

Artificial Intelligence and Moral intelligence

Laura Pana

*Associate Professor, PH.D.
University "Politehnica" of Bucharest
313, Splaiul Independentei, Bucharest
lcpan20032000@yahoo.com*

Abstract: We discuss the thesis that the implementation of a moral code in the behaviour of artificial intelligent systems needs a specific form of human and artificial intelligence, not just an abstract intelligence. We present intelligence as a system with an internal structure and the structural levels of the moral system, as well as certain characteristics of artificial intelligent agents which can/must be treated as 1- individual entities (with a complex, specialized, autonomous or self-determined, even unpredictable conduct), 2- entities endowed with diverse or even multiple intelligence forms, like moral intelligence, 3- open and, even, free-conduct performing systems (with specific, flexible and heuristic mechanisms and procedures of decision), 4 – systems which are open to education, not just to instruction, 5- entities with "lifegraphy", not just "stategraphy", 6- equipped not just with automatisms but with beliefs (cognitive and affective complexes), 7- capable even of reflection ("moral life" is a form of spiritual, not just of conscious activity), 8 – elements/members of some real (corporal or virtual) community, 9 – cultural beings: free conduct gives cultural value to the action of a "natural" or artificial being. Implementation of such characteristics does not necessarily suppose efforts to design, construct and educate machines like

human beings. The human moral code is irremediably imperfect: it is a morality of preference, of accountability (not of responsibility) and a morality of non-liberty, which cannot be remedied by the invention of ethical systems, by the circulation of ideal values and by ethical (even computing) education. But such an imperfect morality needs perfect instruments for its implementation: applications of special logic fields; efficient psychological (theoretical and technical) attainments to endow the machine not just with intelligence, but with conscience and even spirit; comprehensive technical means for supplementing the objective decision with a subjective one. Machine ethics can/will be of the highest quality because it will be derived from the sciences, modelled by techniques and accomplished by technologies. If our theoretical hypothesis about a specific moral intelligence, necessary for the implementation of an artificial moral conduct, is correct, then some theoretical and technical issues appear, but the following working hypotheses are possible: structural, functional and behavioural. The future of human and/or artificial morality is to be anticipated

Keywords: Text

Acknowledgement: Text

1 An Artificial Ethics for the Artificial World. Introduction

The setting up of an artificial technical environment and its actual transition to an intellectual technical environment, by the transformation of the information machine into the intelligent machine, generates a new practical and theoretical field of research, that of machine ethics.

The ethical impact of machine action was until now neglected and the responsibility for its consequences was devolved upon its user, designer or owner. Human and artificial beings are acting and interacting now in new manners and the machines cumulate multiple and meaningful functions related to man and society. Artificial agents are created not only to assist but to replace humans in such processes as fabrication, business, services, communication, research, education and entertainment. Their conduct thus receives a moral significance.

This increasingly artificial world, generated by the man-machine interaction, produces not just a growing complexity of the machine, but of humanity itself and of its moral values. The human species evolves in all its dimensions: biotical, psychical, social and cultural. Now it evolves towards artificiality.

As a biotical (natural) being, man evolves towards artificiality because in our process of aging and decreasing, we receive more and more artificial components and we become more and more robot-like, by means of combinations between the biotic and the technical tissues, by the invention of artificial muscles and artificial blood and by using the artificial nervous system or creating mixed intelligent systems.

On the other hand, artificial intelligent agents receive super-sensorial properties (such as infra-red vision and ultra-sound detection) and an efficient form of abstract intelligence. As computer scientists are anticipating, Artificial Intelligent Agents will also develop by simulated “natural” evolution and by processes inspired from garden tillage or cake baking (Hillis 2001, p.171 and 179-180). According to other authors (Gregory 2000, p.59) artificial intelligent systems will develop by learning from experience and by the assimilation of culture following the model of a child’s education.

Society evolves too, mainly under the influence of the advances in information and management technologies: the cultivation of instrumental and intellectual automatisms transforms humans and human institutions in more or less efficient automata. Humans, machines and organizations are now less and less different in their specific activities, because of the shift in emphasis of the different types of behaviour: machines are gaining in intelligent conduct and humans are reduced to an automatic one, while we know that humans have always been characterized by the highest structural levels of conduct: intellectual and spiritual.

The human cultural dimension is also transformed and a technical man is born. All types of values are renewed, because of the emergence of new human needs, often satisfied by virtual relations and virtual means, in a virtual environment, which is a technical and mainly intellectual artificial environment.

This artificial world needs an artificial ethics. Starting from the fact that human and artificial agents are now going to explore and populate a virtual global intellectual environment, we anticipate that this environment will be governed by a new ethics – a manifold, multi-layered, partially virtual and artificial ethics. Some parts and aspects of this new ethics are already born and developed.

1.1 Four New Ethics for Human and Artificial Agents

Ethics of Computing is useful for all who use a computer in the new technical intellectual environment. It is not a professional ethics, but one destined for computer and net workers with diverse professions, who process and transmit information in this way. The following issues are considered characteristic for this ethical field: software property protection, ensuring user identity and intimacy and even the sharing and preservation of a netiquette.

Computational Ethics uses computers in the ethical field of philosophy for theoretical and practical moral problem solving; computer-based teaching and learning means and methods are adopted and developed. This ethical domain uses results of IT research and advantages of the telematic, distributed, flexible and multimedia learning, assisted by knowledge-based AI techniques. Accredited moral theories are studied by computational methods which allow also for foundation-of-decision analysis in difficult moral problems. Moral effects of meaningful social decisions are anticipated by evolving modelling and simulation.

Danielson (1998, p. 292) revealed that important parts of morality have always been artificial; using the computer in this field, we just extend the artificial characteristic of the morality.

Machine Ethics concerns the computer itself; the intelligent machine induces changes in the world, like humans do. All human activities have moral significance. A machine with similar possibilities needs moral functions. Machine ethics is a new domain of research, where generalists from philosophy and specialists in AI collaborate; the result will be an Artificial Ethics which will constitute a part of Artificial Philosophy. Machine Ethics tends, actually, to become a field of the research in Artificial Intelligence, but it cannot be

conceived and constructed without a strong philosophical (ontological, axiological, pragmatic and ethical) foundation. Actual philosophy can deal with specific problems raised by artificial ethics because it evolves towards an Artificial Philosophy, even by multiple ways (Pana,2005 b). The term “Artificial Philosophy” was first used by Fr. Laruelle and designated the science of thinking developed by mathematical and technical methods (Laruelle, p. 235). Earlier still, P. de Latil had studied “synthetic” or artificial thinking.

Global Information Ethics was been generated by the globalization of knowledge, communication and work. Global Information Ethics can be conceived, simply, as the ensemble of the above mentioned new domains of ethics, but more accurately, as a super-structured new level of ethics, as a result of their synthesis under the circumstances of the informatization and intellectualization of all human activities. Important aspects of Global Information Ethics were emphasized by Bynum (1998), who showed that in this domain of ethics moral values and conducts become objects of debate and adjustment, beyond geographical, social, political and even cultural differences.

2 Characteristics of Artificial Intelligent Agents (AIA)

Human Ethics cannot be an efficient model for Machine Ethics because of the nature of human morality and because of a fundamental incongruence between human and machine ethics derived, mainly, from some characteristics of artificial intelligent agents. The theoretical and practical difficulties in the implementation of a moral code will be analyzed in Section 4.

Because of their level of complexity, their specific functionality and mainly the necessity of a degree of freedom, artificial moral agents must be conceived as individual entities also endowed with other necessary qualities.

Artificial moral agents can/must be treated as 1 - individual entities (complex, specialized, autonomous or self-determined, even unpredictable ones), 2 - open and even free-conduct performing systems (with specific, flexible and heuristic mechanisms and procedures of decision), 3 - cultural beings: free conduct gives cultural value to the action of a human (natural) or artificial being, 4 – systems which are open to education not just to instruction, 5 - entities with “lifegraphy”, not just “stategraphy”, 6 - endowed with diverse or even multiple intelligence forms, like moral intelligence, 7 - equipped not just with automatisms and intelligence, but with beliefs (cognitive and affective complexes), 8 - capable even of reflection (moral life is a form of spiritual, not just of conscious activity), 9 - components/members of some real (corporal or virtual) community.

Machine Ethics has to solve a difficult problem, that of the joining of a theoretical requirement to a practical imperative concerning moral freedom. In human conduct, morality supposes the practice of moral freedom: only with this assumption and the application of moral norms can human conduct have a cultural value. Also, in the case of the machine, the freedom–responsibility correlation must function both in a broad and a restricted sense:

- a) responsibility is the condition of freedom;
- b) freedom is the cause of responsibility.

More explicitly, in (b), the degree of responsibility is not just dependent on, but even determined by kinds and levels of freedom.

Thus, Machine Ethics must consider the need to ensure the freedom of choice in the application of moral norms in determined action domains and situations. Man must assume the risks which result from allowing freedom to machines, and he must also accept the possibility that machines increase their degrees of freedom: its conduct depends not just on humans, but on its own decision and even on the decisions of other machines.

Implementation of the above mentioned characteristics does not necessarily suppose efforts to design, construct and educate a machine as a quasi-human being. On the other hand, even human conduct is perfectible just by the construction and application of a philosophically founded and scientifically deduced ethics. This will be an artificial ethics, which may be applied by humans and machines, who can meet midway between the natural and the artificial, in a common, better ethics.

3 Intelligence as a System with Internal Structure

The ethical conduct of Artificial Intelligent Agents (AIA), as well as the human moral conduct, needs, as we will see, many and diverse “psychical” qualities, not just intelligence, for the achievement of any ethical value. Intelligence itself is needed in its specific forms in different domains of human or artificial activity, and can be cultivated and developed starting with its internal analysis.

Human intelligence is diversified into human activity. Our hypothesis regarding the existence of moral intelligence and functionality is allowed by the application of systemic methodology in psychology and is sustained by an integrative philosophical vision of the forms of culture.

Research concerning the structure of human intelligence has evolved in two directions, each of which leaves an incomplete map of its representation. In an inductive way, mathematical, linguistic, descriptive, interpretative and theoretical forms of intelligence have been studied, but not scientific intelligence. The literary, musical and plastic intelligence have been inventoried but the artistic one has not. In a deductive way, a general intellectual functional availability (general intelligence) has been identified.

Technical intelligence has been analyzed as a form of practical intelligence. Moral and political intelligence are also predominantly action-oriented and strongly controlled by norms, but differentiated by the specific values pursued and by the kind of means utilized. Moral intelligence, however, cannot be integrated without difficulties in the group of practical actions.

We have, thus, a general intelligence which ensures the specific, intellectual level of any human conduct, then particular forms of intelligence which include abstract and practical types of intelligence and finally, specific forms of intelligence, generated by distinct domains of action, oriented by specific values and developed by educational technologies and by experience in adequate environments.

The actual stage and rate of progress in AI studies is generated by their orientation to develop mainly an abstract form of intelligence. A complex form of human and artificial intelligence is needed for an efficient practice or a successful implementation of a moral code.

4 Implementation of a Moral Code for Human and Artificial Agents Conceptual and Technical Difficulties

Computer scientists and AI researchers are not the only ones confronted by difficult and complex problems in the areas of conceptualization and invention. Philosophers and representatives of the humanities are concerned with a necessary re-foundation and a structural reconstruction of their research domains, according to the perspectives of the present informational environment. Examined in accordance with the present exigencies of scientific and technical knowledge, many accredited explicative and interpretative models of the social sciences and reflection require serious reconsideration.

In practice, human morality is a morality of preference, constraint and irresponsibility; as moral theory, human ethics presents a set of internal contradictions. Neither Human Morality nor Human Ethics can serve as a model for Machine Ethics.

It is now time to conceive and construct a scientific and technical basis for a common, entirely invented, and in this sense artificial ethics, one that might be implemented by both humans and machines.

Specialists in computer science and dedicated scholars in AI research are trying to “implement” a moral code:

1. without giving attention to the real complexity of moral culture, the spiritual level of the moral system and the real environment of AIS activities;
2. forcing ahead the development of an abstract artificial intelligence;
3. forgetting that
 - intelligence and the intelligent environment are the most efficient answers to complexity; see the consideration of these problems in Pană (2003, 13-22);

- efficiency supposes specificity (adequacy of agents, motivations, objectives, conditions, means, strategies, evaluations, ends as well as their dynamic connection in a cybernetic system);
- 4. eluding the specificity of the form of intelligence which can render efficient artificial conduct in an ethical context.

It is necessary then to achieve:

- a comprehensive analysis of the moral system, with its actual characteristics, with
- a study of Moral Intelligence as an answer to the problem of complexity and specificity of moral conduct and culture.

4.1 The structure of the moral system

Moral systems include a hierarchy of structural levels and suppose, for living, understanding and renewing them, a large diversity of more or less well-understood forms of intelligence.

Levels of moral conduct Forms of intelligence involved

I. Moral relations and activities Practical Intelligence, Concrete Intelligence

(Moral practice or morality) Imitative Intelligence

II. Moral community Inter-personal and Communicational Intelligence

III. Moral conscience Intra-personal, Emotional, Evaluative Intelligence

IV. Moral science

(Scientific ethics as: bioethics, Logical, Mathematical, Technical Intelligence
techno-ethics, machine ethics)

V. Moral philosophy

(Ethics and meta-ethics) Abstract and Theoretical Intelligence

VI. Moral spirituality

(Moral values understanding, Descriptive Intelligence

learning, Crystallized Intelligence

inventing, Fluid and Creative Intelligence

applying) Interpretative Intelligence.

As we can see, the moral system is integral to social life. To live in a moral system we must use, in diverse combinations, different forms of intelligence. Philosophers and scientists who search for conceptual and technical ways to implement a moral code for intelligent artificial agents must acknowledge the complexity of moral life.

4.1.1. Moral Consciousness

Moral conscience, especially, is a complex moral state and process. Moral conscience is highly structured, but sequentially formed and unequally evolved.

The structure of moral conscience includes, as constitutive levels and components, a set of habits, emotions and feelings, opinions and beliefs as well as specific reflections, which represent the source of moral philosophy and the start of the spiritual level of moral life.

As the elementary level of moral conduct, moral habits are induced from the outside, they are imposed by the system of negative and positive sanctions; the life-long moral behaviour is therefore enrolled in an obligation system which calls forth negative emotions, the entire moral life being thus subjectively experienced as living-in-constraint.

Moral options and opinions are, in their turn, internalized generally not from personal experiences, but by “apprenticeship”, from the culture of past generations, and are very stable or even inflexible, because their content, evolution and finality are rarely perceived.

Beliefs are cognitive, affective and evaluative complexes, which can be structured, in any domains of culture, at diverse levels, comprising the proportion of these components. Moral beliefs are also relatively stable, being held together by logic and strengthened by affects. But inconsistent beliefs can coexist, their successive or alternative application being possible. These constituents of moral conscience have been

proved easily transposable within diverse domains and situations (which can be an advantage), but they are also often too malleable, and other times just declarative or even false.

Moral conscience, as an aspect of human activity, manifested both in low and high level conduct (from elementary habits to spiritual life), implies not just cognition and affectivity, but free will too. Conscience means, mainly, auto-determined conduct, the capacity to avoid command and control, and dexterity to ensure communication and cooperation, as well as the ability to establish new goals, means and values through action.

Conscience represents, then, an important resource of creativity; it is also a very difficult objective to accomplish for the artificial mind-makers. But many well-known psychologists consider that the unconscious mind is the permanent and infinite source of creativity. A long and exotic inventory of this structural level of our mentality was made, for example by Ey (1982 Part 4, *The Unconscious*). Can we conceive an artificial mind characterized not just by consciousness, but one also gifted with unconsciousness?

4.1.2. . Moral Cognition

Recent research in AI, which started from psychological, psycho-analytical and even psychopathological studies (Arieti & Bemporad 1978), allowed a better understanding of the representation of social structures and permitted some successes in the recognition of human affective states, as in their description, classification and analysis. The localization of cerebral areas responsible for social cognition and the description of the cognitive structures revealing the specifics of the cognition of social space, and of diverse other kinds of social relations have been attempted.

But even authors such as Sloman (1990), who have developed notions of affectivity as elements of organization-and-control within the agent, recognize that emotion is not a special sub-system of the mind, but one of its penetrating features. Delving into the initially-mentioned ideas, the researcher finds that diverse emotions are associated with different types of motivational control, and develops strategies for comparison and selection of motives, as decision procedures for autonomous systems with limited time and effort resources but with multiple or even opposite emotions and motives (Sloman 1990, pp.237-238). Decisional procedures are based, in his vision, on moral criteria.

A hierarchy of structural and dynamic prototypes, appropriated by language, has been advanced as a model of social cognition. It overtakes the cognition of natural or social laws or the usage of recognised rules, and permits the advance of knowledge in conditions of complexity, ambiguity or incompleteness (Churchland 1998, p.153).

Specialists in computer science have also studied, in addition to problems such as recognition and reproduction of forms, themes concerning the cognition of cognition as the non-sentential, diagrammatic representation of knowledge. Thus, technical knowledge, based on cognitive psychology and other cognitive sciences, can solve problems of philosophical psychology or of epistemics.

Moral cognition remains, at first sight, an aspect of philosophical cognition, but for decades, it included scientific cognition and has now become even a technical cognition, implied, for example, in a moral code implementation in the field of machine ethics. Scientific moral cognition itself was manifested as empirical, theoretical and meta-theoretical cognition. Ethical forms of empirical cognition are today augmented by multimedia.

Even the above mentioned computational ethics (a philosophical ethics), which studies inclusively the history of ethics and the main ethical theories is, to a certain extent, an empirical ethics, because it applies to individuals and their ethical problems, it studies documents, articles, legal texts, presents biographies of pathfinders, and encourages modelling of problematic moral situations or proposes suitable decision procedures.

The study and use of the spiritual level of the moral system is more complex. Moral and technical cognition need to be joined for the conception and creation of a Machine Ethics. The resulting synthetic form of cognition can be founded inclusively by the comparative study of human and machine ethics. Such

a study may emphasize even the perspective of their co-evolution as the possible alternative: the evolution of human ethics itself towards an artificial ethics.

4.1.3. Moral Spirituality

Spirituality is another characteristic of moral life, but not an obligatory one. If it were not so, morality would not be a field of social life. Moral spirituality is grounded in moral culture, which is lacking from the life of the majority. Culture becomes a form of life by the subjective living of some specific values. Thus, even if values represent a structural level of the social system, they do not constitute a separate realm but a component of every element of the social system:

1- any social unit is a product of the social action; 2 - value functions both as goal and result of an action; 3 - action includes, even in its internal structure, value (as object, kind of motivation, criterion of choice between means or conditions, element of decision substantiation, basis for evaluation of the results and as norm for the opportunity to restart action in case of lack of congruence between goals and ends); 4 - value exists because it initiates actions (still as Plato conceived it).

Each subsystem of society is built by the corresponding type of social action, oriented by its own value system (economical, technical, political or moral), which constitutes the specific cultural field.

Moral culture is a system of moral values, structured around a central value, which can be common for all moral times and spaces. Other moral values can change their content and significance through modifications of their system of reference. All moral values are interpreted from different perspectives by ethical theories, which represent the cultural dimension of moral spirituality.

4.2 Moral Values

If we attempt to create and implement a moral code for machines (and possibly for humans) we might base our work on specific values, as in any other domain of action. Efficiency is the foundational value for each technical type of action, it is not specific (and, consequently, efficient) for machine ethics building.

Does an appeal to moral values help when building a new, universal morality, efficient and felicitous, for all of us, citizens, "netizens" or "intellizens", and neighbours in our artificial (technical, virtual, intellectual) world ?

We can observe that

a) moral values are general, synthetic, vague or fuzzy and evolving;
 b) the achievement of any moral value needs to use and cultivate one or many psychological aptitudes / properties:

- intelligence permits valorization of our capacity to do what is right
- imagination accomplishes the avoidance of possible bad acts
- consequence can found the habit of sincerity
- will can aid the doing of one's duty
- capacity for effort facilitates the maintenance of moral freedom
- attention, capacity to learn and intuition can ensure dignity.

Are these aptitudes or qualities accessible to intelligent machines? Possibly, but:

- c) embodiment of moral values is also conditioned by personality traits and cultural attitudes;
- d) central moral value accomplishment needs the achievement of all other moral values;
- e) the link between values and norms is not univocal;
- f) the mentioned discontinuity between values and norms is prolonged in the imprecision of norms; these are imperatives with a content that is determined only in concrete contexts of the real moral and lived life, so, only by life experiences;
- g) moral norms are related also to the characteristics of each moral/cultural community

4.3 Moral Norms

The solution of conceptual problems, identified by us in the process of searching possibilities to implement a moral code for machines, depends on our capacity to understand the moral phenomenon itself in the completeness of its components: spirituality, norms and activity.

However, the gap between cognition and action, values and norms can be attenuated by changes that have happened within (moral) theory itself: if theory is now a synthesizing and systematizing of empirical results, the norm can be determined as the result of the transposition of assertive sentences into prescriptive ones and, as an expected step in the process of the endless standardization of human activities, in their practical and intellectual forms.

Norms are then continuations of attainments, cognition in prescriptive or even imperative form. Theories become operational using norms of applicability, derived even from theory, and theory is proved if a functional model can be built and tested, along with norms elaborated on the basis of theory itself.

Moral culture is part of the culture of action, besides technical, political or juridical culture, and has a strong normative character.

Moral norms have some already known specific traits and also ones which have been less well-studied, but which become very important in this context. By eluding implementation of abilities supposed by such traits, our 'moral agents' can be in fact just ordinary robots, useful only for a very limited and very accurately defined job set.

Such are the characteristics:

1 – moral norms are available for any type of (human) action: all (human) activities have a moral dimension and can/must be evaluated from this perspective; 2 – morality is an obligatory requirement of all (human) activities, but application of moral norms in different domains needs i) the use of complex, for now exclusively human, abilities and ii) the acquisition of some cultural implements; 3 – the general, transposable and perennial moral norms can be correctly interpreted and efficiently used only by cultivating and using not just intellectual, but spiritual and creative qualities.

Here, then, are five of the most important difficulties of founding moral conduct, both for human and artificial agents:

I. Human ethics presents its own internal inadequacies; human morality cannot be corrected from the standpoint of moral theories. As an example: the ethics of the majority is not able to eliminate failures of individualist or collectivist theories.

II. A doubly profound incongruence is manifested between:

i) the theoretical and practical level of human morality; the behavioural and the spiritual levels of morality are evolving separately: towards science – human and technical, and respectively, towards anomy. Morality subsists almost only as professional deontology.

ii) the human and machine ethics: human morality cannot function as a model for machine ethics.

III. Some of the implicated intelligence forms are unstudied (such as the communicative and the evaluative intelligence), less studied (the interpretative intelligence), or only recently approached (emotional intelligence).

IV. Moral Intelligence is not considered and studied as a prerequisite to the implementation of a moral code.

V. Enhancement of specific or particular intelligence forms as the viable alternative to the extreme development of an abstract intelligence form, which is not even a general intelligence form, is not a priority for AI research, although it is already observable that, persevering in this, AIA can only gain in speed, manageability and universality. But these successes can also be interpreted as losses in profoundness or in self-determination and differentiation capacities.

5 Exploring Moral Intelligence - a Step towards a Moral Code Implementation

By studying the complex structure of the moral system we can more clearly observe the set of difficulties in the implementation of a moral code; analyzing (human) intelligence, we can more realistically evaluate the possibilities for creating efficient AIA.

Artificial agents which will work not only according to technical but to moral values and norms need to be endowed with adequate “aptitudes” and skills and, if possible, even with a “spiritual life”, not just with emotions, feelings or with simple (but difficult to recreate artificially) life.

In the specialized domain of psychology the structure of intelligence has been analyzed. An unequal attention was given to different forms of intelligence, the more easily-accessible (analyzable and quantifiable) ones being distinguished first; their classification has been made in conformity with ever-changing criteria.

Thus, some forms of intelligence appear in more than one class (as the case of the mathematical, linguistic or inter-personal intelligence), and diverse forms are associated in different kinds of classifications, for example, descriptive, imitative, interpretative and creative. But many real and important forms of intelligence are completely overlooked. Examples might include those used and developed at highly abstract and complex levels of any types of activities, and which are all treated simply as expressions of “creative intelligence”.

These theoretical imperfections are accompanied by educational failures, beginning with the absence of the specific objective to form and to develop intelligence: as a result, people cannot know, evaluate and use all the forms of intelligence available to them, even if every healthy human being has, potentially, all possible intellectual capacities. Those AI workers who want to keep the initial GOFAI design must also be aware of this.

We propose three new, non-psychological but philosophical classifications of forms of intelligence, all of which permeate the study and cultivation of moral intelligence, as the condition for the cultural existence of both human and artificial agents.

The first proposed criterion can be represented by the fields of culture in which diverse human abilities are improved, trained and manifested. According to this criterion, we can distinguish the scientific, artistic, technical, political or moral forms of intelligence.

The second criterion is a practical one: intelligence is constituted in conformity with requirements, norms and values from diverse domains of activity. Moral values form one of the value systems present in every social organization and their realization needs the building, cultivating and exerting of an ensemble of (human) aptitudes, such as moral intelligence. As we have shown, other aptitudes needed for achievement of every moral value can be identified, and vice versa, development of every inner, or acquired, quality necessitates specific activities, means and environments.

The third theoretical strategy to integrate moral intelligence in a coherent explicative system is not to distinguish, but to gather up all its necessary, specific, constant – defining – features, those which characterize it among all other complex forms of intelligence and not to separate it as an exceptional human ability. Thus, we think that moral intelligence includes all other human intelligence forms, in variable measure and intensity. It penetrates, for its formation and by its ways of manifestation, all intellectual spaces and processes.

Moral Intelligence, as a complex of concrete manifestations of these other forms of intelligence, is synthetic. Moral intelligence is not an abstract form of intelligence, nor is it even a special form of intelligence. It is neither a general nor a particular (abstract or practical) form of intelligence. Until now this form of intelligence has not been studied. Thus, it is, unfortunately, also under-cultivated.

Moral intelligence cannot be mistaken for general intelligence (the general intellectual functional disposition) and the degree of manifestation of this type of intelligence can be very different from the possibilities offered by the level of development of the first. Moral intelligence is active and efficient mainly in specific (moral) contexts, but such circumstances occur in all domains of activity and each person has moral experiences at all levels. Moral intelligence is not a special form of intelligence, just like some forms of artistic intelligence (as plastic, motor or musical ones) and scientific ones (mathematical, linguistic, analytical or synthetic, theoretical or practical).

Since it is impossible to reduce moral intelligence to one or other of the above-mentioned forms of intelligence, it cannot be isolated from them. Yet, as a complex, synthetic form of intelligence, moral intelligence seems to depend, not just on mentioned forms and levels of intelligence, but mainly on the level of development of moral conscience and on the degree of presence, interdependence and functionality of its analyzed components. In more imperative terms, moral intelligence is conditioned not principally by other more or less complex aptitudes (psychical factors), but by spiritual, educational and social (cultural factors).

However, the implementation of a moral code in a machine can be facilitated by using results of the study of some common aspects of human verbal and numerical intelligence, and the capacity for reasoning and memory. But we can remember findings of older psychologists as A. Ray, who showed that every successful (human) action is a result of the activity of the psyche as a whole. It will also be necessary to identify adequate factors to describe functions which can simulate complex behavioural abilities and cultural attitudes characteristic of moral conduct.

We do not have to implement a moral code, but to create a moral intelligence, we can aspire to a condition of potentiality, not the generation of some fixed reality. For humans potential intelligence is more important than a crystallized one. The gaining of a complex intelligence form will not necessarily draw the machine nearer to humans, but Machine Ethics might help to overcome some difficulties of human morality.

In conclusion, a Machine Ethics can be a) directly deduced from a moral theory, b) assisted by intellectual techniques, c) based on objective evaluation of possibilities related to necessities, d) implemented by technical means which assure precision, transparency and efficiency, and e) achieved by knowledge based technologies and cognitive robotics. But the belief generating or conditioning and controlling problems will appear at the spiritual level of the eventually artificial moral conscience.

A better understanding of Human Morality leads to the conclusion that Machine Ethics will not be a result of a simulation of human morality, but of a moral invention: an Artificial Ethics. But this ethics could be suitable even for the improvement of human ethics, as another expression of the present tendency towards artificiality.

References

- Arieti S., & Bemporad J.R. (1978), *Severe and mild depression: the psychotherapeutic approach*, New York: Basic Books.
- Bedau, M. A. (1998) „Philosophical Content and Method of Artificial Life”, in Bynum, T. W. and Moor, J. H. (eds.), *The Digital Phoenix: How Computers are Changing Philosophy*, Blackwell Publishers, Oxford
- Bynum, T. W. (1998) „Global Information Ethics and the Information Revolution”, in Bynum, T. W. and Moor, J. H. (eds.), *The Digital Phoenix: How Computers are Changing Philosophy*, Blackwell Publishers, Oxford
- P. M. Churchland (1998) „The Neural Representation of the Social World”, in Bynum, T. W. and Moor, J. H. (eds.), *The Digital Phoenix: How Computers are Changing Philosophy*, Blackwell Publishers, Oxford
- Danielson, P. (1998) „How Computers Extend Artificial Morality”, in Bynum, T. W. and Moor, J. H. (eds.), *The Digital Phoenix: How Computers are Changing Philosophy*, Blackwell Publishers, Oxford
- Darwall, S. (1998) *Philosophical Ethics*, Westview Press, Colorado, Oxford
- Ey, H. (1982) *Conștiința (The Consciousness)*, Editura Științifică, București
- Gregory, R. (2000) *Viitorul creatorilor de inteligență (The Future of Mind-Makers)*, Editura Științifică, București
- Hillis, W. D., *Mașina care gândește (The Pattern on the Stone. The Simple Ideas that Make Computers Work)*, Editura Humanitas, București, 2001
- Laruelle, Fr. (1990) *Théorie des identités, fractalité généralisée et philosophie artificielle*, P. U. F., Paris
- Londen, R. B. (1997) „Vices of the Virtue Ethics”, in Crisp, R.; Slote, M., *Virtue Ethics*, Oxford University Press
- Pană, L. (2002) *Cultura tehnică și industria culturală (Technical Culture and Cultural Industries)*, Editura Tehnică, București.
- Pană, L. (2005) „Filosofia artificialului și filosofia artificială” (The Philosophy of the Artificial and the Artificial Philosophy), in *Academica*, nr. 34, ianuarie, Anul XV, 171.
- Pană, L. (2000) „Cultura morală” (The Moral Culture), in Laura Pană, *Filosofia culturii tehnice (The Philosophy of Technical Culture)*, București, Editura Tehnică
- Pană, L. (2004 b) „Etica artificială”(Artificial Ethics) in *Filosofia informației și a tehnicii informaționale (The Philosophy of Information and Information Technology)*, Editura Politehnica Press, București

- Pană, L. (2004 a) "Modelarea unor aspecte ale evoluției sistemelor de valori al culturii românești prin prisma culturii tehnice" (A Model for the Evolution of some Aspects of the Actual Value Systems from the Perspective of the Technical Culture) in Laura Pană (ed.) *Evoluția sistemelor de valori sub influența culturii tehnice (Actual Evolutions of Value Systems under the Influence of the Technical Culture)*, Editura Politehnica Press, București
- Pană, L. 2005 "Moral Intelligence for Artificial and Human Agents", *Machine Ethics, Papers from the AAAI Fall Symposium Series*, Arlington, Virginia, 2005, November 4-6, AAAI Press, Menlo Park, California.
- Pană, L. (forthcoming) "The Intelligent Environment as an Answer to Complexity", *Proceedings of the XV IUAES Congress "Humankind/Nature Interaction: Past, Present and Future"*, Florence, 2003, in the volume *The Trans-disciplinary Flow of Our World*
- Rey, A., (1924) "Invention artistique, scientifique, pratique" in Dumas, G (ed.) *Traite de Psychologie, Tome II (Les fondements de la vie mentale)*, Premier livre, Chapitre VI, Librairie Felix Alcan, Paris
- Sloman, A. (1990) "Motives, Mechanisms, Emotions", in M. Boden (ed.), *The Philosophy of Artificial Intelligence*, Oxford University Press.