Machines and Us: The Comparison of Machines and Humans at the Test of the Problematic of Solipsism

Timur Cengiz Uçan
Bordeaux Montaigne University, Pessac, France
Sciences, Philosophy, Humanities Mixed Research Unit
Social Sciences of Contemporary Changes Department

Abstract:

The first objective of this article is to propose a reflexion about the limits of the comparison or analogy or metaphor between humans and machines. This comparison which runs through the history of European philosophy (Aristotle, 1995, 1253b23; Descartes, 2006, pp. 157-159; Onfray de la Mettrie, 1996, 3-39; Kant, 2007, §65; Lewis, 1934, p. 144; Sartre, 2003, p. 248; Wittgenstein, 1947, Ts-229, 448), is basic for functionalism, and central for the development of medical sciences. For the distinction between parts of living bodies, in particular, between organs, involves the consideration of distinct and mutually compatible biological ends, whose coordinated functioning together renders satisfaction possible. However, although the affirmation of the comparability of these two types of cases is not problematic as such, the affirmation of the identity or indistinctness of these relations is not without posing problems, whether conceptual or practical. If humans are under some aspects *like* machines and inversely, as some tasks are realizable by humans or machines, another thing is to suppose affirming that humans *are* machines, or that machines *are* humans (see C. I. Lewis, 1934). The stake of this point is considerable, for its range is not only the literality of the personification involved by the humanization or biologization of machines as robots (for we are not surprised by saying that such robot sweeps, achieves actions, smiles), but also that the depersonification involved by the machinization or metaphorical dehumanization of humans (whether to express an appreciation of the realization of a task by a person or to express the horror

and the inhumanity, the absence of emotions involved by the realization of an action by a person). But its range also concerns: the extension of our concept of autonomy, the asymmetry of our relations to rules, principles, laws, of humans and machines, and in fact to a stronger extent our concept of *relation*. The question is thus whether this comparison, pertinent under some aspects in some contexts for certain ends, could have been adequate, turned out not be a comparison at all, such that the metaphorical could have become in such cases, literal. This affirmation could have seemed entirely incompatible with new possibilities of liberation rendered possible by technological innovations. In reality that is not the case since these possibilities are understood as such against the background of precedent possibilities. The problem we then shall pose is the following: to which extent does the comparison or metaphor or analogy of human machine render possible the necessarily nonrestrictive limits of intelligibility? What are the limits of this comparison? To which extent does the recourse to this comparison turn out beneficial? To contribute to the resolution of this problem, I shall propose to put the comparison between machines and us and of us with machines at the test of the problematic of solipsism. To achieve this task, I present the criticism made by Lewis of solipsism (1934), and then present Turing's critical reconception of solipsism (1950). I then attempt to establish the way in which Wittgenstein, with his criticism of solipsism (1953), functionalism, and reductionism, solves the problems encountered by the conceptions of solipsism of Turing and Lewis.

Keywords:

artificial intelligence, consciousness, C. I. Lewis, machines, solipsism, A. Turing, L. Wittgenstein.

Introduction[1]

This first objective of this article is to propose a reflexion about the limits of the comparison or analogy or metaphor between humans and machines. This comparison, which runs through the history of European philosophy (Aristotle, 1995, 1253b23; Descartes, 2006, pp. 157-159; Onfray de la Mettrie, 1996, 3-39; Kant, 2007, §65; Lewis, 1934, p. 144; Sartre, 2003, p. 248; Wittgenstein, 1947, Ts-229, 448), has been studied (cf. Kennedy, 2022), and is basic for functionalism and central to the development of medical sciences. The distinction between parts of living bodies, in particular, between organs, involves the consideration of distinct and mutually compatible biological ends, whose coordinated functioning together renders satisfaction possible.

However, although the affirmation of the comparability of these two types of cases is not problematic as such, as the comparability and eventually the similarity of relations between wholes and ends is involved by scientific and engineering practices (for example, the wing of the plane is *like* the wing of the bird and inversely), the affirmation of the identity or indistinctness of these relations is not without posing problems, whether conceptual or practical. If humans are under some aspects *like* machines and inversely, as some tasks are realizable by humans or machines, another thing is to suppose affirming that humans *are* machines, or that machines *are* humans. And it is uncertain that whoever achieved or even, strictly speaking, *tried* or could have tried such affirmation[2].

The stake of this point is considerable, for its range is not only the literality of the personification involved by the humanization or biologization of machines as robots (for we are not surprised anymore by saying that such robot achieves actions as sweeping, smiling, etc.), but also that the depersonification involved by the mechanization or metaphorical dehumanization of humans

[1] Many thanks to Donald Cornell, to the reviewers, and to the editors of this volume for their helpful remarks and criticisms about this text.

[2] This negation might seem incompatible with some uses of the metaphor between humans and machines, as that, for example, of Wittgenstein (1947, Ts-229, 448), but one central purpose of this article is to propose the epistemological elucidation that it is not. On this see also Bouveresse (2022, pp. 259-260).

(whether to express an appreciation of the realization of a task by a person or to express the horror and the inhumanity, the absence of emotions involved by the realization of an action by a person). Its range also concerns: the extension of our concept of autonomy, the radical asymmetry of our relations to rules, principles, and laws, of humans and machines, and in fact to a stronger extent our concept of *relation*. Another way to formulate the conceptual difficulty (as is ordinarily, frequently, commonly "verified" that we are "humans" in our ordinary internaut uses), is that of the indeterminacy of what we do when we *lend* to machines what we know of other humans, and of what we do when we lend to humans what we could *not*, strictly speaking, have ignored of machines, conceived to render possible either the better execution of some tasks, or the simple execution of some tasks (strictly speaking unrealizable by humans without their intermediacy).

The question is thus whether this comparison, pertinent under some aspects in some contexts for certain ends could have been adequate, turned out not to be a comparison at all, and the metaphorical could have become in such cases, literal. Surely, numerous technological innovations (biological computers, interfaces, and tools adjunctive to human bodies) render, for some conceptions, to some extent porous (cf. Kennedy, 2022) conceptual distinctions that could have seemed sealed, and mutually uncommunicative. Yet, if the open-endedness or intrinsic evolutivity of language is undeniable, it is uncertain that in the case of the comparison of humans and machines, we could have had to grant that this comparison could have ceased to be one, and became a unique literal means of expression. This affirmation could have seemed entirely incompatible with new possibilities of liberation rendered possible by technological innovations. In reality that is not the case since these possibilities are understood as such against the background of precedent possibilities: the intelligibility of history as social and objective science is tied to this point. The problem we shall then pose is the following: to what extent does the comparison or metaphor or analogy of humans and machines render possible the necessarily unrestrictive limits of intelligibility? What are the limits of this comparison? To which extent does the recourse to this comparison turn out beneficial?

The response to this question is also important to think about some structural similarities of debates about ecological or climatic catastrophisms in relation to the development of artificial intelligence: similarly to ways in which catastrophistic narrations about climate provide occasions to think of the reality of the ecological emergency, catastrophistic narrations about artificial intelligence provide occasions to think the reality of the possibility of conceptions and detrimental uses of artificial intelligence. This is not unrelated to the fact that environmental or technological misuses are too often causes of environmental or technological catastrophes. But equally important is to remark that such catastrophisms should not be held as the presentation of some paralyzing aspect of reality in any sense whatsoever. Not only the transformations of facts (by contrast notably with the analyses or the explanations of facts) but also the misleading presentations of false facts as true, exaggerations (as under-evaluations) neither substitute nor could have substituted for the conception of artificial intelligences or for ways in which artificial intelligences can contribute to the resolution of environmental problems.

To contribute to the realization of this task we shall propose to put the comparison between machines and us and of us with machines at the test of the problematic of solipsism. To achieve this task, I shall first present the criticism made by Lewis of solipsism in "Experience and Meaning" (1934). I shall then present Turing's critical reconception of solipsism in "Computing Machinery and Intelligence" (1950) and propose a critical assessment of this conception against the background of philosophical results achieved earlier by Lewis. I will then attempt to establish the way in which Wittgenstein's criticism of solipsism, functionalism, and reductionism in the *Philosophical Investigations* solves problems centrally encountered by the critical conceptions of solipsism of Turing and Lewis.[3]

[3] The notion of machine is not used in any theatrical sense throughout the text. The proposed approach is both critical of the very reductive criticisms of theatricality involved, for example, by Fried's notion of theatricality, and of the very inflationist conceptions of theatricality involved in some conceptions that allegedly would continue or have achieved the criticism of the Enlightenment. Theatricality is neither a problem nor a solution per se. But, as I shall attempt to render clearer in the third part of this text, in which I shall present a study of Wittgenstein's criticism of solipsism, self-estranged theatricality is

1. The critical conception of solipsism of C. I. Lewis

1.1. The problem of the solipsistic supposition according to which we '*are*' machines

> "To repudiate all such transcendence is to confine reality to the given, to land in solipsism, and in a solipsism which annihilates both past and future, and removes the distinction between real and unreal, by removing all distinction of veridical and illusory" (Lewis, 1929, p. 183)

> "Descartes conceived that the lower animals are a kind of automata; and the monstrous supposition that other humans are merely robots would have meaning if there should ever be a consistent solipsist to make it. The logical positivist does not deny that other humans have feelings; he circumvents the issue by a behaviouristic interpretation of "having feelings." He points out that your toothache is a verifiable object of my knowledge; it is a construction put upon certain empirical items which are data for me – your tooth and your behaviour. My own toothache is equally a construction." (Lewis, 1934, p.144)

C. I. Lewis expressed in 1929 a critical diagnosis of solipsism: Solipsism would be a position where one would arrive as the result of a repudiation – of a refusal – of "transcendence". Such a would-be position, solipsism, *thusly* reached (inasmuch as Lewis presupposes that we can distinguish solipsisms), would involve: the annihilation of past and future, and the removal of the distinction between the real and the unreal as the outcome of the removal of any distinction between the veridical and the illusory. Among conceivable and eventually conceived solipsisms, such solipsism would be peculiarly unbeneficial, and delusory. For the rejection of every distinction between the real and the illusory, incompatible with the reality of past and future, can seem to leave as our only

---

delusory.

option a self-contradictory assumption according to which only the present and whatever is presently and sensorially available could exist (metaphorically "given" to mind). But if whatever is sensorially available to us is all that is real, then whatever is not sensorially available to us is not real. So according to the conception of solipsism, devised and critically diagnosed by Lewis, we either would have to acknowledge "transcendence", that is, that the real could not possibly reduce to the sensorially available, or could not think a distinction between the veridical and the illusory, the real and the unreal, the past and the future. Conceptual distinctions of relevant opposites could but should not be entirely suppressed. There would be the possibility of delusory entrapment within a possibility that is not a possibility.

Lewis attempted to this extent to account for a distinction between a conception of solipsism which is inherently delusory, from a philosophical acknowledgement of (the reality of) reality: the idealism presented by "the world is my idea" could not ultimately but turn out to be acknowledgement of the fundamental and natural similarity between the idea of the world of an individual person, and the world whose idea is that of an individual person[4]. But Lewis did not render explicit in 1929 the motives of his critical conception and diagnosis of solipsism. He proceeds to such clarification in 1934, in the above quoted passage. He there argues that among solipsisms, a much more problematic solipsism would consist in the supposition that humans could not and would not be anything but robots. Thereby Lewis leaves aside the traditional characterization of solipsism, which would consist in the claim that a single person could be the only reality there is, and of which Schopenhauer had earlier argued that it would be claimed only in psychiatric institutions. According to the conception criticized by Lewis, any attempt to identify another human would necessarily fail and amount to an attempt to misidentify a robot – and not the opposite[5]. Any attempt to distinguish other humans from robots and robots from other humans

[4] An approach which is relevantly comparable with that of Wittgenstein in the *Tractatus* (2003, 5.62) which inspired Lewis.

[5] Considering the direction of the use of the comparison of machines and humans to explain the criticism of Lewis matters – as remarked by Bouveresse about Wittgenstein's approach (2022, p. 259).

would necessarily fail. Lewis does not unfold his diagnosis, but the difficulty is easily expressed: unlike humans, robots are tools conceived and produced to achieve the automatic achievement of tasks according to human desires, some of which are unachievable otherwise. The result of the negation of the conceivability of a distinction between robots and humans, the affirmation of the reducibility of humans to robots cannot but raise multiple problems concerning our relations. For, although some humans have engendered some other humans, no human has engendered every other human. No robot has engendered or could engender a human. Every robot has been produced by humans or by other robots, themselves produced by humans. Humans could not be reducible to tools, may have their own conceptions of which they are more or less conscious, and have their own desires and ends. Conceptions according to which humans could be produced for nothing but the satisfaction of the desires of other humans, and peculiarly, of their genitors, are abnormal: human procreation could not be reducible to slave production.

Two important aspects of the criticism made by Lewis of solipsism are to be considered. Lewis not only argues in favour of a distinction among solipsisms, but also among solipsists, according to the eventual consistency of claims and actions. The mere affirmation of the reducibility of the only reality there is to oneself by a person is not coherent, as earlier brought out, and such incoherence is prima facie manifest: such solipsism would involve simultaneous negation of the previously considered solipsistic affirmation and inversely. But Lewis (as Sartre in 1943 with *Being and Nothingness*, Part 3) also considers the eventuality of solipsistic maintaining of (solipsistic) inconsistency. Indeed, Lewis considers a difficulty with respect to the activity in which "supposing", and its results – "suppositions" – consist. A supposition results from an eventually expressed and eventually collective activity of thinking a truth. And in many ordinary, unproblematic and desirable cases, the truth of a fact is not and could not be dependent upon the decision of someone else. Thus, at first sight, Lewis can seem to be claiming that, as the achievement of a supposition by someone is directly dependent upon the action of only one person and no one else, the supposition that humans are merely robots can successfully be achieved by

whoever thusly supposes. All cases considered: either a person supposes that humans are robots, or a person does not suppose that humans are robots. If we grant that the negation of the conceivability of a distinction between humans and robots is monstrous, in the sense of problematically abnormal, then its achievement cannot be really successful. But then the true answer of the question "Can one relevantly and successfully achieve the negation of the distinction between robots and humans?" could seem to remain indeterminate, as could seem relevant to negate the relevance of the previously expressed conditional for practical purposes. This is the difficulty addressed by C. I. Lewis just after the quoted passage[6]. The question of the determinacy of the true answer to the question "Can the distinction between robots and humans be negated?" could be, according to logical positivism, circumvented by means of a behaviouristic interpretation:

> "The logical positivist does not deny that other humans have feelings; he circumvents the issue by a behaviouristic interpretation of 'having feelings.' He points out that your toothache is a verifiable object of my knowledge; it is a construction put upon certain empirical items which are data for me – your tooth and your behaviour. My own toothache is equally a construction." (Lewis, 1934, p. 144)

According to such a picture, human relations could be reducible to partially communicative behaviours of humans which would consist in the sensorially accessible part of otherwise inaccessible data of humans about each other. The difficulty brought out by the circumventing pointed out by Lewis is that the affirmation that the feelings of others can be accessed only indirectly – through behaviours – cannot but have consequences with respect to the evaluation of a human's own feelings by oneself[7]: if the feelings of others are mental objects, constructions which are forever only partially accessible to an individual person, then one's own feelings are also constructions which are forever only partially accessible to others, and eventually to that individual

---

[6] This difficulty is also addressed by Sartre who explicitly presents behaviourism as solipsism put into practice (2003, 253) and also, as we shall see, by Wittgenstein (2009, §420).

[7] See Uçan (2016) on this.

person oneself. Multiple difficulties arise from such an unreflexive "strategy": among which notably mutual alienation, devaluation of knowledge, and possibly destructions[8].

1.2. Is the moralistic rejection of the comparison of humans and machines philosophically receivable?

Humans-to-machines reductionism is, on Lewis' terms, "monstruous" in that strictly carried out, such conception involves for practical purposes the self-contradictory negation of the conceivability of any distinction whatsoever between (other) humans and machines. The neglect of this problem has consequences with respect, notably, to our understandings of our experiences (as shared common experiences would be unintelligible as such), to our respective knowledges of others (which also would be unintelligible as such). "Reduction", in this sense, ultimately leads to mutual alienation, devaluation of knowledges, and eventually to destructions. To this extent, Lewis raised the question of the identity of methodological solipsism with solipsism, a question to which Putnam, Sartre, Descombes, and Wittgenstein also provided positive answers (Sartre, *Being and nothingness,* 2003, p. 253; Putnam, "Why reason can't be naturalized", pp. 236-7, 1996; Descombes, *La denrée mentale*, 1995, p. 289; Wittgenstein *Philosophical Investigations*, §420): *Methodological solipsism, mere internalism*, *reductionism with respect to mind* (other minds) is not ultimately distinct from solipsism[9].

Remarkably, Turing, with the attempt to render clearer that machines think, has done in 1950 under one description exactly that which Lewis argued against. That is to say, one thing is to compare humans and machines for the achievements of some ends, goals, tasks, and finalities, and another

---

[8] These problems, which are related to *problematic* skepticism, contrarianism, and denialism are further considered and exemplified in the second part of the present article.

[9] This way of expressing their common criticisms could not conceivably reduce to a would-be "argument of authority", and rather involves acknowledgment of the fact that distinct philosophers from different philosophical traditions have at diverse times and places reached independently the same result in diverse ways with respect to the would-be distinction between "methodological solipsism" and "solipsism": ultimately there is no such distinction. As shall be rendered clear in the third part of the present text, the sort of possibility of verification that would allegedly be necessary for the establishment of the truth of solipsism cannot be possibly be verified and is not, could not be, a possibility of verification at all.

thing is to reject that distinctions between humans and machines can be achieved whichever are the considered ends, goals, tasks, finalities. Especially against the background of a tacit agreement to the traditional verticalist *scala naturae* conception of a hierarchy of lives according to degrees of complexity and soulfulness, the comparison of humans and machines may have seemed to unavoidably involve the breaking of a taboo. But does the approach for scientific purposes of relations of parts and wholes of organisms as mutually coordinated involve the negation of the receivability of the moralistic criticism of the comparison of humans and machines?

To reply exhaustively to this question, the precision of the sense of the question, and the consideration of distinct cases will prove beneficial. The comparison of humans and machines is basic to functionalist achievements whose results are undeniable – notably in medical sciences. Inasmuch as we can compare parts of wholes of human organisms with parts of wholes of mechanisms constructed for definite ends or aims, we can distinguish functions and ends or coordinated parts of wholes[10]. Such comparisons contribute to render conceivable the resolution of theoretical problems, required for practical resolutions of health problems, and the conception of preventive and curative practices, which can be institutionalized. Whether such achievements do involve "metaphysics" can be asked. For as we shall see, although Turing rightly called into question moralistic ways of criticizing the achievability of the analogy of humans and machines for scientific purposes, this criticism was achieved by Turing with a misleading and distorted picture of other cultures, and especially of Islamic cultures, while Lewis had earlier argued that the resolution of the problem raised by the solipsistic supposition – "metaphysical solipsism" – required very limited, and more integrative, dependence to "metaphysics":

> "A robot could have a toothache, in the sense of having a swollen jaw and exhibiting all the appropriate behavior; but there would be no pain connected with it. The question of metaphysical solipsism is the question whether there is any pain connected with your observed behavior indicating toothache." (Lewis, 1934, p. 145)

[10] On the compatibility of the criticism of the sufficiency of at least some "'mechanistic' world-view" see Putnam (1975a, pp. 364, 366, 385)

Lewis grants the conceivability of "metaphysical solipsism", which would consist in the question of whether there is and could be any pain connected with an observed behaviour (for example, a behaviour indicating toothache). "Metaphysical solipsism", as a solipsism, implies wrongly calling into question the existence of a connexion between an observed behaviour and pain. Such connexion could be unverifiable and unknowable. From the outset, the conception of "metaphysics" involved by the "metaphysical solipsism" envisaged Lewis is very minimal. Unmoralistically, such conception involves just the acknowledgment of the commonality of the veridicality of the expressions of their pains by humans[11]. Such a conception is compatible with any moralistic conception of the veridicality of the expressions of their pains by humans, that is, any conception according to which one must only veridically express that pain is felt by oneself *because* of some prescription, rule, law internal to a world-conception. Any such conception is indeed compatible with the existence of connexions between behaviours expressive of pains and experiences of pains (by contrast with the cases of machines and robots) and incompatible with fake expressions of pains by persons while no pain is felt by them.

The receivability of the moralistic criticism of the comparison of humans and machines is to this extent debatable: the mere rejection of the relevance of such comparison by appeal to a principle, religious or not, is not receivable since functional achievements (by contrast with functionalism) are not only conceivable but achieved and further will be achieved. The use of such comparison has a central place in the development of medicine, for the autonomous development of persons, individual or collective (institutional). But moralistic criticisms of the rejection of any conceivable distinction whatsoever between machines and humans because of a prescription, rule, or law internal to a world conception present some truth, as the negation of the distinction between robots or machines and humans does not result, could not have resulted, in the indistinctness or abolishment of the distinction between machines and humans. Such criticisms seldom are satisfactory, at least, if the appeal to a prescription, rule, or law, is meant to coerce the

[11] "Any metaphysics which portrays reality as something strangely unfamiliar or beyond the ordinary grasp, stamps itself as thaumaturgy, and is false upon the face of it." (Lewis, 1929, p. 10).

acknowledgment of the expression of pain as such, of the existence of a connexion between a behaviour expressive of the experience of pain, and the experience of pain.

2. The critical conception of solipsism of A. Turing

2.1. “Can machines think?”

As mentioned, Turing achieved, under a description, in “Computing Machinery and Intelligence”, exactly that which Lewis argued against. In this part, I will propose a philosophical and epistemological study of Turing’s conception and criticism of solipsism in that article. I will attempt to render clear that although Turing there established that machines can somehow be unproblematically said to think, that thoughts and actions can relevantly be ascribed to machines, the conception of solipsism there put forward is, to express the point in Lewis' terms, “thin” (1929, p. 30). The reduction of the problematic of solipsism to one and only one of its aspects, socially regrettably enough contributed to the replacement of a philosophical conception of solipsism by an unphilosophical one, whose consequences are yet to be brought out, studied, and criticized. Turing indeed introduces a conception of solipsism, to carry out a criticism of solipsism, in one of the counter-objections to the objections to the argument proposed with *Computing Machinery and Intelligence*, namely, the would-be objection that is called by Turing, “the argument from consciousness”. To critically assess this conception, let us first recall the problem posed by Turing and the replacement strategy proposed as an indirect means to achieve the resolution of the problem.[12]

[12] This problem is deeply related to the relations of our conceptions of common sense with the one of Turing, inasmuch as (quasi-)paradoxically, common sense is necessarily debatable, open both to philosophical and unphilosophical contestations and acknowledgements. In that, Turing’s approach faces difficulties similar to the one of Sartre (2003, pp. 481-489) as their conceptions of common sense are not, at least, *that* common. Yet uncommon claims of common sense can desirably become common. For a historical and philosophical account of the development of Turing’s conception of common sense in relation to Wittgenstein see Floyd (2021).

After having proposed a consideration of the question “Can machines think?”, Turing considers a difficulty concerning an answer to this question (Turing 1950, p. 433). Uncritical adherence to an understanding of the question employing definitions that somehow “reflect so far as possible the normal use of the words” would be scientifically and philosophically problematic. Sciences and knowledge do progress with linguistic uses – uses of words – which are neither necessarily incompatible nor necessarily compatible with, independent from uses that are normal or considered as normal within a community, a society, of linguistic practitioners. Were we to restrict ourselves only to available “normal use of the words”, novelty, improvement, discoveries, and creations, would almost be impossible, creativity could be reducible to exhaustion of combinations of allowed moves predetermined by social norms, and social norms would be unquestionable, whichever these are. But, if we would merely reject available “normal use of the words”, similarly, novelty, improvement, discoveries, and creations would almost be impossible, as novelties, improvements, discoveries, and creations could not be expressed within, and eventually understood, by a community, a society of linguistic practitioners. Thusly posed, everything can seem as if we are unavoidably entrapped in a predicament:

Either we accept that machines can think, reject “the normal use of the words”, the relevance of the examination of meanings involved by common uses of words. But then we might be led to assume that we must to rely on a statistical evaluation of the meanings of “meanings”. But then the justification of the answer could not be provided in any community anyway, and then both the meaning of the question and the end achieved by the asking of the question are lost.

Or we reject that machines can think, accept “the normal use of the words”, the examination of meanings involved by common uses of words as both relevant and sufficient. But then we cannot justify our answer except by reiterating appeals to “the normal use of the words”.

Turing thusly presents a dilemma which could not be resolved and which would result from opposite demands: that of the uncritical adherence to the common meanings of words for the sake of communication and critical rejection of the common meanings of words for the sake of novelty,

discovery, and progress. Whether the phrase "machines can think" is true or false is a question that cannot, as such, be directly and satisfactorily answered. As a means for an indirect resolution of the problem raised by the question "Can machines think?", Turing presents a replacement strategy with "the imitation game" (Turing, 1950, p. 433). In this "game" an interrogator has the objective to identify out of two persons with whom communication is achieved from a distance and without visual contact, a woman and a man, who is a woman and who is a man, provided that the man will attempt to make the identification fail. Such a game should be considered as a correct replacement to the initial question of whether the man is replaced by a machine[13]. Such replacement of the man by a machine in the game can indeed result in a different outcome, which can justify a reassessment of the relative positions of the humans playing the game, and also the way in which both "the imitation game" and the concept of game are to be conceived and understood. Drawing a conclusion from the previously mentioned difficulty related to the use of common definitions of words, Turing replaces the question "Can machines think?" by another "which is closely related to it and is expressed in relatively unambiguous words." The questions, in fact, the allegedly equivalent questions, are the following:

> "We now ask the question, 'What will happen when a machine takes the part of A in this game?' Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, 'Can machines think?'" (Turing, 1950, p. 434)

Turing proposes in this way to reconceive the relations between concepts and applications. A satisfactory answer for the question "Can machines think?" could involve a reconception of our concepts both of humans and machines. That the interrogator is not in the vicinity of both the machine and the woman, rules out a sexist misunderstanding of the expression "who is a human". The remarkable point to which Turing draws attention to is that "A machine can be constructed to

[13] For historical and philosophical accounts of "Turing machines" see (Kennedy, 2021; Floyd, 2021, Mundici and Sieg, 2021).

play the imitation game satisfactorily" (Turing, 1950, p. 435), that is to say, a machine can be conceived and constructed to lure an interrogator into thinking that a woman is a man (no essentialism involved). Turing's objective is indeed to render clear that automated and closely approximate replications of human actions by machines can be achieved (that is, indirectly by humans) (Turing, 1950, p. 438). Turing's argument involves the acknowledgment that a machine that can replicate the behaviour of any discrete-state machine can be produced: "Provided it could be carried out sufficiently quickly the digital computer could mimic the behaviour of any discrete-state machine" (Turing, 1950, p. 441).

A few conclusions can thus be drawn if Turing's clarification that machines can necessarily rightly be ascribed thoughts and actions is accepted: it would be a mistake to suppose the possibility of beneficially reducing Turing's problem to itself without considerations of application. The problem raised by the question "Can machines think?" does not, and could not reduce to the conceivability of the affirmation of the indistinction of machines and humans, or to the negation of the distinction of machines and humans. That there are games at which humans and machines can play together and which can both be won and lost by humans and machines does not imply that the distinction or difference between humans and machines can (relevantly or without loss) be rejected. On the contrary, the commonality of such situations implies that the personification involved by the humanization or biologization of machines as robots (as when we say of a machine or robot that such machine achieves actions, as sweeping or similarly) could not imply its own literality.

It is relevant is to say of a machine or robot that such machine or robot achieves actions (which could be achieved by humans as well), that actions can relevantly be ascribed to robots or machines, since there is no relevant doubt with respect to the availability of a distinction between machines or robots and humans. The ascription of an action to a machine or robot is derivative in the sense that when a human person ascribes an action to a machine or a robot, that person does not ascribe an action (or expression) to a machine or robot which could eventually transform into or turn out to be a human. For then, there would not be any test of concepts in their relation to their uses or

applications[14]. Actions in such cases are ascribed to a machine which has been constructed to render possible the automated (and eventually) better execution of a task which otherwise would eschew to one or several humans, or of a task which otherwise would remain unachieved by humans (as some human actions necessarily involve the mediation of the actions of machines to be achieved). The obviousness of such a point is probably more easily and better understood if one considers that: mechanisation or metaphorical dehumanization of humans, which involves the depersonification of humans at the occasion of the comparison of one or several machines with one or several humans, also has contraries, or "opposite poles". Ordinary language uses do indeed involve distinguishing between: desirable cases in which, humans are appreciated for their mode of realization or achievement of a task as, or even better than machines (as conceiving an artificial intelligence or winning a game of go), and undesirable cases in which, lived horror or inhumanity of humans is expressed due to their realization of a task or action whose realization by a human necessarily implies the rejection of felt or observed shared human emotions. To this extent, Turing might have, on this point, involuntarily underestimated the resources of our (common, ordinary, everyday) linguistic means, our "natural" languages.

Turing did, since the 50s, envisage the evolution of logical space, the space of possibilities, our possibilities, with respect to the fact that the ascription or attribution of thoughts and actions to machines by humans is unproblematic (Turing, 1950, p. 442). But less obvious is that Turing's evaluation of one's own question can be agreed with under one's own terms. Obviously, the "original" question "Can machines think?" should be discussed – for example, regrettably enough, many people could lose their work if such a question is not publicly addressed[15]. Most probably, it is

[14] The analysis of the test proposed with this article is different and independent from the one proposed by Gonçalves (2024).

[15] That liberatory possibilities involved by the conception and the use of artificial intelligences (as the execution of some tasks can be automated and dispensed with) should not make us forget that the challenges thereby raised present social significance: the realization of the antic dream of the liberation from repetitive work is no more than it was, a wish whose realization would be, as such, relevantly available to every one (Aristotle, 1995, 1253b23).

not senseless to consider that, in the 1950s, the question was too remote from most persons' lives, interests and concerns, to be considered as somehow linked and eventually determinative of their own conceptions of their lives. In that sense, Turing's writing about one's own question that this question is "too meaningless" could eventually be understood. Nevertheless, such evaluation does not, and could not imply that there are, or could be, *degrees* of logic, logicity, logicality or logicalness[16]. In that, Turing's evaluation of one's question is arguably in tension with Turings' own achievements in "Computing Machinery and Intelligence". No unacceptable fact was involved in Turing's expressions of one's own conception of computers. No one would in any way deny jointly that computers have been constructed and do not exist. And remarkably enough, even contrarianist conceptions implicitly addressed by Turing in the objections section, would involve as a step of the conception of their destructive efforts the acknowledgment of the existence of a targeted existence (As the buddhas of Bamiyan). In that, Turing's conception of consciousness and solipsism, I will attempt to render clear, is not, and could not be successful, turn out adequate[17].

Let's consider the "*The Argument from Consciousness"* that Turing wants to contest and of which Professor Jefferson is presented by Turing as a notable defender (Turing, 1950, pp. 445-447)[18]. The argument is that if a machine could write a poem or compose a musical piece *because of thoughts and emotions felt*, then at least one machine could think or be considered conscious, and therefore

[16] On this, Lewis' criticism of the alogical is to be reminded: "Sometimes we are asked to tremble before the specter of the "alogical" in order that we may thereafter rejoice that we are saved from this by the dependence of reality upon mind. But the "alogical" is pure bogey, a word without a meaning." (Lewis, 1929, p. 246).

[17] The first, narrowly theological objection considered by Turing, consists in denying that animals or machines can think on the basis of the affirmation that only humans (by contrast with animals and machines) have souls or are soulful, and that thinking is a function of the soul (Turing, 1950, p. 443). This objection is of little interest for the problem posed and addressed in this article, and Turing's question: "How do Christians regard the Moslem view that women have no souls?" at best is expressive of a distorted picture of Islamic cultures. As a clue of a conceivable reply, the falsity of the question can be established by the true affirmation by a person, whether a Muslim believer or not, to know someone who is Muslim and believes that women have souls, that women are soulful.

[18] For a study of the context of the debate between Turing and Jefferson see Gonçalves (2024, Sections 4.6 and 5.5).

machines could think, or be considered conscious. But, and the following was right when expressed by Jefferson quoted by Turing, machines have not achieved such artistic productions. Thus, machines do not think, are not to be considered conscious. Maybe Professor Jefferson would have liked to add: "because machines cannot experience, feel and act as we – humans – do" (underlining mine), but such addition would arguably render clearer a tension internal to Professor Jefferson's conception. Turing's interpretation is that such argument consists in a rejection of the validity of the test. Turing achieves to render clear a difficulty in simultaneously attempting to maintain that machines can fail humans into determinate misidentifications (that is to say, machines not only can make someone believe that someone is anyone else, but also make someone believe that someone is someone else), and that machines cannot think: only if machines think can these achieve an action which is inconceivable without previous reflexions. The objectivation of the realization of such failing of a human by a machine can be successfully achieved and verified, the loss of a human face to a machine noted both by machines and humans. In that, what was later to be called the "Turing test" is formally valid, as its falsification is conceivable, and the criteria of the test are public and publicly acknowledgeable by relevant expert practitioners.

Nevertheless, less clear is that "the argument from consciousness" consists in a rejection of the formal validity of the test, except maybe, the last move involved by Professor Jefferson's reply, which involves presenting something undone as something that cannot be done, the presenting of a limit as a restrictive limit. For, the reflexions involved by the (derivatively) intentional aspect of the failure of humans by machines, are, strictly speaking those of the (eventually) other humans who conceived and constructed the considered machine, rather than only or merely the achievement of the machine considered in isolation from its conceivers and producers.

Otherwise put, the transition from the question "Can humans fail other humans into thinking that machines are humans, by conceiving and constructing machines which can lure humans into thinking that machines are humans *as good as* humans who can lure other humans into thinking that someone is another?" to the question "can machines fail humans into thinking that machines are

humans?" is at least unclear, not to say undue or illegitimate. For, even in the intricate case in which the machines (which can lure humans into thinking that machines are humans) have been conceived and constructed by other machines (conceived and constructed by humans to lure other humans into thinking that machines are humans), it is not rendered true that machines self-conceived themselves by themselves – that is, autonomously in an underivative sense – to lure humans into thinking they are humans rather than machines[19]. Quite the contrary, *only inasmuch as humans conceived machines*, which can conceive other machines, which can lure humans into thinking that machines are humans, can it be rendered true, and not only in a narrow experimental sense but in a historically accurate way, that humans can be *lured* into thinking that the machines conceived and produced by the machines they conceived and produced are humans rather than machines. To this extent, we should probably reject, not that the test is valid, but rather that the imitation game does consist in a test at all. That is to say, if any test is involved by the "imitation game", this test is different from the presented test (that of the testing of the thinking of machines)[20], and strictly irreducible to the "verification" of the humanity or humaneness of humans (as in "tests" which we are, as internauts, frequently asked to achieve). What is at stake is rather whether the production of a luring situation by a source-of-language, of source-of-language conception could be acknowledged (conception(s) according to which (a) "private language" could be conceived).

To reject the validity of the test, according to Turing would be equivalent – under its most extreme form – to defending solipsism, which *could be,* Turing grants, "the most logical view" (Turing, 195,

---

[19] Putnam considers a similar intricate case in which the question whether robots are conscious is posed about robots produced by other robots, and argues that its answer involves a decision concerning the treatment of robots within one's linguistic community (Putnam, 1975b, pp. 406-407) and not a discovery. This article is fully compatible and in agreement with Putnam's rejection that considerations with respect to the applicability of the concept of consciousness are meant to be decided on the basis of a discovery. But this article does not argue in favour of the conception according to which consciousness-ascriptions to robots could have been without truth value *until* a decision is taken with respect to the question whether robots are conscious.

[20] On this see Davidson (2004, p. 83). And for a criticism of misleading uses of the argument see Descombes (1995, p.156).

p. 446). So not only that there could and would be degrees of logic, of logicity, of logicality, of logicalness, but there could also be *consistent solipsism*, with solipsism defined as the thesis according to which the *only* way to know *that* – the fact that – someone thinks is to *be* that (particular or individual) person and feel oneself thinking. A parallelism, an analogy, could be made with the case of machines: the only way to know *that* a machine thinks is to *be* the machine and feel oneself thinking.

The motives of Turing's partial objection to Professor Jefferson's eventual objection (as Turing agrees with Professor Jefferson against solipsism) can then be brought out: verification of whether machines are humans is impossible. Thus, it would be sufficient to lure a human into thinking that a machine is a human to establish that machines can think. And among many reasons that can be provided, some of which have been previously explained, such would both be too much and not enough, especially since the premise according to which verification of whether machines are humans is impossible is left uninterrogated. Prof. Jefferson's defence of the "argument from consciousness", according to Turing, then amounts to verifying whether machines are humans is impossible. But suppose such verification could consist in an artistic production by – literally – a machine. Until such production is achieved, that machines can think will not have been established. If Prof. Jefferson were right at that time, then today, Turing would be right and Prof. Jefferson would be wrong since artistic productions produced by machines have failed even expert juries. But it is remarkable that the victory has nevertheless been attributed to a human (by contrast with cases in which victories were attributed to an artificial intelligence as in, for example, the games of chess or go). Nevertheless, could the production of a luring situation of humans by machines be conceivably determinative as Turing argued for? This is, at best, unclear. A reappraisal of the conception of solipsism presented by Turing will prove important, necessary, and beneficial. For Turing both grants that solipsism could have been "the most logical view to hold", and that the only problem involved by such a "view" would be that communication would be rendered difficult. Not only that Turing does not address the question of the logicality, or logicity, or logicalness, of

solipsism, or of whether solipsism could be logic or logical[21], but also, and more importantly Turing neglects both the initially non-philosophical and philosophical conceptions and criticisms of solipsism.

The problems raised by solipsism were indeed not reducible to difficulties of communication. Even only according to the analyses of Lewis, that the difficulties raised by solipsism are very concrete, as concrete as the negation or denegation of the reality of pain involved by contrarianisms and denialisms, among which behaviourism, has been shown. Not only that the problem posed by non-philosophical and philosophical solipsism(s) is not reducible to Turing's conception, but also undue belief in such reductive conception of solipsism can lead to the neglect of solipsism, even *to* solipsism, and this, even despite Turing's achievements.

> "In short then, I think that most of those who support the argument from consciousness could be persuaded to abandon it rather than be forced into the solipsist position. They will then probably be willing to accept our test." (Turing, 1950, p. 447)

Turing indeed assumes that there would be an exhaustive alternative between two opposite possibilities. Either we abandon the argument from consciousness – as we would need to be persuaded to abandon the argument from consciousness to integrate the results from the Turing procedure, or to be persuaded that we cannot integrate the results from the Turing procedure if we maintain the argument from consciousness. Or we are forced into the solipsist position.

And indeed, if we grant both: that the argument from consciousness and the results of the Turing procedure are not compatible, and, that we need to reject solipsism even if that involves rejecting the argument from consciousness, then quasi-unavoidably, the conclusion seems to follow from the premises: we probably will accept Turing's test, allegedly "our test". But an undue dichotomism, or at least, an undue use of a dichotomy in a non-dichotomic case is involved by Turing's conception both of consciousness and solipsism. Indeed, the whole "pressure" exerted on defences of

[21] The negative answer is involved by the negative answer to the conceivability of an exclusively private language, a philosophical result that is not explained in this paper (On this, see Uçan, 2016; 2023).

consciousness turns around the ambiguity involved in the would-be claim according to which "Machines cannot *feel* thoughts and emotions". That is to say, the phrase can be used both to express that machines do not feel thoughts and emotions as we do (and how could we be surprised about that?), or to remind ourselves that expressions of feelings of thoughts and emotions authored by machines are really produced by machines (as we can be astonished by the similarity of expressions authored by machines and humans). Could we have really meant that machines *lack* the sensibility of humans? This is, I shall try to render clearer, at best unclear.

2.2. Commensurability and incommensurability of the facts of humans and machines

Let us remark that the central range of cases integrated by the Turing procedure, the replacement strategy, is *the range of commensurable actions of humans and machines (automated and eventually automatically)*. That is to say, the Turing test is formally valid in an unproblematic sense, as some actions can be realized by both humans and machines, even if its philosophical relevance can and should be criticized. A machine can perform, realize, and achieve for you exactly that which you could perform, realize, and achieve by yourself (for example, cleaning the floor of a room). But there is a range of cases in which the actions of humans and machines are not commensurable. You can and cannot fly at 900 km per hour at 11 km of height under different descriptions. You can thusly fly with a plane, even if you are not the pilot. You obviously could not thusly fly without such a plane. But the impossibility involved is not 'real' (and even could not be such), and could even less be determinative of a restrictive limitation internal to humans[22]. Turing is, to an extent, clear about this distinction:

[22] On the distinction between the criticism of mechanistic conceptions of the human mind and the usability of a Turing machine as a model for some realizations of the human mind see Putnam (1975a, p. 366; 372).

> "We do not wish to penalise the machine for its inability to shine in beauty competitions, nor to penalise a man for losing in a race against an aeroplane." (Turing, 1950, p. 435)

To be relevantly assessed as successful or failed, won or lost, the actions, performances, and achievements of humans and machines need to be relevantly compared. It would nowadays not belong to our expectations, shared human expectations, for a human to fly without a plane at 900 km per hour at 11km of height. For a human to fly involves the use of a tool or a machine which renders possible the achievement of a flight. Imagining the contrary is not impossible and eventually rather comical. But the missing of the comical in such a case could be tragic. After all, cannot we conceive that the conception and production of planes by humans imply the past acknowledgment of the existence of a restrictive limitation internal to humans? This line of argument, is, I argue, truly addressed by Turing, although relatively indirectly, in "Computing Machinery and Intelligence".

Remarking the range of cases of incommensurable action to humans and machines does not imply that the ends attained by machines (in the sense of the tasks achieved by machines that cannot be done by humans, e.g. exhaustively surveying the data of a mega-database) are *not* the ends of some humans. Such ends can be and are attained by humans only since the attainment of such ends has been envisaged, and conceived, and machines and robots constructed along the lines of such conceptions to render possible their attainment by some humans. The realization of such ends could not be possible otherwise, that is, without the mediation of the past conception and construction of the machines which rendered possible the attainment of ends whose attaining was previously impossible. Machines are practically necessary for the attainment of some ends in this respect. Although some actions achieved by machines are incommensurable with actions achieved by humans, with respect to their realization, as the former can do what the latter could not, the same does not apply to the ends of these actions, which are commensurable. We considered that the ascription of thoughts and actions to machines by humans is derivative (of human ascriptions of

thoughts and actions among themselves, rather than from past conceptions and productions of machines). The ascription of ends to machines by humans, the self-ascriptions of ends by machines, the ascriptions of ends by machines to other machines, or even to humans, are likewise derivative of humans' ascription of ends to themselves. The ends of the machines could not, as such, be alien to those of humans. To this extent, the evaluations of the commensurability or incommensurability of the facts of machines and humans is circumstantial. Humans, as such, would neither be limited nor unlimited without machines. And similarly, without humans, as such, machines would neither be limited nor unlimited.[23]

2.3. The limits of metaphorical expression

We considered that the affirmation that machines (sometimes) (metaphorically) think is unproblematic. The important pivotal point is the distinction between the metaphorical and the literal senses of our comparative claims about machines and humans, eventually via the mediation of a comparison of some of their aspects. To say that someone is a machine is neither necessarily problematic (consider the case of the use of a metaphor to express the appreciation of a modality of an action's achievement) nor necessarily unproblematic (case of would-be unmetaphorical use for expression of depreciation, eventually expressive of lack of expectable emotion). To say that a machine is someone is neither necessarily problematic (case of the use of a metaphor to express the appreciation of the similarity, of the accurateness of the replications by a machine, of someone's human behaviours) nor necessarily unproblematic (case of would-be unmetaphorical uses involving identity confusion, a case different from the cases considered by Turing). Similarly, to say that a machine thinks is neither necessarily problematic (case of the acknowledgment of the machine-

[23] These remarks are entirely compatible and in agreement with Putnam's rejection of the unavoidability of a trilemma concerning the application of the concept of consciousness to robots: it is at best unclear that we could have been bound either to affirm that robots are conscious, or deny that robots are conscious, or express our unavoidable ignorance with respect to the eventual truth of the question whether robots are conscious (Putnam, 1975b, p. 407).

mediated realization of a task or action, whose realization without one or several machines sometimes is and sometimes is not conceivable), nor necessarily unproblematic (cases of the depreciation of a human by others, and of the solipsistic others-as-tools conception).

To this extent, although Turing was right about the unproblematicity of the affirmation that machines can think and (sometimes) think, the philosophical acknowledgability of Turing's reconception and displacement of the problematic of solipsism can and is to be contested. The problem with solipsism never was and could not have been merely reducible either to the correct identification of the thoughts of a person or to the correct observation of the achievement of the activity of thinking by a person.

The least that can be said is that the central aspects of solipsism, brought out by notably by Wittgenstein (also by Sartre and Putnam, but it is unsurprising that Turing did not discuss their works) are neglected.[24] The aspect that is centrally neglected, and which is elucidated by Lewis' critical conception of solipsism (and as we shall study, by the one of Wittgenstein as well) is the consideration of the eventuality of the experience of pain. Too much (or not even anything) is done by Turing about solipsism by granting that solipsism could be "the most logical view". Although, under one's own terms, Turing's focus on an aspect of solipsism is understandable and relatively beneficial, such focus and such reconception of solipsism has arguably contributed to the substitution of a thin non-philosophical conception of solipsism to previous philosophical and critical conceptions of solipsism, if we consider the influence of "Computing Machinery and Intelligence", and of Turing's works and achievements. In this sense, the reconception and the displacement of the problematic of solipsism proposed by Turing is not philosophically receivable, or acceptable. For rejection of asymmetrical pain ascriptions resulting in delusive false impossibilities does not imply, could not imply the rejection of the relevance of the

[24] No would-be « argument of authority » is involved by such expression of the issue. Putnam himself achieves the criticism of the intelligibility of the argumentative dimension of the would-be argument of authority, which nevertheless does not, could not consist in a mere rejection of specialization, authorship, authority, truth, and history (Putnam, 1996, p. 233).

acknowledgment of asymmetries between humans and machines with respect to ascriptions of pain. The phrases “machines cannot feel” and “machines do not feel” could not conceivably be reduced to each other and attempts to reject such irreducibility, I shall try to render clearer in the third part of this paper, cannot but turn out delusory.

To answer the question of the limits of the comparison of machines and humans thus involves considering the dimension of successfulness of the achievement of the comparison – its “performative dimension” – so to speak. Comparisons can be successfully achieved. Reflexion concerning circumstances in which the realization of comparisons of humans and machines turns out to be successful can obviously also be achieved. This point matters to remark and address several important difficulties in Rorty’s account of metaphor explained and used by Kennedy, in a wonderful article entitled “Gödel, Turing and the Iconic/Performative Axis”, difficulties which have been only partially addressed so far, to evaluate the place of the machine metaphor in our languages, cultures, societies, forms of life:

> Rorty’s elaborate account of metaphor, of the way metaphor operates in language, is useful here. Metaphors, for Rorty, are “private acts of redescription” originating “outside” of language -“outside”, metaphorically, in the sense of unintelligibility; and his account turns on the idea of the literalized metaphor, literalization being what happens when a metaphor breaks into sensibility; when a phrase like, for example, “point of view” comes to mean something like an attitude toward something—becomes, in other words, literalized:
>
> > Between . . . [between living and dead metaphor] we cross the fuzzy and fluctuating line between natural and non-natural meaning, between stimulus and cognition, between a noise having a place in a pattern of justification of belief. Or, more precisely, we begin to cross this line if and when these unfamiliar noises acquire familiarity and lose vitality through being not just

> mentioned . . . but used: used in arguments, cited to justify beliefs, treated as counters within a social practice, employed correctly or incorrectly.
>
> Rorty sees the creation and literalization of metaphors as the "fuel of liberalism", and "a call to change one's language and one's life". As such, metaphors are a sign of the viability of a shared social practice; evidence of the ability of that practice to continually transform itself, to produce new meaning, through the creation of metaphors. (Kennedy, 2022, pp. 3-4, underlining mine)

Rorty's conception, as explained by Kennedy, involves several assumptions concerning the place, the origin, and the integration of metaphors in ordinary linguistic practices. Metaphors would be acts of redescription, new and different acts of description of whatever has been described utilizing a nonmetaphorical expression: the imaged or metaphorical expression would be a new description. Such redescriptions would somehow be "private", and correlatively, the origin of such acts could be mysterious as such. However, what would such redescriptions be redescriptions of? Of a non-imaged, non-visual, non-metaphorical expression, of a possible or actual literal use of language? On Rorty's account explained by Kennedy, the reply to such questions is relatively ambiguous and yet would somehow turn out successful (as metaphors would be signs of the viability of a shared social practice).

But, as earlier raised by Sartre (2003, pp. 536-538), the question whether language could be an author, by itself, ought to be raised anew: Could the personification of language be as such relevant at all? Even more problematically, a difficulty, earlier evoked, that I will not attempt to address in this paper, but of which an aspect is relevant our consideration of the difficulty raised by such account of metaphor, is the problematic of private language. For distinct ways of considering the relevance of the use of metaphors are, for explanatory purposes, distinguishable.

One thing is to acknowledge metaphorical expressions as direct or indirect expressive means, another thing is to suppose considering metaphorical expressions means as unavoidably indirect

expression means (one could not but use a metaphor in some would-be noncontrastive sense). Then, to present metaphors as “private acts” could seem to call into question the availability of our medium of expression, language, if any such distinction is supposed. And for a very simple reason, in fact, as one thing is to affirm that we indirectly use a metaphor to express whatever we could not have affirmed, cannot affirm otherwise (for we have not found a non-metaphorical expression to our metaphorical expression, although we can), and in such case there is no such thing as an implicit would-be exclusion of a possibility that is not possibility involved.

But another thing would be to suppose ourselves able to affirm that we use a metaphor to express whatever we could not have conceivably affirmed otherwise. The concept of the literalization of metaphors, of literal metaphoricality is to this extent “a double-edged sword”, that is to say, certainly not a risky weapon, due, allegedly, to the sharpness of both of its edges. But a concept that is similar to a weapon, a sword, whose edges both cut, and which has relevant ways of effective handling.

To suppose ourselves to be granting that we unavoidably have to use metaphors because we cannot express – ourselves – whatever we suppose ourselves able to be willing to express, otherwise than by employing metaphors, amounts to underestimating both our possibilities of expression and our (eventual) successfulness in our searches for new or better means and ways of expressing ourselves. The Rortian approach advocated for, to an extent, by Kennedy, is thus not satisfactory ; not that such a conception is ‘risky’ or ‘dangerous’ as such (the criticism I make is one of intelligibility, but not moralistic), but that such a conception of the evolution of language (acknowledged in a way that is partially congruous with Wittgenstein’s philosophically pragmatic remarks in *Philosophical Investigations* concerning the fact that language, language-uses, change), may lead to confusions, sometimes somehow involved by aspects of Turing’s article, such as that of the confusion of humans with machines and inversely.

Although the limits of intelligibility indeed extend with our expressions, our actions, our doings, such a remark could not have implied that any linguistic use, any action, or any doing necessarily

consists in an extension of the limits of intelligibility by itself in a relevant sense. Remarking that a fact is historical (its historicity, so to speak), by means of the remark of the compatibility of the expression of a fact with relevant, accurate, more comprehensive and extensive historical narrations could not relevantly be equated with the creation, the conception, or the production of a new way of understanding, doing, explaining, or achieving. The difficulty with the assumption or the supposition of "degrees" of logic, logicity, logicality, logicalness is not 'after all' a difficulty related to our incapacity to distinguish between conceivable or actual courses of actions which are more or less relevant, or even adequate for the attainment of some ends. Quite the contrary, the difficulty is rather that there could not be such a difficulty and arises from an eventual tension between the acknowledgment of the existence of diverse systems of logic, world-conceptions, and the uniqueness of a way that is our own to understand.

3. The critical conception of solipsism of Wittgenstein in the *Philosophical Investigations*

3.1. Dissolution of the problem raised by functionalism and reductionism: A 'thought experiment' by Wittgenstein.

> "420. But can't I imagine that the people around me are automata, lack consciousness, even though they behave in the same way as usual? — If I imagine it now—alone in my room—I see people with fixed looks (as in a trance) going about their business—the idea is perhaps a little uncanny. But just try to keep hold of this idea in the midst of your ordinary intercourse with others, in the street, say! Say to yourself, for example: "The children over there are mere automata; all their liveliness is mere automatism." And you will either find these words becoming quite meaningless, or you will produce in yourself some kind of uncanny feeling, or something of the sort.

> Seeing a living human being as an automaton is analogous to seeing one figure as a limiting case or variant of another; the cross-pieces of a window as a swastika, for example."

Wittgenstein expresses in §420 that assuming some symmetry between pain ascriptions to humans and machines, between consciousness ascriptions to humans and automata, results in false and eventually delusory impossibilities. Before pressing this point, let us recall that:

(1) Automata are machines which have been built to achieve some actions by themselves, once somehow activated. Once built and activated, the realization of foreseeable and foreseen actions of automata does not depend anymore on (although it is eventually controllable by) their conceivers, productors, and activators. This contrast is involved by the very intelligibility of our eventually successful ascriptions of failures to automata. In such cases whatever was to be relevantly considerable was considered at some stage to render possible the achievement of an action by an automaton, and yet the predicted outcome, the successful achievement of an action by an automaton, did not result from its conception, production, and activation. In this sense the failure of the achievement of an action by an automaton is ultimately intelligible and understood by us derivatively. Nowadays automata, robots, come with and under warranty. We would not take responsibility for each conceivable failure of the functioning of an automaton, of a robot, even if under some description we are the one or ones who have failed to make the automaton function.

(2) "Consciousness" as used by Wittgenstein in this paragraph of the *Investigations* both is and is not used as in phenomenological conceptions under different descriptions. If by consciousness we mean, as in many phenomenological conceptions and accounts, a moment of mental life eventually correlated to irreducibly lived moments (by us or others), as in expressions such as "consciousness of happiness", "consciousness of joy", "consciousness of sadness", that we could express also otherwise, then Wittgenstein's use of the notion of consciousness in this passage is not phenomenological in the sense previously defined. But if by consciousness we mean, the fact that

we can gain consciousness, take consciousness, that at such time and place, I, you, us, them is happy, joyful, or sad, rather than allegedly remarking from within 'isolated' or 'separated' ourselves that happiness, joyfulness, sadness is somehow 'happening' in ways in which not only are remote but cut from ourselves, unavailable to ourselves, separated from ourselves, then surely Wittgenstein's use of the notion of consciousness is at least compatible with such phenomenological conceptions of consciousness.

(3) According to the traditional conceptions of soulfulness or consciousness, to have a soul or to be conscious *is* to be a soul or to have (a) consciousness (See Sartre, 2003, pp. 127-129; pp. 310-315, p. 619). The derivative and metaphorical "property" of a soul would have eschewed to each of us as a result of some attribution about which nothing could conceivably have been done by us – humans, an attribution about which several narratives exist. And in any case, as a result of such an endowment, we necessarily would *have* in ourselves, and ourselves *be what* necessarily could *not* be *had* in themselves by such existents which are not human, and could anyway not have *been* such existents. Consciousness thusly conceived could be some sort of additional ingredient or substance presented by some existents eventually encountered within visual space, and which could and would be in itself *lacking* from other existents eventually encountered within visual space. To render the point clearer: no essentialization of consciousness is involved by such an expression: in fact, quite the contrary.[25] Such lack both can and cannot be observed by us humans who are soulful or conscious, as we could understand that we are provided, endowed, or gifted with exactly the soul or consciousness that could not have been provided, endowed, or gifted to other living existents. Correlatively, we could not have provided the soul or consciousness that we were – as humans – to tools, or objects we construe, as automata, as machines, as robots. This would be an impossibility

[25] On this, Sartre was and is right against Heidegger: if anything is metaphorically 'essential' to consciousness, that is non-coincidence with "itself". Expressed otherwise: according to the traditional picture, animals lack a soul or conscience; their reality is intelligible and accessible to us only negatively and privatively. But although the consciousness of animals might arguably be firstly intelligible and accessible to us negatively, it is at best unclear that such consciousnesses could need to be rendered intelligible and accessible to us privatively.

we could not but acknowledge were we to understand our ‘natural’ place. But we can nevertheless imagine how wonderful would be for such existents to be provided with – like us – a soul or a consciousness.

This is the sense of ‘lack’ involved by Wittgenstein’s ‘thought experiment’ at the beginning of §420. That we can analogically or metaphorically envisage that artificial existents, as automata, robots, or machines, *could* have been provided a soul or consciousness, if these existents *had been* humans involves our acknowledgement that, in fact, these existents could not have been provided a soul or consciousness, as these existents are not humans. Even if we can imagine that these existents could have wished to be provided a soul or consciousness, although these could not have had a soul or consciousness, a soul or consciousness could not have remained unwished-for by these existents if these existents could have imagined consciousness or soulfulness. To this extent, such existents would lack precisely the soul or consciousness each of us is or has. Our assessment of these existents would have but to remain oscillating, once and for all (we could be condemned to idle so to speak). Wittgenstein invites us not to remain constrained by such exercises of our imagination. Yes, we also can imagine that other humans are automata, “lack consciousness”. That is to say, we can imagine that people around us, at an occasion, are wrongly assumed by us, not to be automata, machines, or robots. It is sufficient to imagine that the substitution or replacement of humans by automata, machines, or robots would have been achieved with automata, machines, or robots whose actions would replicate, mimic, or reproduce the behaviours – actions – of persons whose behaviours – actions – are replicated so well, so accurately, that the substitution or replacement would remain undetectable by us. However, if we try, such automata would be ‘logically’ indistinguishable from humans and inversely (at least according to the traditional conception of logic addressed by Frege, Russell and Wittgenstein). Importantly, whether we are imagining to be with others (who in fact are not others but automata) while we are not, or are with others (who in fact are not others but automata) while we are, does not change the ‘thought experiment’ and its

outcome. For in neither case, could be rendered true that humans are automata, or automata are humans, in ways in which we so far, until now, could have failed to notice, to discover.

Yes, we can *imagine* that we wrongly assume that machines *are* humans, but imagining such a case involves reconceiving what the holding of such a case, the happening of such a fact, would consist in. The distinctions between humans and automata would not thereby be rendered unavailable. The availability of such distinctions would remain implied by the intelligibility of the situation as such (one's hesitation with respect of the identity of the existents in the surroundings, one's discovery of a failure to identify a human or a robot). That we do not know could not imply in such cases that the truth about the eventual identification mistake could not conceivably be known by us – the realization of the replacement itself would involve the concerted action of several persons. To this extent, the 'merely direct' reading of this passage remains superficial. If §420 only addressed the 'risk' involved by a superficial conception of solipsism and functionalism, then §420 could have been ended with its first question. But this is not the case. Wittgenstein envisages 'in the first person', or invites us to envisage by ourselves, one way in which we could conceive the result of the imaginative exercise of our imagination, in determinate circumstances, the first range of cases considered above.

Let us imagine that we are not with others who in fact are not others but automata, and that we are alone in our rooms, in one's room, and *imagine* that we are with others who in fact are not others but automata. Wittgenstein then expresses a conceivable result of such an imaginative exercise of our imagination, which can eventually be considered as quite deceptive: "I see people with fixed looks (as in a trance) going about their business". Such a description of an imagined situation could be either very similar or very dissimilar from our ordinary experiences (not experiments): after all, one might or might not have experienced cases in which the focus of persons with whom one works seems very irrelevant or very relevant. But, more centrally, could not one have expected the outcome of the imaginative exercise of one's imagination to be seeing-automata-and-not-people? Was not the case envisaged, the case in which one is wrongly assuming that humans are automata

‘after all’? But importantly enough, such an imagined case does not involve such a conclusion – another case could, but one independent from the former and that we would have to imagine.

Wittgenstein does call attention to the openness and necessarily public conceivability of the result of the ‘thought experiment’. If there is no conceivable way of discovering – and especially as we are considering imagined cases – that humans – conscious existents – are automata – existents which supposedly lack consciousness – then there also is no conceivable way of discovering that automata – existents which supposedly lack consciousness – are in fact humans. Then the realization of the delusiveness of the would-be result of the would-be attempt to distinguish people and automata by presenting asymmetries with respect to attributions of consciousness to humans and to machines as involving reciprocal (and necessarily restrictive) impossibilities, is rendered conceivable: the reductive and ingredientist conception of consciousness, the conception according to which consciousness could exist as an ingredient of some bodies, is necessarily misleading.

Wittgenstein then invites us to interrogate ourselves with the eventual feelings we could experience, if we would imagine such a result, and notably the feeling of uncanniness. One might ‘after all’ remain unconvinced by one’s own realization that the imagination of a case of delusory confusion of humans with automata, or machines or robots, does not, and could not amount to the establishment of the eventuality of the relevance of such confusion as such. Could not, and should not some feelings constitute (metaphorical) grounds on the basis of which we could and should reject that humans could be machines or that machines could be humans?

That moralistic resolution of the problem is rejected by Wittgenstein. For we can realize the meaninglessness of the feeling of uncanniness produced by means of the meaningless use of some of our words (the case of would-be attempt of reduction of children to mere automata by *consideration* of their reducibility to mere automata, turns out ‘ineffective’ in would-be ‘optimal’ circumstances, that is, in the vicinity of children), and thus the non-conclusiveness of the delusory outcome of the ‘thought experiment’ can be realized. That is to say, the words by means of which we supposed an understanding of the reality of the situation to be rendered available to us, lose their

sense as we understand that such use of words were not rendering anything available except a misunderstanding of the reality of the situation to ourselves. And we can also realize the correlative meaningless effectivity of the meaningless use of some of our words in the production by us in ourselves of the feeling of uncanniness.

The liberation from the would-be disjunctive entrapment within a dilemma between the meaninglessness of the feeling of uncanniness and its meaningless production by us in ourselves, does not consist in a conclusion, could not be drawn on the basis of premises, and is nevertheless not unargumentative. Rather, appropriation or reappropriation is realized by us by exhaustive consideration or reconsideration.

This realization renders available a non-psychologistic or non-psychological and philosophical achievement with respect to seeing: to use Wittgenstein's examples, though we could use other examples as well, we can see the cross-pieces of a window as a swastika, see that another figure (necessarily imagined) could be obtained by subtraction of some of its elements to a figure (necessarily perceived) in some cases. Importantly, the example put forward by Wittgenstein is a case in which the figure from which another figure can be obtained presents the dimensions and the elements from which the other figure could somehow be obtained. The figure which can be obtained yet is not, and could not, be reducible to an ingredient of the figure from which such figure can be obtained. For the figures and the ends, if any, achieved by production are not necessarily dependent on each other. Not every figure could be, or is meant to be, obtained from every other figure anyway. Some figures could be obtained by us by using some other figures. But some figures could anyway be obtained from each other.

This remark does not, could not imply restriction, or acknowledgement of restrictive limitation. We can also imagine the figure of the swastika to be completed so as to form the figure of the cross-pieces of a window. To this extent, we can see the figure of a swastika as the variant of the cross-pieces of a window and inversely. But would we consider the realization of the completion – not its eventuality – of a figure, to produce one of its variants, or, the subtraction of the elements of a

figure to produce one of its variants, then each figure is seen by us or imagined by us as a limiting case of the other, *for the operations which are to be achieved to produce one from the other could not conceivably be the same*.

We neither unavoidably could have had to construe the figure of a swastika to construe the figure of the cross-pieces of a window or the opposite, contrary to the assumption of the ingredientist conception. Constraints about figure productions could not have had to be unavoidably thought of as signs of restrictive limitations, and can be thought as unrestrictive limits of modes of conception, production, in cases in which figures are conceived by us, and of constraints – unrestrictive constraints – concerning the modes of conception, production of figures from each other, in cases in which is envisaged the obtaining of a figure from another. The relation between seeing the cross-pieces of a window as a swastika and seeing a living human being as an automaton are similar.

We can imagine, to an extent, the obtaining of the later from the former: the conception and construction of machines, robots, automata have been rendered possible by the subtraction of aspects and dimensions of the lives of humans. It is possible to produce a robot, an artificial intelligence, an automata that replicates aspects and dimensions of the lives of humans. But it is also possible to produce a robot, an artificial intelligence, an automata that does not replicate aspects and dimensions of the lives of humans, for the life of a human, or for the lives of humans. There is not and could not be a common ingredient – consciousness – that would need to be added to some and not others so as to render possible the reversion of the relation: such concept of consciousness is delusive.

To this extent, §420 not only addresses the risk of solipsism involved in the reductionist and functionalist conception, but also the would-be attractiveness of a contrarianist form of reductionism and functionalism, namely “methodological solipsism”. That is to say, if one difficulty is that of the credulity related to a naive form and conception of solipsism, another one is that of the incredulity related to a sophisticated form and conception of solipsism which is “methodological solipsism”, whose distinction from solipsism needs, as earlier remarked, needs to be criticized.

The determinacy of Wittgenstein's concern with solipsism has been in some sense unhelpfully neglected. The recent publication of the Whewell's Court Lectures (Wittgenstein and Smythies, 2017) provided us with important passages in which Wittgenstein expresses one's critical stance concerning solipsism, and the relation between the criticism of solipsism and the problematic of the philosophical relevance of pain:

> "Suppose someone said: 'I am having pain: the other person hasn't got real pain' – Solipsism, solipsistically speaking.
>
> We are up against one definite use of language. If I say, 'Lewy hasn't got real pain', he'll be offended. I'm belittling his sufferings. This I don't want to do.
>
> The answer would be: 'Sometimes yes, sometimes no.' It would be a distinguishing property of language as we know it." (Wittgenstein and Smythies, 2017, p. 115)

Solipsism is definable under its own terms as the denegation of the reality of the pain of others. Such denegation is according to Wittgenstein one "absolutely definite use of language". The first important aspect of the case of pain, the reason for which this case constitutes a hard case, is that the case of pain is a (unrestrictively) limiting case of paradigmaticity and verification. An important paradigmatic aspect of pain is that pain has *degrees*, but the objectivation of (the experience of) pain does not necessarily involve reliance on quantification. This could not mean that a quantificational system cannot be used in order to render objective or objectivate the reality of pain, but that there is no such thing as an unavoidable use of a quantificational system to objectivate and objectively agree about the reality of the eventually high degree of the pain of someone (and for example, to evaluate the need for the use of some drugs to attenuate someone's pain). Pains and degrees of pain can be expressed and measured in diverse ways, and whatever the used measure system is, provided public criteria, the results of the measure will be translatable into other measure systems, eventually with some little loss in accuracy, but negligible loss (and eventually undefinitive) with respect to the ends in which the measurement activities are carried out.

However, the objectivation and the eventual measures of pain imply the acknowledgment of the necessary secondariness of the denegation of the reality of pain. That is to say, we can well imagine or observe that someone fakes feeling or resenting some pain. But such cases are understandable as such against our having internalized the available intelligibility of a primary range of cases, in which, pain is felt and is somehow expressed by someone. It is as pain is felt that pain is expressed and not as pain is expressed that pain is felt. With respect to verification this might seem to cause, generate, induce, or raise a problem: by contrast with other cases of measurement activities, not only that someone's pain is not necessarily perceived, but it also is not always relevantly expectable to be observed or objectivated, except by the mediation of our acknowledgment of the words of others. Verification of pain thus can at least sometimes be assumed to be impossible.

That was the position of the problem addressed, as we earlier studied, by Lewis, and involved by the criticisms made both by Lewis and Wittgenstein of verificationism. For, if it is acknowledged that sometimes verification of the feeling of pain by someone is impossible, then it is not inconceivable that such verification could always be lacking. If words could conceivably be used by others as by oneself to affirm that pain is felt although that is not the case (for some ends, whichever are these), maybe conceivable doubt concerning the expressions of pains of others could always be relevant. If such doubt can at least seem to be always relevant, then a verification could always be missing in the case of pain. But then, even in one's own case, pains could eventually be unverifiable, always probable only, although, one does not see the way in which one could be wrong in expressing one's pains, which are not, strictly speaking, ascribed to oneself by oneself, but expressed by ourselves[26].

[26] On this see Putnam (1975a, p. 362).

3.2. The resolution of the problem posed by C. I. Lewis

Wittgenstein's analysis of the philosophical relevance of pain, and aspects of his dissolution of the problematic of private language responds to a central aspect of the problem addressed by Lewis: the negation of the reality of pain, involved by the solipsistic claim as earlier defined.

Let us recall that Lewis does grant the conceivability of "metaphysical solipsism" and argues in favor of the relevance of a minimal sense and conception of "metaphysics" which is meant to provide some grounding to the false rejection, the wrong calling into question of the existence of a connexion between observed behaviour and pain. Such connexion could be unverifiable and unknowable in the absence of the acknowledgment of the existence of a "metaphysical" connexion which would provide the ground, ensure the existence, of a connexion between an observed behaviour and pain.

We considered that one centrally beneficial aspect of Lewis' conception, which is congruent with Wittgenstein's criticism of solipsism and methodological solipsism, is that his minimally "metaphysical" conception is compatible with any moralistic conception of the veridicity of the expressions of their pains by humans, any conception according to which one must only veridically express that pain is felt by oneself because of some prescription, rule, law internal to a world-conception (or form of life, in Wittgenstein's terms). Indeed, any such conception is compatible with the existence of connexions between behaviours expressive of pains and experiences of pains (by contrast with the cases of machines, robots) and incompatible with fake expressions of pains by persons while no pain is felt by them.

But the force of this conception is also in some sense a weakness. For every connexion between a behaviour expressive of pain and the experience of pain should arguably be *grounded*, inasmuch as if such groundings did not exist, then the claim of which the grounding constitutes the basis would not be grounded. The force and contextual relevance of such conception stems from the establishment of the necessary compatibility of each true conception with each other with respect to shared human needs and interests. The weakness of this conception is related both to the modes of

the conception and to the reply which would arguably be required to be made to “metaphysical solipsism”, to be refuted under its own terms.

For we already considered that in some sense an exhaustive generalization would be, according to Lewis, involved by the legitimate acknowledgeability of a relevant doubt of the existence of a connexion between a behaviour and a pain. That is to say, if such connexion can relevantly sometimes be assessed to be lacking, then nothing precludes that such connexion could always be lacking. A dichotomous approach should nevertheless, according to Lewis, enable us to settle the question: we should be able to assess that: either a pain is connected to a behaviour and reciprocally, or a pain is not connected to a behaviour and reciprocally. A pain cannot be connected and not be connected with a behaviour in the same sense and reciprocally, a behaviour cannot be connected and not connected with a pain in the same sense. A pain can sometimes be connected to a behaviour (for example, one sometimes tells others about one’s headache; others sometimes tell us about their headaches). A behaviour can sometimes be connected to a pain (for example, someone might consider that such and such behaviours and actions are done by a person when that person feels pains in one’s knees, which are not similarly achieved by each one else in such case). Nevertheless a pain is not each time connected to a behaviour (sometimes one does not tell others about one’s headaches; sometimes others do not tell us about their headaches). And neither is a behaviour each time connected to a pain (for example, one can truly consider that another person faked *again* being in pain).

“Metaphysical” anti-solipsism is meant to provide an infallible response to “metaphysical” solipsism: the false denial of the existence or the inexistence of a connexion between a behaviour and a pain and reciprocally must always be wrong. And as a result of the consideration of the comparison between robots and humans, we studied that according to Lewis there is no such thing as relevantly rejecting each conceivable distinction between, or affirming the indistinctness of, humans and robots. That is to say, according to Lewis there should never be a “consistent solipsist” who could make “the monstrous supposition that other humans are merely robots”, as also this

could have for outcome or result the provision of meaning to solipsism although solipsism should not be provided any meaning at all.

In a sense, Wittgenstein invited us in §420 to make exactly the supposition that Lewis invited us to reject – but in way which is different from that of Turing – and which has for first result to liberate us, if required, from the tacit acknowledgement of the eventuality of an event, which all things considered, could not have happened anyway: the transformation of humans into robots and reciprocally as the result of our 'thought experiment'[27]. Imagination is not meant to be restricted in any sense if the issue raised by solipsism can be addressed at all. But Wittgenstein's conception enables us to solve the problem posed by Lewis, with a radically different account of generality, a different account of relations between solipsism and skepticism, and a different account of the requirements internal to the intelligibility of the metaphor of humans as machines.

First, on Wittgenstein's approach, the possibility for a person to fake an expression of pain does not, could not invalidate or disprove, that we express our pains. Quite the contrary, in fact. As mentioned, Lewis's conception does not imply that the first range of cases we need to consider when observing the expression of pain by someone are cases of persons who are faking being in pain. And ultimately, Lewis also rejects methodological solipsism. Methodological solipsism is also as considered by Wittgenstein a sophisticated form of solipsism according to Lewis, but a "metaphysical response" should nevertheless be provided according to him to "metaphysical solipsism".

That is the sense in which "metaphysical" solipsism should be at least in principle always be established to be wrong by "metaphysical" anti-solipsism which necessarily is common to every conception of "metaphysics" compatible with human needs and interests. A relevant contrast between the approaches of Lewis and Wittgenstein can then be spelled out: if on Wittgenstein's approach, it is unclear that solipsism, understood under one's own terms, could be right, in the way

[27] And obviously although the cyborg is neither a mere human nor a mere robot, the consideration of the cyborg case is not as such a sufficient answer to the problem which is not that of the lack of an intermediate case between robots and humans.

solipsism requires to possibly be, according to Lewis, to be right *at all*, then it is no more clear is that the wrongness of solipsism, understood under one's own terms, should be establishable, according to Lewis, for the wrongness of solipsism to be established at all.

In other terms, while Lewis grants the possibility of the truthfulness of a "metaphysical" sort of solipsism to render explicit that any coherent "metaphysical" anti-solipsism can establish its falsity – except if the solipsist is coherent enough in the self-production of one's solipsism, Wittgenstein does not grant the possibility of the truthfulness of a "metaphysical" sort of solipsism which would await its refutation. Solipsism does never, could not start to present the relevance which would justify the acknowledgement of the existence of its false grounds. The negative replication of the difficulties generated, produced, raised, and posed by solipsism could not be relevant at all, and even less, philosophically.

3.3. The resolution of the problem posed by Turing

We have thus reached another crossroad. Both Turing and Wittgenstein invite us to do what Lewis invites us to reject, but not in the same senses. Wittgenstein invites us to reject metaphysical anti-solipsism and methodological solipsism with solipsism, but not the notion of consciousness.[28] While Turing does not invite us to reject metaphysical anti-solipsism and methodological solipsism with solipsism, but does neither preserve the notion of consciousness.

From the outset it can be remarked that some counter-objections envisaged by Turing to one's own argument idle. The contrast between machines and humans could not be blurred or rendered less accurate by the acknowledgment that machines (also) think. Our concept of consciousness does not necessarily, could not have necessarily implied reliance on the presenting of unrestrictive limits

[28] Wittgenstein used the comparison of humans and machines when he defined "Turing's 'machines'" as "humans who calculate" (Wittgenstein, 1947, Ts-229, 448). On this passage see Floyd (2012a, p. 40; 2012b) and Shanker (1987, pp. 615-623), and on the relations of Turing and Wittgenstein see Floyd (2021, pp. 123-126).

brought out at the occasion of the comparison of machines and humans as (restrictive) impossibilities. No one is or should be considered as eventually *forced* or coerced into a solipsistic position, and especially not for the sake of the establishment of the truth of an argument. The criticism of solipsism can be more direct and should be more direct to be addressable at all.

Further, Wittgenstein's evolutive conception of language renders conceivable to think the possibility of compatibility or agreement with respect to the ascription of actions to machines (and artificial intelligences) without calling into question the relevance of the notion of consciousness which is central in world-conceptions (by contrast with the notion of subjectivity), and to account for the distinction between humans and machines whenever required. The production of a luring situation is not, could not be conclusive in the way Turing presented, and Turing's achievements are (hopefully!) independent from an argument that is no more, and could not have been conclusive anyway. One way to express the point made by Wittgenstein is to remark that natural history is both *natural* and *historical,* that our history is not the history of men, but of humans.

Conclusion : Independences and Forms of Life

This article proposed a reflexion about the limits of the comparison, analogy or metaphor between humans and machines. As such, the comparison could not be problematic: humans and machines present common aspects, and instances of such comparison are implicit in ordinary, engineering, scientific, and medical practices. That many progresses have been rendered possible also with, or in ways compatible with the use of this comparison could not have had to be established again. But the extent to which the comparison can be metaphorically literally understood, if it can be metaphorically literally be understood at all, is, as studied, a question whose stakes are of primary philosophical importance. Thematizations of this comparison of the XXth century, whether entirely philosophical – such as those of Lewis and Wittgenstein, or presenting philosophical significance – such as that of Turing, are indeed intertwined with the problematic of solipsism. As much as

linguistic practices are concerned, we considered that there exist appreciative and depreciative ordinary linguistic uses which do involve this comparison, and testify of the available intelligibility of distinct ranges of cases which do not involve, and are not compatible with the confusion of machines with humans.

The first part of this paper presented Lewis' critical conception of solipsism against this background. The affirmation of the indistinction, or the negation of any distinction between machines and humans is necessarily problematic, necessarily misleading or delusive. According to Lewis, a minimally "metaphysical" conception is required so as to disprove "metaphysical" (and methodological) solipsism which consists in the negation of the existence of connexions between pains and behaviours expressive of pain.

However, we considered in the second part of the article, that the successful establishment by Turing that machines can be unproblematically be said to think – notably by means of the introduction of the "imitation game" – involves the assumption of a disjunctive entrapment between either defending that machines think, or, defending both consciousness and (a reconceived and unphilosophical conception of) solipsism. Fully acknowledging Turing's criticism of a traditional conception of consciousness according to which machines would be *deprived* of thoughts and emotions, we nevertheless considered that Turing's reconception of solipsism contributed to the substitution of an unphilosophical conception of solipsism to a philosophical conception of solipsism in an undue way.

Indeed, philosophically accounting for consciousness could neither necessarily involve to grant that machines *could* be *deprived* of thoughts and emotions, nor to defend solipsism. The criticism of such unavoidable disjunctive entrapment is achieved by Wittgenstein in the *Philosophical Investigations,* as studied in the third part of this article. Indeed, reciprocal asymmetries with respect to attributions of pains and consciousnesses to humans and machines are inconceivable when presumed as involving reciprocal and necessary restrictive impossibilities. Infallible response to "metaphysical" solipsism then is no more than "metaphysical" solipsism, required or relevant to

address the problematic of solipsism. At stake is no less than our conceptions of science, of diversity and of forms of life: scientism could not substitute for science, exclusion could not be compatible with diversity, forms of life could not be compatible with solipsism. The relatedness of some forms of life could not imply the mutual dependence of each form of life with each other. This can seem to be incompatible with ecology, but, on the contrary, is not: wholistic reflexion is not and could not be based upon mechanistic reductionism.

Bibliography:

Aristotle. *Politics*. (1995) Edited by R. F. Stalley. Translated by Ernest Barker. Oxford & New York: Oxford University Press.

Bouveresse, J. (2022). *Les vagues du langage*. Lonrai: Seuil.

Copeland, J. (2004). *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life: Plus The Secrets of Enigma*. Oxford: Oxford University Press.

Descartes, R. (2006) *A Discourse on the Method*. Translated by Ian Maclean. Oxford & New York: Oxford University Press.

Davidson, Donald. (2004). "Turing's Test." In *Problems of Rationality* (pp. 77–86). Oxford: Clarendon Press.

Descombes, V. (1995). *La Denrée Mentale*. Minuit: Paris.

Floyd, J. (2012a). Wittgenstein, Carnap, and Turing: Contrasting Notions of Analysis. In P. Wagner (ed.), *Carnap's Ideal of Explication and Naturalism* (pp. 34–46). London: Palgrave Macmillan. https://doi.org/10.1057/9780230379749_4.

———. (2012b). Wittgenstein's Diagonal Argument: A Variation on Cantor and Turing. In P. Dybjer, S. Lindström, E. Palmgren, and G. Sundholm (eds.), *Epistemology versus Ontology* (pp. 25–44). New York & London: Springer. https://link.springer.com/book/10.1007/978-94-007-4435-6.

———. (2021) Turing on 'Common Sense': Cambridge Resonances. In J. Floyd and A. Bokulich (eds.), *Philosophical Explorations of the Legacy of Alan Turing*, (pp. 103-149). Cham: Springer International Publishing. https://link.springer.com/book/10.1007/978-3-319-53280-6.

Gonçalves, Bernardo. (2024). *The Turing Test Argument*. New York & London: Routledge.

Kant, Immanuel. (2007). *Critique of Judgment*. Edited by Nicholas Walker. Translated by Martin Meredith. Oxford & New York: Oxford University Press.

Kennedy, J. C. (2021). Turing, Gödel and the 'Bright Abyss.' In J. Floyd A. Bokulich (eds.), *Philosophical Explorations of the Legacy of Alan Turing* (pp. 63–92). Cham, Switzerland: Springer International Publishing, 2021. doi.org/10.1007/978-3-319-53280-6.

———. (2022). Gödel, Turing and the Iconic/Performative Axis. *Philosophies* 7 (6), 141. https://doi.org/10.3390/philosophies7060141.
Lewis, C. I. (1923). A Pragmatic Conception of the A Priori. *The Journal of Philosophy* 20 (7), 169-177. https://doi.org/10.2307/2939833.
———. (1929). *Mind and the World Order: Outline of a Theory of Knowledge*. New York: Dover Publications.
———. (1934) Experience and Meaning. *The Philosophical Review* 43 (2), 125–46. https://doi.org/10.2307/2179891.
———. (1970). *Collected Papers*. Edited by J. D. Goheen and J. L. Motershead. Stanford: Stanford University Press, 1970.
Onfray de la Mettrie, J. (1996). *Machine Man and Other Writings*. Translated by Ann Thomson. Cambridge, New York, Melbourne: Cambridge University Press.
Mundici, D., and W. Sieg. Turing, the Mathematician. In J. Floyd and A. Bokulich (eds.), *Philosophical Explorations of the Legacy of Alan Turing*, (pp. 39-62). Cham: Springer International Publishing, 2021.
Putnam, H. (1975a) Minds and Machines. In *Mind, Language and Reality* (pp. 342–61). Cambridge, New York, Melbourne: Cambridge University Press, 1975.
———. (1975b) Robots: Machines or Artificially Created Life?. In *Mind, Language and Reality*, (pp. 386–407). Cambridge, New York, Melbourne: Cambridge University Press, 1975.
———. (1996). Why Reason Can't Be Naturalized. In *Realism and Reason* (pp. 229–47). Cambridge: Cambridge University Press, 1996.
Sartre, J.-P. (2003). *Being and Nothingness*. Translated by H. Barnes. London: Routledge.
Shanker, S. G. (1987). Wittgenstein versus Turing on the Nature of Church's Thesis. *Notre Dame Journal of Formal Logic* 28 (4), 615-649.
Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind* LIX (236), 433–60. https://doi.org/10.1093/mind/LIX.236.433.
Uçan, T. (2023). "Autonomy, Constitutivity, Exemplars, Paradigms." *Conversations: The Journal of Cavellian Studies*, (10), 52-79. https://doi.org/10.18192/cjcs.vi10.6613
———. (2016) The Issue of Solipsism in the Early Works of Sartre and Wittgenstein 2016. *University of East Anglia Digital Repository.* https://ueaeprints.uea.ac.uk/id/eprint/62314/1/2016UcanTUPhD_%282%29.pdf.
Wittgenstein, L. (2009). *Philosophical Investigations*. Edited by P. M. S. Hacker and J. Schulte and translated by G. E. M. Anscombe, P. M. S. Hacker, and J Schulte. Oxford: Blackwell.
———. (1947) Ts-229,448 Facsimile 1947. http://www.wittgensteinsource.org/
———. (2003). *Tractatus Logico-Philosophicus*. Translated by C.K. Ogden. New York: Barnes & Noble.
———. Tractatus Map. University of Iowa Tractatus Map. Accessed December 21, 2018. http://tractatus.lib.uiowa.edu/.

Wittgenstein, L., and Y. Smythies. (2017) *Wittgenstein's Whewell's Court Lectures*. Edited by V. Munz. and B. Ritter, Malden and Oxford: Blackwell.