonstrating that this existence is itself a *precondition of the very attempt* at putting this existence in question. I cannot doubt my own existence, because the fact that I am doubting it proves that I exist. And even the demon cannot deceive me regarding the fact that I exist, because in order to be deceived at all I must exist as the precondition of the deception's success. This argumentative principle is obviously Augustinian.

As Miran Božovič (Božovič 1990) and Slavoj Žižek (Žižek 1998) have both shown in their respective studies, the key difference between the two arguments lies in the fact that, at the end of the second deduction, the subject's existence actually *emerges as an object*. In the first argument, the meditator simply observes that it is impossible that he has not existed while asking himself whether or not he exists. The second argument leads to a different result: the subject himself is deduced as the *objective precondition* of any act of deception. And precisely because it has been deduced as an object (of deception), the subject can be regarded as *something* that belongs to the "objective coordinate system." While in the first argument the meditator merely finds absolute *subjective* certainty in his own irrefutable existence, the second argument actually changes this subjective certainty into "something" that exists objectively (or "publicly," as Hintikka would have it). Although nothing has changed at the level of the argumentative substructure, the argument thus clearly attains a quality that is not represented in Augustine's argumentation.

At this point, I would like to briefly turn once again to the second anticipation, namely to Aquinas's formula of self-evidence: "No one can assent to the thought he does not exist. For in thinking something, he perceives that he exists."

I must admit that this formula always appealed to me more than any of the other cogito arguments. There is something

magnificent as well as horrific about a philosophical formula that changes a possible misunderstanding into the structure of its understanding. However, we must now interrogate the difference between Aquinas's formula and the Cartesian cogito. But perhaps we have posed the question in the wrong way. For once we pose this question in terms of difference, it seems we are set on the wrong track, at least if we perceive this "difference" as moving away from the formula of self-evidence. The real question is: How did Descartes *break into this formula*? Or, how did he eventually colonize the formula of self-evidence? For Descartes did break into this formula of self-evidence and in doing so he changed it into a kind of real-life philosophical experience. Perhaps all his philosophy relies on this initial act of burglary. However, in order to break into Aquinas's formula of self-evidence, Descartes had to actually *invent* doubt of one's own existence, and not to resolve it.

Doubt and rationality are prerequisites for one another in the *Meditations*. Before attaining the point of certainty, rational thinking is kept safe from delusion only in its relation to doubt. However, considered on its own, doubt is also not without a certain "delusional potential." This could easily have been one of the reasons that Descartes so famously excluded madness from the list of reasons for legitimate doubt. Doubt must be rational. As a matter of fact, every single episode of doubt in *Meditations* is rationally founded. For instance, at the same time that the meditator starts doubting even the simplest of mathematical truths, he contemplates the fact that all truth is contingent on God's will. God created those mathematical truths and God can subsequently change them. None of the episodes of doubt consists in doubting alone; each doubt must be complemented by the "positive" act of reflecting upon the reason why this particular doubt is possible. The only exception appears to be the doubt in the meditator's own existence. But is this really so?

But I have convinced myself that there is absolutely nothing in the world, no sky, no earth, no minds, no bodies. Does it now follow that I too do not exist? No: if I convinced myself of something then I certainly existed. But there is a deceiver of supreme power and cunning who is deliberately and constantly deceiving me. In that case I too undoubtedly exist, if he is deceiving me; and let him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something. (Descartes 1985b, pp. 16-17)

It is true that Descartes's dismissal of doubt in the meditator's existence appears prior to the introduction of the deceiver. What is it then that the hypothesis about the deceiver stands in for if not some sort of extension of this same doubt into "my own" existence that has just been raised? After all, if this were not so, why would Descartes say: "let him deceive me as much as he can, he will never bring it about that I am nothing so long as I think that I am something." Clearly, Descartes would not need to state this if there were no remaining doubt in the meditator's existence.

The deceiver cannot shatter my existence directly. He is not a person who would appear in order to convince us of something. And even if he did appear, Descartes would be obliged to dismiss him as a false authority. What the deceiver can do is "*claim*" the *cogito argument* for himself. And he can do so with authority, for even though his own existence is rather obscure, he allegedly could be perceived as the mediator between myself and my own thoughts. "Is there not some God, or some other being by whatever name we call it, who puts these reflections into my mind?" (Ibid., p. 16) All thoughts of mine can easily be those of God or of the deceiver.

What I propose here is that *once I deduce my existence from thinking, I actually link the existence itself to the same thoughts* that the demon could possibly control. In this precise sense, the doubt of my own existence becomes possible only after the cogito itself has already been articulated. For the demon would have no means whatsoever to threaten my existence if the existence has

not already been shown to be ontologically dependent on thinking. Only once this has been proven can I start asking myself: But what if even the seemingly unshakable idea that I exist is in fact just a delusion controlled by the demon? In short, my existence can only be suspect to doubt if it is shown to be dependent on thinking. And this is what the cogito argument does. This is why the cogito must be repeated once more from another angle, which shows that even the hypothesis about the demon actually confirms the truth that I exist while thinking—instead of threatening it.

Bibliography

- Aquinas, St. Thomas (2006) *Summa Theologiae, Questions on God*, trans. Brian Davis (Cambridge: Cambridge University Press).
- (2008) *Truth*, Vol. II, trans. Robert W. Mulligan (Eugene, OR: Wipf & Stock).
- St. Augustine, Aurelius (1958) *City of God*, trans. Gerald G. Walsh et al. (Garden City: Image Books).
- Blanchet, Léon (1920) *Les antécédents historiques du "Je pense, donc je suis"* (Paris: Felix Alcan).
- Božovič, Miran (1990) "Descartes – dvom, gotovost, norost," *Dvom in norost* (Ljubljana: Društvo za teoretsko psihoanalizo).
- Bunta, Aleš (2016) *A-cogito* (Ljubljana: Društvo za teoretsko psihoanalizo).
- Descartes, René (1964-76) *Œuvres de Descartes*, ed. Charles Adam and Paul Tannery (Paris: Vrin).
- (1985a) *The Philosophical Writings of Descartes*, Vol. I, trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press).
- (1985b) *The Philosophical Writings of Descartes*, Vol. II, trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press).
- Frankfurt, Harry (1999) *Necessity, Volition, and Love* (Cambridge: Cambridge University Press).

- (2008) *Demons, Dreamers, and Madmen* (Princeton and Oxford: Princeton University Press).
- Hintikka, Jaakko (1962) “*Cogito, Ergo Sum: Inference or Performance?*” *The Philosophical Review* 71, 1:3-32 (Durham: Duke University Press).
- (1996) “*Cogito, Ergo Quis Est?*” *Revue Internationale de Philosophie* 50, 195:5-21 (Paris: Vrin).
- Matthews, Gareth (1992) *Thought's Ego in Augustine and Descartes* (Ithaca: Cornell University Press).
- Menn, Stephen (1998) *Descartes and Augustine* (Cambridge: Cambridge University Press).
- Popkin, Richard H. (1964) *The History of Scepticism from Erasmus to Descartes* (New York: Humanities Press).
- Žižek, Slavoj (1998) “The Cartesian Subject versus the Cartesian Theater,” *Cogito and the Unconscious*, ed. Slavoj Žižek, 247-74 (Durham and London: Duke University Press).