Artificial Intelligence

Mariano L. Bianca

Abstract Before analyzing Agazzi's conception of Artificial Intelligence, an historical overview is offered of the different ways in which the problem of imitating or reproducing human thought has been approached since the Middle Ages till now. My review of Agazzi's position is mainly based on two essays: "Alcune osservazioni sul problema dell'intelligenza artificiale" (Some observation on the problem of the artificial intelligence, 1967) and "Operationality and intentionality: the missing link of the artificial intelligence" (1981). In these Agazzi carries out a deep analysis and criticism of the different approaches to this topic (especially computationalism and functionalism) and outlines his own articulated conception. In his view today's computers (or intelligent artificial systems in general) cannot be said to have thought or feeling analogue to those of human beings because they are not endowed with intentionality. In such a way Agazzi anticipated of about 15 years this celebrated thesis maintained by John Searle. It is possible to build complex machines capable of realizing performances very similar to those of human reasoning or goal-oriented behaviors. This similarity, however, is confined to the capability of performing certain operations, while the difference that still remains is the capability to give meanings to such operations and their results, and to intentionally propose goals to themselves.

1 A Preliminary Note

Evandro Agazzi has dealt with artificial intelligence, both from the critical and the philosophical point of view, in a few essays, the most salient of which are: "Alcune osservazioni sul problema dell'intelligenza artificiale" (*Some observation on the problem of the artificial intelligence*, 1967a) and "Operazionalità e

M.L. Bianca (✉)
University of Siena, Siena, Italy
e-mail: mariano.bianca@unisi.it

intenzionalità: l'anello mancante dell'intelligenza artificiale" (*Operationality and intentionality*: *the missing link of the artificial intelligence*, 1991). We say "salient" because the 1967a paper contains the first extensive critical treatment of the issue of artificial intelligence proposed by Agazzi, with particular insistence on intentionality as the decisive characteristic that distinguishes human intelligence from machine or artificial intelligence. In such a way he was anticipating of 15 years this thesis that is often credited to John Searle. Actually Agazzi had presented his paper in English at the "Wiener Memorial Meeting on the Idea of Control" held in Genoa in 1965, whose proceedings however never appeared. Therefore in 1967 he published an Italian translation of that paper, that was also reprinted in Agazzi (1978), and in addition presented his ideas in a shorter paper at the 21th National Congress of Philosophy (1967b). Agazzi had the opportunity of taking up again and expanding his ideas in the paper "Intentionality and artificial intelligence" presented at a meeting on "The Mind-Body Problem" organized by the International Academy of Philosophy of Science in 1980 and whose proceedings appeared as a special issue of *Epistemologia* (see Agazzi 1981). An Italian translation of this paper was later published (Agazzi 1987).

Agazzi's treatment of artificial intelligence reflected also, from the very beginning, his strong focus on operations in scientific epistemology. Therefore he published a paper in which both intentionality and operationality are stressed as central concepts in the debate on artificial intelligence (Agazzi 1991) closing in such a way the circle of his discussion of this topic. The 1991 paper was newly reprinted as Agazzi (2010).

This brief historical reconstruction explains why, in the present contribution, quotations are taken only from Agazzi (1967a, 1991), that is from the initial and the final points of this trajectory. Indeed the 1967a seminal paper contains the priorities and originality of the whole treatment devoted by Agazzi to artificial intelligence, while the 1991 paper stresses that strict connection between intentionality and operationality that is the backbone of the whole reflection on human knowledge developed by Agazzi in his work, a connection whose significance is well exemplified in the treatment of artificial intelligence. Of course in Agazzi (1981) one finds a more systematic, detailed, analytically deepened and also thematically enriched treatment of the issue, but the real novelties are rather linked with Agazzi's investigations in the domain of logic, semantics, foundations of mathematics to which are specifically devoted other contributions contained in the present volume, so that we did not consider essential to pay them a particular attention.

2 A Few Introductory Remarks

Before examining Evandro Agazzi's theses about artificial intelligence, it is useful to dwell upon the idea of reproducing in any possible way the mental activities of humans and, more particularly, the so called rational or cognitive activities.

The idea of imitating, simulating or emulating human cognitive activities is very old, and can be dated back to the medieval logicians, and particularly to the logic of Ramon Lull and afterwards to the combinatorial logic of Pierre de la Ramée. The works of these authors were based on Aristotle's logic, which basically claimed that human thought (the strictly cognitive thought, at least) develops by logical procedures, like deduction and induction, which could always be reformulated in a rigorous and even formal way. Certainly Aristotelian logic did not include the idea of a reproduction of human cognitive activities, but it contained the idea that such activities are formal procedures, which could have been adapted to any type of cognitive content; in general, this idea was adopted by those who first, in the modern age, formulated the concept of reproducibility of the cognitive activities of man: they considered these activities as the result of formal processes.

Such a conception was also adopted by Ramón Lull, who was certainly the first in the history of philosophy and of logic to pass from considering the cognitive activities as logical-formal activities to believing in the possibility to imitate such activities by mean of techniques or more or less mechanical devices. Lull's logical machines, in fact, were a kind of primitive computers formed by various disks of papers laid one on top of the other, capable to rotate and so to generate different combinations of the symbols which were printed on them. Such *logical machines* allowed to formulate different types of syllogisms concerning any kind of symbolically representable content.

He used these logical syllogistic machines to demonstrate the existence of God, and more particularly the existence and the primacy of the Christian God; his logical machines are the first, even if primitive, material form of the idea of imitating some activities of the human mind; they attracted much interest, and were resumed by other logicians in the following centuries. Such machines were able to formulate mechanically the syllogisms, and their application to theology might have been acceptable within Christian theology as a means to offer a logical proof of the existence of God; unfortunately, however, it proved unacceptable to Muslim theologians, who after a short period of interest decided that it was probably better to abandon these logical machines, and Ramón Lull was stoned to death.

But the idea of the mechanization of the human mind, and more particularly of its logic, was not abandoned, and it survived in various stages till modern logic, the formulation of axiomatic systems, and the construction of intelligent machines. This course, up to the era of the first computers like the Eniac, manufactured among others by Arthur Burks, had many stages, and it is useful to remind some of them. The first was the combinatorial logic of Pierre de la Ramée, and afterwards Leibniz's logical studies and the famous machine devised by Pascal to perform mechanically some elementary arithmetical operations. If mathematical thought was typical for mankind, and if it was mechanically reproducible, then at least some part of human thought was also reproducible by a machine: in other words, a machine could be able to perform some human mental operations.

This concept of the reproducibility of human thought was further used in devising logical or mathematical machines like those of Babbage; since then the

possibility of reproducing human thought has made progresses, up to the manufacture of machines capable to perform not only mathematical operations, but also logical operations, in particular deductive operations similar to those of humans: the underlying idea has always been that of considering the mental processes as rational and logical processes grounded on rules and applicable to different contents.

Nevertheless, we must remember that, together with formal logic, in the Middle Age also other kinds of logics were formulated, where the contents of the propositions involved in logical processes were also relevant: just think, for instance, of Boethius' logical conceptions.

Within the formal view the first researchers on Artificial Intelligence (AI) held that it was possible to reproduce, or at least to simulate or emulate, human thought by a machine. Hence, their purpose was not to understand, by the use of machines, the functioning of the human brain, but to use machines to perform the same operations (even if not all) that the human mind is capable to perform. So, they drove to the man-automaton equivalence.

Once again, this conception was based on the idea that human thought was the result of elementary computations which were capable to generate very complex processes. This conception was little by little abandoned within AI, in favor of a more limited objective, the construction of machines capable to realize computations useful to perform specific tasks: therefore, little by little the term 'intelligent' was abandoned, and attention was devoted to the construction of expert systems useful to perform different tasks, like: visual recognition of objects; auditory recognition of sounds; recognition of texts; management and control of very complicated processes, like those of airplanes, of space ships, or space probes; the performance of activities like those of robots, drones, or similar apparatuses.

However, it is useful to remind that computationalism, which took the computations performed by machines as a model of the human mind, brought to a reductionist philosophy of mind; in its strongest form this conception was abandoned as a consequence of results obtained in the neurosciences, according to which the human mind is much more complex than mere computational processes based on few rules applied to basic elements. This point, as we shall see, was stressed by Agazzi since the sixties of the last century.

The first theorists of AI reasoned from the premise that the mind operates always and only on the basis of logical or formal rules; hence, for them simulating the human mind meant building technological apparatuses capable to perform some of its operations, and this entailed that "machines can think". Agazzi dwells upon this statement clearly distinguishing two aspects of the problem.

3 Evandro Agazzi's Approach

Agazzi believes that the claim that an artifact is capable to think like humans think is not in itself wrong, if considered as a guideline idea from which to move in trying to formulate simulations or emulations of human thought. But this thesis,

or if you like, the affirmative answer to the question 'can the machines think?', cannot be considered as a scientifically validated hypothesis, because this would require supporting it by adequate evidence, not only as to the results, but also as to the procedures needed to achieve such results. But until now no such evidence has been brought up in a conclusive and rigorous way, because even if machines can obtain results similar to those of humans, this does not mean, as Agazzi points out, that minds and machines have the same nature. On the contrary, the neurosciences have revealed that not only inside the black box of the human mind there are components and elementary constituents substantially different from those of machines, but there are also procedures which in most cases are not reducible to merely computational processes, understood in their standard conception: a finite set of formal rules applied to elementary components and relative to any kind of mental information, be it cognitive or not.

Agazzi underlines that often the suggestion of the equation man-automaton and the related statement 'machines can think' is due to the use of an anthropomorphic language to indicate operations performed by machines: talking of a machine like an electronic computer we use propositions like 'it learns', 'it remembers', 'it makes choices' etc., but the operations of this automaton are very different from those of a mind; nevertheless, the use of these terms belonging to our ordinary language induces one to imagine that the machine performs precisely the operations which are performed by a mind. According to Agazzi's this linguistic practice is apparently justified by the analogy with the talk concerning *transitive* operations of machines: *transitive* operations are defined on the ground of the results that any agent (be it a man or a machine) can achieve starting from an initial state of affairs, just as when it is said that a machine sews, another mows and another washes. But when we use expressions like 'answering', 'learning', 'remembering', 'feeling', 'seeing', etc., we are not talking of *transitive* operations, but of *immanent* operations: immanent operations are not defined through their "external" results but through the modifications they induce "within" the agent; we shall revert further on such distinction when we shall examine the thesis defended by Agazzi in his essay "Operazionalità e intenzionalità: l'anello mancante dell'intelligenza artificiale".

In connection with the behaviorist perspective, which focuses on the analogy between behaviors/results of a machine and those of the mind, Agazzi states as follows:

> The majority of reductivist positions inspired by artificial intelligence is based on the fact that the actual computers are able to perform a certain number of functions that man can perform thanks to the use of the intelligence, and on the persuasion that even those functions that have not been imitated up to now will be imitated with the progresses of technology. This is the position inspired by the famous apologue of Turing known as 'the imitation game', consisting in hypothesizing a machine able to give to a human interlocutor answers that don't allow her to understand whether such answers come from a machine or from another human interlocutor to which the same questions have been asked. The game is a clear representation of the stimulus-answer scheme of behaviorist psychology, and is subject to the critical remarks received by the latter concerning its adequacy as a methodological attitude to investigate cognitive activities. To these criticisms

must be added (in case the Turing method is adopted in a reductionist point of view) the mention of a further methodological incorrectness, which consists in holding that the identity of one or more functions justifies the claim of identity of nature between the entities displaying such functions. Such incorrectness can be noticed just on the pure logical level: identity of nature in fact is a condition *sufficient* but not *necessary* for the identity of function, therefore it is not possible to infer the first from the second. Very obvious examples prove this truth in a clear way: both an airplane and a bird can fly, but they have a very different nature; both an old mechanical computing machine and an electronic computer can add two numbers, but they have nearly nothing in common as far as their nature and structure is concerned. The very developments of artificial intelligence are a probatory confirmation of this fact: the progresses of machines in 'emulating' some intelligent operations of man have been obtained by giving up any 'simulation' of the cognitive activities of man: i.e., by using our advancements in electronic engineering more than those in cognitive psychology (Agazzi 1991, p. 5).

Besides this functionalistic or behaviorist perspective Agazzi takes into consideration another, that he names *structuralist*, characterizing it in the following way:

In order to overcome the difficulties of the behaviorist (or so to say 'functionalistic') formulation the imperative seems to be that of checking what really is inside the 'black box'. But even this proposal is not exempt from reductivist risks: indeed it depends on what one is looking for inside the 'black box', more particularly utilizing the tools of the artificial intelligence. Now it happens that, very often, those who intend to imitate not just the functions, but directly the structure of the human thinking activity, identify such activity with the human brain, of whose functioning the thought would be just a product. If one succeeded in building an 'artificial brain', that would therefore be able to think. At the basis of this perspective there is however an unfounded assumption: the identification of thought with an immaterial product of the brain (a product therefore not merely physiological, like the many products which accompany the metabolism of the cerebral activity). Even in this case a logical mistake can be noticed: exchanging a necessary condition for one which is *also* sufficient: it is fully plausible that there is no thought without brain (at least as long as we deal with human thought), but this is not to say that having a brain is sufficient to have a thought. Even the most spiritualistic philosophies, in fact, have never denied that the brain is an essential *condition* for the exercise of the thought, but this neither means that it is the *cause*, nor that it coincides with thought. But there is something more: strictly speaking the construction of an artificial brain will never consist in the exact reproduction of a natural brain, and all the progresses of the bionics can succeed to produce is an artifact which *works* in a way very similar to the natural brain. But then we really fall in the earlier situation: even the 'structuralist' path in the end is a 'functionalist' path aimed to a lower objective, i.e., to 'simulating' the neurophysiological functions, rather than the cognitive ones (ibid., pp. 5–6).

However, there are some relevant aspects of the functionalist thesis which are focused on by Agazzi in the following way:

the Western philosophical tradition not seldom has held the immateriality of thought (hence also the spirituality and the immortality of soul) because humans are able to perform activities which do not imply any contribution of materiality (typically, knowledge of the universal), and from this it was possible to infer the existence in humans of an immaterial principle necessary to perform such activities (ibid., p.7).

This conception is grounded on the metaphysical principle *operari sequitur esse*. From this principle Agazzi draws a question crucial to the topic of the difference between humans and machines: "if it is incorrect to infer from the identity of

the functions of two entities the identity of their nature, is it not equally incorrect to infer, from the fact that a certain activity leaves out matter, that also the entity capable to perform such an activity can exist without matter?" (ibid.). And he answers this crucial question as follows:

> Notwithstanding the appearances the questions are not on the same level. '*Operari sequitur esse*' means in fact that any being must possess in its nature all the conditions necessary and sufficient to perform its functions, so that if for a certain function a particular condition (in this case that of materiality) is not necessary (besides not being sufficient), the related necessary condition will be of different nature (immaterial in the specific case). On the other hand, if it is granted that the first condition is not necessary, and it is evident that the function takes place, it is possible to infer that the second condition, beside being necessary, is also sufficient. Naturally the application of this principle brings to the acknowledgment of two 'distinct' conditions for the exercise of a certain function: that they are also ontologically 'separated' is a much more complicated problem which is of no interest here (although rather important for the statement of the separation of the soul from the body) (Agazzi 1991, pp. 7–8).

With reference to the simulative aspect of the functionalist thesis in the discussion on the artificial intelligence Agazzi points out that

> if really the simulation of *all* the most intellectually elevated human activities (i.e., those activities which usually are considered not 'depending from matter') would perfectly succeed without residuals by means of material electronic calculators, one couldn't any more claim that the materiality is *not sufficient* to perform such functions; so, it would no longer be necessary to find conditions different from matter for the realization of those functions in humans. Certainly it would still be true that the identity of functions does not logically imply identity of nature, but in order to affirm the difference of nature one should produce *other* arguments, different from those classically considered the more conclusive (ibid., p. 8).

Agazzi, therefore, has pointed out that the two theses (behaviorist/functionalist and structuralist) fail to show that machines think in a way similar to the mind, not only with reference to behaviors or functions, but also to the nature of machines and of the mind.

Agazzi' arguments can also be referred to the strictly logical operations performed by a machine. In this case his analysis appears to be particularly convincing and articulated. There are, Agazzi points out, two essential aspects: on the one side the involvement of the notion of 'truth', and on the other side "the request to take into account all the *possible* interpretations of our assertions" (Agazzi 1967a, p. 19). Neither aspect is present in the same way in humans and in machines. This is how Agazzi argues: "We substantially say that machines succeed in imitating our deductive reasoning since we succeed in reproducing in them that whole of formal schemes and rules of logically thinking that we have considered as typical of our deductive way of *operating*" (ibid., p. 224). But Agazzi points out that

> it must not be forgotten that these formal systems of rules are only an artificial surrogate of what we mean by the terms 'demonstration' and 'logical consequence'. The link of logical consequence has an intuitive nature, and it is possible to say that a proposition P is the logical consequence of a set propositions S when it is not possible to suppose in any case that the S are true and that P is false; in other words, as we say more precisely, when *any* 'interpretation' which verifies the S also verifies P (ibid.).

To avoid the difficulties entailed by the reference to the infinite totality of the interpretations contained in the intuitive notion of logical consequence, this has been replaced by that of *deduction*, that is, "of an *operating* on the ground of rules in which the concept of truth does not appear but which is practically checkable step by step" (ibid., p. 19). For this reason, Agazzi remarks, it is plausible to admit that, as it is possible to manufacture washing or writing machines, it is also possible to manufacture machines to perform material operations on symbols and therefore also logical operations like the deductive ones. But also in this case we only have a surrogate of what the mind does, since also in the case of deduction we always have to do with a living being, with the structure and the contents of his mind, and this makes the nature of these operations different from those of an automaton.

To point out the difference between the human way of operating and that of a machine Agazzi proposes a clear example related to vision. Let's think of the difference between the perceptual image of an object and its image on a photographic plate: in the latter case we don't use the term 'sees', but the expressions 'it fixes' or 'it records' the image. For which reason is this terminological difference made? It consists in the fact that in the case of the visual perception of humans, unlike the recording of an image by a camera, there is something *more* which is indicated by Agazzi by the term *intentionality*; by such term he alludes to the fact that in the cognitive or simply perceptive activity of the living being, and particularly of humans, there occurs a kind of participation or of identification of the subject with the objects "which even remaining themselves, in some way become part of the subject" (ibid., p. 16). In other words, in general, reference is made to the subjectively lived character of perception.

In fact, this argument by Agazzi, in the current state of the neurophysiological researches, allows to claim that the perception of an object is not only the result of the ocular apparatus and various cortical and not cortical areas of the brain used to the reception and to the elaboration of the photons, such as the geniculated nucleus brain or the primary and secondary visual cortex, but also of the involvement of the cognitive and associative cortical areas located in the temporal and frontal regions of the brain, which add information to that received by the stimulus, thus producing the perception.

That something *more* differentiating humans from cameras is constituted by intentionality and consciousness, and this means that in order to have a visual perception we must add to the information coming from the stimuli some information already present in the mind, which allows both to cognitively recognize the object, and to assign the image different meanings, like, for instance, 'it is beautiful', 'I like it, 'it reminds me of a past experience', 'it makes me feel a certain mood'. For this reason we say that an object is *seen* by us (in this wide and conscious sense), while it is simply 'recorded' by a camera.

In even wider terms, according to Agazzi's conception, it is not the eyes, but the whole subject which 'sees', and this is naturally and substantially different not only from what happens in a camera, but also in an artificial apparatus capable of recognizing objects, as in the case of *pattern recognition*; in this sense,

in machines there is no subject operating and endowed with intentionality and consciousness. So, the seeing of an object by a human is not a process similar to the recording by an artifact; in this case even the results are different, since the perception of an object involves information possessed by the percipient subject, and such information is not present in the processes of recognition of an object by a machine.

This topic, which is central in Agazzi's analysis of the artificial intelligence, has been more deeply tackled in the essay *Operazionalità e intenzionalità*: *l'anello mancante dell'intelligenza artificiale* (1991), of which I shall report some passages which allow a more complete understanding of Agazzi's conception of artificial intelligence. First of all, Agazzi introduces the already mentioned distinction between transitive and immanent operations. With reference to the former Agazzi underlines that

> The majority and maybe the totality of man's transitive activities can be expressed by indicating a specific *operation* which has been generated. We have said 'has been generated' and not 'in which consists', because the operation, as we want to define it now, is characterized through two static moments and is independent of the modality of transition from the one to the other. In general, in fact, it is possible to define an operation *O* as *any process whatever* which, applied to an initial state *I*, leads to a final state *F* in which the transition from *I* to *F* doesn't make any difference. For instance 'sewing' is an operation defined by the passage from an initial state in which the two pieces of tissue are separated to a final state in which the pieces are strictly connected with a thread. Just because the operations concern states of the world, they result univocally defined with reference to such states, therefore they remain *the same* whether they are performed by humans or by animals or by machines. On the other hand, in this case they *say nothing* either about the process of transition from *I* to *F*, or about the agents which have performed the process. This is the fundamental reason for which humans, even if having continuously realized machines capable to perform operations identical with those of their transitive activities, never considered this fact as a threat to their human identity, nor they feared to be assimilated to machines, even when these were more efficient than humans in performing certain operations (Agazzi 1991, pp. 8–9).

According to Agazzi, we run into a very different situation when "the machines have started to do some activity which is considered immanent (like those connected to the exercise of thought and of intelligence)" (ibid., p. 9). However, Agazzi points out that in this case it is necessary to carefully consider whether "what is imitated by the machine is really the immanent aspect of such activities" (ibid.). In which way can we answer this question? According to Agazzi we must answer it in the negative, because "even for some of their immanent activities humans have been able to devise some *operations*, and it is only the latter that machines can imitate (or more exactly, can carry out)" (ibid.). An immanent activity is always a change, but it does not concern the state of the external world, but that of the subject which performs such activity. In some cases however, continues Agazzi,

> the subject may reflect on his internal states and associate them with material objects *representing* them as signs, then establishing material *operations* on the signs, so that when from the internal state I he passes to the internal state F, the operation O applied to the material sign which represents I leads to the material sign which represents F. (ibid.).

This condition occurs since humans and their minds operate on symbols, for which they can always produce "the translation into material signs of their internal states and the reconstruction of internal states starting from material signs, in the extremely varied range of the 'languages' in which they express themselves" (ibid., p. 9–10). However this complex activity of the human mind according to Agazzi may be understood only by stressing that it is always guided and controlled by intentionality, which "allows to pass from the internal state 'intentionated' to the material sign which *represents* it, and it is again intentionality which allows to *interpret* some material signs as representing certain internal states" (Agazzi ibid., p. 6).

At this stage of the argumentation we must face a crucial question: is it possible that these mental activities are carried out without the constant control of intentionality? In other words, for Agazzi, we should ask "whether the chain of internal transitions which bring me from a state I to a state F passing through a number of intermediate states (also of intentional nature) can be *represented* by a material *operation* (not intentional as such) taking from the symbol of I to the symbol of F" (ibid., p.10). The answer to this question (which is in general negative) leads to a non-functionalist position and to the rejection of the equivalence between humans and robots. In this sense Agazzi points out that

> the answer to this question is positive only in few cases, that is, in those in which the transition from I to F complies with some well specifiable requirements and in which, in addition, it would be possible to devise some symbolic representations of the operational manipulations on the symbols which, when they are re-interpreted, comply with these requirements. The more relevant example in this sense is represented by formal logic, in which I and F can be respectively considered as the premise and the conclusion of a reasoning in which the requirement of the transition is that the truth of the premise passes to the conclusion. I and F are contents of thought … but they can be appropriately translated in linguistic expressions, that is, in material signs which express them, and formal logic institutes operations which from the expression translating I lead to the expression translating F while complying with the condition of the persistence of the truth. Equally well known is the example of mathematical operations in which one materially operates on symbols according to rules which have been instituted on the ground of an 'intentional' reflection, but which are also applied in a purely material way. From what has been said it is clear that those operations which, following the double application of intentionality, can be performed to *represent* immanent activities, have a transitive nature (they manipulate the external world) and, *qua operations*, can be performed in any way whatsoever by any agent as long as they correctly take from I to F. It is therefore completely obvious that they can also be performed by machines, perhaps even more efficiently than by humans, and it is also clear that their carrying out does not require *intelligence*, just because this, thanks to the intentionality that characterizes it, stays upstream and downstream the carrying out, and it consists in devising the correct manipulations on the material signs, in the appropriate symbolization of I by means of material signs accessible to those manipulations, and in the interpretation of the result of the operation as a certain F (ibid., pp.10–11).

This shows that only a limited and specific number of mental activities, owing to their formal nature, can be simulated by an automaton; therefore Agazzi puts a strong constraint on the concept of simulation of human mental activities by an artifact. There are no difficulties of principle for transitive activities, but there

are limitations as to the possibility of current electronic computers to represent immanent activities. In fact such computers "limit themselves to perform material operations which can be used by humans to replace some part or applications of human immanent activities only thanks to intentionality" (ibid., p. 11). A further unavoidable question arises at this point:

> even admitting that intentionality is the characteristic which differentiates the operations of humans from those of a robot, can we consider it as a qualitative difference justifying the application of the principle *operari sequitur esse*? In other words, can we attribute it to a dimension which goes beyond materiality? (ibid.).

From this follows the crucial question for the entire argumentation of Agazzi about artificial intelligence: "is it possible to realize intentionality within a robot?" (ibid.). Agazzi replies negatively, for the following reason: "intentionality is not a physical state, is not the simple 'presence' of something, but a special *way* of such presence" (ibid.). To clarify this remark Agazzi considers the image of a house; "the image of a house is present *physically* in a film and in the retina of an eye, but it is present *intentionally* to the perceptive faculty of a man or of an animal" (ibid.). In this sense, therefore, he points out that "intentionality needs physical conditions to reveal itself, but it does not coincides with them" (ibid.), and this holds both for perception and for the other activities which are considered intelligent. From these remarks Agazzi concludes that it is possible to realize

> for instance by means of sensors and analogic computational procedures governed by specific negative feed-back, according to the ideas of McKay, a 'simulation' of perceptive activities more advanced than the very unsatisfactory attempts realized with digital methods; but it will always be a 'physics' of the perception, even if very interesting. A physics which will help us to understand better *through* which processes intentional cognition adapts itself to the external world and absorbs it in itself, but will not be able to physically explain away intentionality" (ibid.).

For this reason once again we are inside a behaviorist/functionalist point of view of second level "in which … the behavior becomes internal and concerns the evolution of the information and of the programs" (ibid., p. 12), therefore, it still belongs to the physical level and does not belong to the cognitive perspective proper.

The analysis becomes more complex if from perception we pass to the strictly cognitive activities characteristic of thought, where we also find "the capability to intentionate the abstract, the possible, the true and the false, the ought to be, the necessary, the mandatory, the prohibited and all the ranges of the *intentiones secundae* which are at the ground of creativity and of free will" (ibid.). Is it possible that these *intentiones secundae* are realized in an automaton? To this question Agazzi replies in an articulated affirmative and negative way:

> in a certain measure yes: i.e., in the measure in which through acts of *self-consciousness* we *objectify* to ourselves some of those abstract entities, we analyze them and we operate a formalization that afterward becomes symbolizable and translatable into programs for the machine. But the intentionality through which we carry out such objectification and such translation remains behind the shoulders of the program, it does not appear in it, it does not enter into the machine, remaining as unanalyzable and inexpressible than the

intentionality which is at the basis of the perceptive act and which confuses itself with the content of perception (ibid.).

In this regard Agazzi claims that it is this intentionality which allows "the demonstration of all the semantic meta-theorems of mathematical logic (completeness, incompleteness, categoricity and non-categoricity, internal limitations of the formalisms such as Gödel's theorem, etc.)" (ibid.). Hence, in this respect, the answer is negative; indeed, Agazzi holds that:

> it is not difficult to show that reflections and demonstrations of this kind are not replicable in an automaton, since they presuppose the availability of a *semantic meta-language* which the automaton lacks just because is based on intentionality, for without a *starting* intentionality it is not possible to construct the programs which enable a robot to compute algorithms … and without an *arrival* intentionality it is not possible *to evaluate* the obtained results (ibid.).

The arguments grounded on intentionality, however, do not refer only to the immanent activities of humans, since intentionality is actually present in

> *any* human activity, making it specifically human. Even in the transitive activities it is always humans who do something by *using* material instruments which can initially be their two hands, then simple tools like a spade, and finally very complicated machines. It is always humans who *intentionate* the result they meant to obtain, who *plan* the operations to obtain it, who plan and manufacture any machines needed to carry out such operations, who finally *evaluate* whether the result is consistent with what they intended to obtain. The same happens in the case of mental activities: humans *intentionate* what they want to achieve (for instance, the sum of two quantities), they *formalize* it conceptually and *symbolize* it materially (for example, by taking pebbles as proxies for each element of the two quantities, and by the contrivance of an algorithm consisting in throwing the pebbles, one after the other, into one and the same container); finally, once the *program* has been completed, they *interpret* the result (by counting the pebbles which are in the container). If instead of using pebbles they use numerals written on a sheet of paper, or the states of magnetization of electronic components of a computer, and if instead of accumulating the pebbles they use algorithms to manipulate the numerals according to fixed rules or computer programs, nothing relevant changes in this proceeding. Similarly, nothing changes if, in order to solve a much more complicated problem, they must materially memorize some intermediate results, or, during the execution of new algorithms, they must use appropriate sub-algorithms at the appropriate moments, so that at the completion of the process they can read the solution of their problem by *interpreting* the terminal configuration of the material signs. It is always humans which, intentionally using the machines, carry out their operations (ibid., pp. 12–13).

From these arguments follows Agazzi's fundamental thesis according to which

> it is never possible to speak of any human activity without appealing to intentionality, while there is never need of intentionality to describe the operations of a machine, no matter how complicated and improved it is. Not even the computers which perform the fascinating technological programs of the A.I. are an exception to this rule (ibid., p. 13).

Concerning this condition, and the great diversity between man and the machine, Agazzi uses an expression taken from biological evolution: "the *missing link* … between natural and artificial intelligence … undoubtedly seems to be intentionality" (ibid.). Therefore, this complex argumentation of Agazzi not only confutes, as we have seen, all forms of physicalistic, computational and functionalistic

reductionism, but it "shows the inadequacy also of those forms of emergentism which consider the most elevated levels of reality only as complexification of those of the lower levels, losing the real nature of the new and of the different" (ibid.).

4 Conclusion

In conclusion, it is useful to point out that Agazzi does not assume an a priori point of view against artificial intelligence, but he limits its feasibility to the cases of transitive or even immanent operations based on formal rules; nevertheless, also in this case formal operations in humans and in machines are substantially different, since they are embodied into largely different systems. Mechanical systems differ from the human mind, in which beside intentionality and consciousness, we find very complex processes, where even simple formal operations involve non logical, or non strictly cognitive aspects. In fact, such operations are expressions of a living being, of a complex cerebral system and of an even more complex mind, in which operations, including formal ones, are always interwoven with meanings and guided by specific goals. Actually, it is just the meaningfulness and teleologicity of mental processes which differentiate them from the processes of an artificial system which does neither elaborate nor generate meanings, nor intentionally proposes goals to itself; the distinctive characters of the operating of a mind are just the intentional goals and the processes which create meanings and interpret the stimuli coming from the world; this does not mean that is impossible to build always more complicated machines which allow to obtain results similar to those of human cognitive activities.

Such machines could also be useful to understand some aspect of the human mind, but one cannot claim that they are intelligent like the human mind, nor that they are a reproduction of the human mind; in fact, in order to be such they should possess a biological mind like that of humans, be inside of a body like that of humans, and finally, have a life like that of humans, with all its existential, psychological, affective, emotional, relational and significant aspects: these last, the significant ones, are characteristic of subjects which possess a mind and intentionality.

However, finally, what is really interesting is not so much emulating the human mind, but building machines able to perform various very complicated tasks that the human mind cannot perform. In this way one can construct artificial minds different from the biological ones, and forms of intelligence different from the biological ones: they will pursue goals that, in Agazzi's terminology, are certainly intentional goals of humans, and this is not transferring human intentionality to the machines, or placing in them an intentionality autonomous from that of humans. Whichever is the plausibility of such scenarios, there is also one more possibility: that the robots can autonomously develop, self-generate, self-reproduce and form their own artificial intentionality. But the analysis of the latter scenario is beyond the analysis of Agazzi's positions, although perhaps he would not exclude it.

References

Agazzi, Evandro. 1967a. Alcune osservazioni sul problema dell'intelligenza artificiale, *Rivista di Filosofia Neoscolastica,* 59: 1-34.

Agazzi, Evandro. 1967b. Simulazione del pensiero e intelligenza artificiale. In *L'uomo e la macchina* (Atti del XXI Congresso Nazionale di Filosofia). Torino: Edizioni di Filosofia, II: 45-48 and III: 109-110.

Agazzi, Evandro. 1978. Alcune osservazioni sul problema dell'intelligenza artificiale. In *Cibernetica e teoria dell'informazione,* ed. Paolo Aldo Rossi, 199-244. Brescia: Editrice La Scuola (reprint of Agazzi 1967°).

Agazzi, Evandro. 1981. Intentionality and Artificial Intelligence. *Epistemologia* IV, Special Issue ("The Mind-Body Problem"): 195-227.

Agazzi, Evandro. 1987. Intenzionalità e intelligenza artificiale. In *Informatica e pastorale*, ed. Pietro Prini, 87-105. Brescia: Morcelliana (translation of Agazzi 1981).

Agazzi, Evandro. 1991. Operazionalità e intenzionalità: l'anello mancante dell'intelligenza artificiale, in *Intelligenza naturale e intelligenza artificiale*, ed. Salvino Biolo, 1-13. Genova: Marietti.

Agazzi, Evandro. 2010. Operazionalità e intenzionalità: l'anello mancante dell'intelligenza artificiale. In *Natura umana, natura artificiale,* ed. Maria Cristina Amoretti, 63-77. Milano: Angeli (reprint of Agazzi, 1991).