Gianmarco Veruggio/ Fiorella Operto: Ethical and societal guidelines for Robotics

To introduce our topic, which is a discussion on Roboethics, let us start from robotics as such, and from a statement by yours: You hold that Robotics is a new Science. Is this claim true? Or, is it a wish of some roboticists, who are trying to attribute higher dignity to their studies?

GIANMARCO VERUGGIO
In 2004, roboticists and scholars of humanities gathered in Sanremo, Italy, to lay the foundations of a new applied ethics, which I, as the Chair of the Symposium, had called "Roboethics". This word did not exist before, nor was in any Encyclopedia neither on Google. The two days workshop took place in a historical location, the studying room of Villa Nobel, and the mansion-house where Alfred Nobel lived his last years, and where he wrote his famous testament.

From 2004, five years have elapsed, and today Roboethics is a subject of authoritative discussion and studies; it is the topic of an ad hoc IEEE Robotics&Automation Technical Committee, and headline of many books.

In the next decades in the Western world – in Japan, United States, Europe – humanoids robots will be among us, companions to elderly and kids, assistants to nurse, physicians, firemen, workers. They will have eyes, human voices, hands and legs; skin to cover their gears and brain with multiple functions. Often, they will be smarter and quicker than the people they ought to assist. Placing robots in human environments inevitably raises important issues of safety, ethics, and economics. Sensitive issues could be raised by the so called "robotics invasion" of many non-industrial application sectors, especially with the personal robot; and the surveillance and military applications.

In many instances, I have tried to demonstrate that Robotics is indeed a new science, of a special kind. And that in the making of this new science we can understand in-depth many new fields of physical disciplines, as well as of Humanities. In the main, Robotics is in fact considered a branch of Engineering dealing with intelligent, autonomous machines. It shares knowledge with other disciplines, and it is somehow the linear sum of all these studies. On the other side, some of us regard Robotics as new science, in its early stage. Ultimately – we say – it

is the first time that humanity is approaching the challenge to replicate a biological organism. That is why Robotics holds this special feature of being a platform where Sciences and Humanities are converging – an experiment in itself.

To discuss this matter, let us start from a question: How is a new science born?

Thomas Kuhn says that "under normal conditions the research scientist is not an innovator but a solver of puzzles, and the puzzles upon which he concentrates are just those which he believes can be both stated and solved within the existing scientific tradition".[1]

However, he adds in another locus of the same work, that " (..)I think, particularly in periods of acknowledged crisis that scientists have turned to philosophical analysis as a device for unlocking the riddles of their field. Scientists have not generally needed or wanted to be philosophers".[2]

Let us think of chemistry, of physics, sciences originating from many original and even weird sources, and later on systematized by famous scientists whose mission was to order the knowledge in laws, principles and rules, applying mathematical methodology to structuring the cluster of confirmed experiences and cases. Sciences are syncretic creatures, daughters of rationality, non rationality and of societal forces.

Back to Robotics. As said before, it is the result of melting knowledge from many fields: Mechanics, Automation, Electronics, Computer Science, Cybernetics, and Artificial Intelligence. It also stems from Physics & Mathematics; Logic & Linguistics; Neuroscience & Psychology; Biology & Physiology; Anthropology & Philosophy; Art & Industrial Design. And, the more it develops, the more it floods into other disciplines, exceeding schemes and borders. A proof of the complexity of robotics comes from the 1600 pages of the monumental "Springer Handbook of Robotics"[3], the first encyclopedic volume existing in the literature devoted to advanced robotics, edited by Bruno Siciliano and Oussama Kathib.

There is another important element of development, and it is the boost in robotics' applications, which in turn is controlled by the so-called forces of the market: huge investments are funneled into it, from Japan's Meti 40 billion yen in the humanoids challenge, to the 160 billion dollars in the US Future Combat Systems program.

We are just on the brink of the development of our science, and it is hard to envisage its future. It may

happen that Robotics swallows up other sciences; or that, like the giant red stars, it will explode into many other sciences, originating from the intersections of adjoining fields.

FIORELLA OPERTO

Robotics: Much talking about it, but little known. Actually, despite investments, efforts and results, penetration in our societies and media scoops, Robotics is a science which is still relatively unknown, or little known, and often misrepresented. Seldom is the keyword ***Robotics*** read in the institutional Programmes, being mainly hosted in the ICT cluster, or hidden under different initials.

Sometimes I linger to ponder the under-studied inferiority complex of some engineers which prevents them attributing universal qualities to their work. This so called inferiority feeling derives – as the Italian scholar of studies in history and philosophy of science, Paolo Rossi, says – from ancient times, when *mechanicus* meant a vile and not noble man. Paolo Rossi writes:

> "At the roots of the great scientific revolution of the 17th century is the union between technology and science that has marked, for the good and the not so, the entire Western civilization. This union, that became marked in the 17th and 18th centuries and which perpetrated all over the word, was, however, absent in ancient and medieval civilizations. The Greek term *banausia* means mechanical art or manual labor. In Plato's *Gorgia*, Callicle states that a machine manufacturer ought to be despised; insulted, by being called a banausos; and that no one would consent to the marriage of their daughter to him. Aristotle had excluded the mechanical workers from the citizens'society and had said that they differed from slaves only due to the fact that they care for many individuals' needs whilst a slave only cares for one. The divide between slaves and free individuals tended to be made manifest by the division between techniques and science, the division between practically-orientated knowledge and knowledge dedicated to the contemplation of truth. The disdain with which the slaves were treated was equally transferred to their areas of work. The seven liberal arts of the trivium (grammar, rhetoric and dialectic) and of the quadrivium (arithmetic, geometry, music and astronomy) are so named liberal due to their belonging to free individuals, and not to the non free individuals, or to the slaves who practiced mechanical or manual arts. Knowledge not directed towards a specific end but collected for its own intrinsic value is the only key to discovering the true nature of humankind. The practice of

sophia requires wealth and the presence of life's fundamentalities. Philosophy needs the mechanical arts upon which it is based, however, they are inferior forms of knowledge that are immersed between the material and the sensible and which are linked to manual and practical labor. The wise and learned individuals ideals tends to coincide (as it does in Stoic and Epicurean philosophy and later in Thomas Aquinas' thoughts) with the image of one who dedicates his life to contemplation while waiting for (like the Christian thinkers) the bliss of contemplating God".[4]

GIAMARCO VERUGGIO
In fact, it was in the Italian Renaissance that the profession, and word, "engineer", or, more precisely, geometrician, architect indicated a profession of equally importance as scientist, artist and of socially acknowledged leadership. Maybe, one of the reasons for this underestimate is the paradox of Engineering, which is, on the one hand, an arid, stark, abrupt and operative science; on the other side, the making of a craftsman, often of a true artist.

Even Ove Arup, the leading Anglo-Danish engineer, said that: "Engineering is not a science. Science studies particular events to find general laws. The projecting activity of the engineer uses those laws to solve particular problems. In this, it is closer to the art or handicraft: problems are under-defined and there are many solutions, good, bad and indifferent. The art is finding a good solution through a compromise between media and scopes. This is a creative activity, that requires imagination, intuition and deliberative choice."[5]

I believe that roboticists should get a sense of their creative potential and of the importance of their scientific contributions, with method and rigor. We are perhaps witnessing some hints of getting this sense.

I see, robotics is more known through media exaggerations and novelists' stories, Terminators and Wall-e robots?

GIANMARCO VERUGGIO
It is really true! And, here, you have another "siding-mission" of Roboethics.

In the 18th century, one of the aims of scientists working in the field of electromagnetism was to remove magic from the physical phenomena they were interpreting. And we roboticists have to do just that, freeing our field from the magical conception still dominant today in many layers of the population. We are suffering from the heavy burden of literature and fiction, which overimposes on our products their profile and patterns. A tough life for

Robotics: the less people know about it, the more they talk about, and demand from it.

The general population knows about Robotics what it watches in the Sci-Fi movies, which feed any kind of irrational fear about robots being disobedient, rebelling, good or evil souls, conscious and in love, emotional creatures lacking only the quality of total freedom. Robotics stimulates some of humanity's most fundamental questions. This means that we shall not expect some simple answers to be conclusive. The undertaking of discovering the essence and origin of human' intelligence and self-consciousness is as tough and troubling as the challenge around the unification of physical forces, or the research on the origin of the Universe. Simplistic answers could lead to gross mistakes and we cannot obtain correct answers if we ask the wrong questions.

I had hard times witnessing discussions on the rights for robots; on robotics' superiority to humans; on the development of robots to other biological, top-dog species. The sad side of the story is that often it is us, the roboticists, who are responsible of repeating, or fostering such legends, for narcissism or fashion of being philosophers. I believe that we have to use clear thinking, from now on. We would need other myths, images, and metaphor, which are truly intrinsic and proper to Robotics, and not to the anthropology of the human/automata tragedy and legend. Real robotics is far more exciting that fantasy!

For instance, one of the problems to be addressed, with the support of scholars of humanities, is that in robotics, current uses of words such as knowledge, intelligence, representation, intention, emotion, social agent, autonomy, and humanoid are potentially misleading – insofar as it is thereby suggested that typically human mental properties can be indifferently and unproblematically attributed to technological artifacts, disregarding from the current limitations of state-of-the art robotic systems.

But, ultimately, what is a robot?

GIANMARCO VERUGGIO
A robot is an autonomous machine that is capable of performing a variety of tasks, gathering information about its environment (senses) and using it to follow instructions to do work. Nothing really romantic about it! On the other side, robots are the machines which are more similar to humans than anything we've ever built before, and this makes it easier, for people who don't know the subject, to speak about robots as if they were humans.

This peculiarity has favored the rise of all the legends about robots: That

they will rebel against humankind; that they can "evolve" becoming humans, super-humans, and so on. One day, we could also be witnessing the birth of weird "robot worshipper" sects claiming some nutty visions about robots. But, this misconception could also be generating suspects and diffidence in the traditional cultures that could in turn lead them to raise obstacles to the research and development of the most advanced robotics applications.

Let us discuss concisely one of the most popular myth: the one we could call *Pinocchio principle*, that is the idea that humanoid robots could evolve into humans. Basically, the legend embodied in the *Pinocchio principle* is that reproducing ever more perfectly the human functions coincides with producing a human being. Although it is picked up by many scholars, we recognize in it an acknowledged flaw of reasoning and of composition. In fact, even if we could design and manufacture a robot endowed with symbolic properties analogous to those of humans, the former would belong to another, different species.

Actually, human nature is not only the expression of our symbolic properties, but also the result of the relationships matured during our extra uterine development (we are Nature AND Culture). There is a very important concept that is ***embodiment***, which means that an intelligence develops in a body and that its properties cannot be separated by it. A very enlightening article was written by José Galvan in the December 2003 issue of IEEE Robotics & Automation Magazine, "On Technoethics", where it is said, among other things: "The symbolic capacity of man takes us back to a fundamental concept which is that of free will. Free will is a condition of man which transcends time and space. Any activity that cannot be measured in terms of time and space can not be imitated by a machine because it lacks free will as the basis for the symbolic capacity".

It is quite obvious that when a machine displays an emotion, this doesn't mean that it feels that emotion, but only that it is using an emotional language to interact with the humans. It is the human who feels emotions, not the robot! And attributing emotions to the robot is precisely one of these human emotions.

We humans understand the world around us (people, nature, or artifacts) through emotional interaction. Long interaction can result in attachment (it may also provoke boredom). Interaction stimulates humans, and generates motivations for behaviour. Human interaction with the world always involves emotions.

There are useful objects and aesthetic objects, each of them evoking

different emotions in humans. Machines are also artifacts. Different from the aesthetic objects, machines have been designed and developed as tools for human beings. But usually, machines are passive, so human interaction with them is limited. But when a machine expresses its ability to act in a semi-voluntary way (as in the case of robots, which have been designed and programmed to learning), they have much influence on human emotions because the machine's behaviors may be interpreted by humans as emotional and voluntary. Furhermore, a machine with a physical body is more influential on the human mind than a virtual creature.

In the field of human-robot interaction, there are many studies on all these topics. Mit's kismet is one; also, all the projects involving robot pet therapies (for instance, the robot Paro, designed by Japanese roboticists Takanory Shibata[6], or those robotic playmates which can help children with autism.

It is a truly complex field, because the results depend very much on the cultural context and on the-background of the human actors involved.

From the Roboethics point of view, the sensitive issues concern the human right for dignity and privacy. In the case of robots employed by children, the concern pertains to the domain of the relationship of the kids with the world, their ability to distinguishing robot from living creatures and the danger of technological addiction.

In no way, however, in my opinion, a robot can "feel" emotion, at least, not in the way we do it.

Much of the mess about the robot's consciousness, robot's emotions, and robot's rights are based on the confusion generated by the use of the same words for intrinsically different items. That's why, discussing with philosophers in Europe and United States, we agreed that it could be worth expressing these ontological differences through a specific notation.

This is not a very original idea! For instance, in mathematics, the estimate of the variable x (exact or "truth" value) is referred to as "x hat", while its measure is indicated as "x tilde".

I am an engineer, and I am talking as a scientist, aiming at applying – when reasoning about philosophy of science – the same rigor I should employ in my daily work. For this, I would propose that we indicate with a "star" the properties of our artifacts, to distinguish them from those of the biological beings.

- HUMANS have INTELLIGENCE
- ROBOTS have INTELLIGENCE* (INTELLIGENCE STAR)

This could be a first, very simple way to keep us aware of these ontological differences, and at the same time it can help in avoiding flaws in our reasoning, like this:

- DOGS have four legs
- The THING that I see here has four legs

Therefore

- The THING that I see here is a DOG

For the sake of thruth, it necessary, even when we discuss the philosophy of our science, that we engineers apply the same sharpness as Galileo recommended in his synthesis of the Scientific Methodology: "Necessary demonstrations and sense experiences".

From all this, the necessity for the robotics community to become the author of its own destiny, in order to tackle directly the task of defining the ethical, legal and societal aspects of their researches and applications. Of course not alone, but collaborating with academics in the field of philosophy, of law, and general experts of human sciences. Nor should we feel relegated to a merely techno-scientific role, delegating to others the task of reflecting and taking action on moral aspects. At the same time, it is necessary that those not involved in robotics keep themselves up to date on the field's real and scientifically predictable developments, in order to base the discussions on data supported by technical and scientific reality, and not on appearances or emotions generated by legends.

I understand, from what you said, that Roboethics is, in your view, more than some deontological guidelines for designers and users?

GIANMARCO VERUGGIO

Roboethics is not the "Ethics of Robots", nor any "ethical chip" in the hardware, nor any "ethical behavior" in the software, but it is the human ethics of the robots' designers, manufacturers and users. In my definition, "Roboethics is an applied ethics whose objective is to develop scientific – cultural – technical tools that can be shared by different social groups and beliefs. These tools aim to promote and encourage the development of Robotics for the advancement of human society and individuals, and to help preventing its misuse against humankind.

Actually, in the context of the so-called Robotics ELS studies (Ethical, Legal, and Societal issues of Robotics) there are already two schools". One, let us refer to it as "Robot-Ethics" is studying technical security and safety procedures to be implemented on robots, to make them as safe as possible for humans and the plant. Roboethics, on the other side, which is my position, concerns itself with the global ethical studies in Robotics and is a human ethics.

FIORELLA OPERTO

Roboethics is an applied ethics that refers to studies and works done in the field of Science&Ethics (Science Studies, S&TS, Science Technology and Public Policy, Professional Applied Ethics), and whose main premises are derived from these studies. In fact, Roboethics was not born without parents, but it derives its principles from the global guidelines of the universally adopted applied ethics This is the reason for a relatively substantial part devoted to this matter, before discussing specifically Roboethics' sensitive areas.

Many of the issues of Roboethics are already covered by applied ethics such as Computer Ethics or Bioethics. For instance, problems – arising in Roboethics – of dependability; of technological addiction; of digital divide; of the preservation of human identity, and integrity; the applications of precautionary principles; of economic and social discrimination; of the artificial system autonomy and accountability; related to responsibilities for (possibly unintended) warfare applications; the nature and impact of human-machine cognitive and affective bonds on individuals and society; have already been matters of investigation by the Computer ethics and Bioethics.

A few lines about the "history" of Roboethics can be useful here to understand its aims and scope.

In January 2004, Veruggio, myself, in collaboration with roboticists and scholars of humanities organized the First International Symposium on Roboethics (Sanremo, Italy). Its aim was to open a debate, among scientists and scholars of Sciences and Humanities, with the participation of people of goodwill, about the ethical basis, which should inspire the design and development of robots.

Philosophers, jurists, sociologists, anthropologist and moralists, from many world's Nations as well as robotic scientists, met for two days contributing to lay the foundations of the Ethics in the design, development and employment of the Intelligent Machines, the Roboethics.

In 2005, EURON (European Robotics Research Network) funded the Research Atelier on Roboethics (project leader was School of Robotics) with the aim of developing the first Roadmap of a Roboethics. The workshop on Roboethics took place in Genoa, Italy, 27th February – 3rd March 2006. The ultimate purpose of the project was to provide a systematic assessment of the ethically sensitive issues involved in the Robotics R&D; to increase the understanding of the problems at stake, and to promote further study and transdisciplinary research. The Roboethics Roadmap – which was the result of the Atelier and of the following discussions and dissemination –

outlines the multiple pathways for research and exploration in the field, and indicates how they might be developed. The Roadmap embodies the contributions of more than 50 scientists, scholars and technologists, from many fields of science and humanities. It is also a useful tool to design a robotics ethic trying to embody the different viewpoints on cultural, religious and ethical paradigms converging on general moral assessments.

In the meantime, in the frame of the IEEE Robotics&Automation Society was organized a Technical Committee on Roboethics which is currently co-Chaired by Atsuo Takanishi, Matthias Scheutz, and Gianmarco Veruggio.

GIANMARCO VERUGGIO
One of the most ambitious aims of Robotics is to design an autonomous robot that could reach – and even surpass – human intelligence and performance in partially unknown, changing, and unpredictable environments. Artificial Intelligence will be able to lead robots to fulfil the missions required by the end-users. To achieve this goal, over the past decades roboticists have been working on AI techniques in many fields.

From this point of view, let us consider the fact that one of the fundamental aspects of the robots is their capability to learn: to learn the characteristics of the surrounding environment, that is, a) the physical environment, but also b) the living beings who inhabit it. This means that robots operating in a given environment have to distinguish human beings and living creatures from inorganic objects.

In addition to performing a learning capability about the environment, robots have to understand their own behaviour, through a self reflective process. They have to learn from the experience, replicating somehow the natural processes of the evolution of intelligence in living beings (synthesis procedures, trying-and-error, learning by doing, and so on).

All these processes embodied in the robots produce an intelligent machine endowed with the capability to express a certain degree of autonomy. It follows that a robot can behave, in some cases, in a way, which is unpredictable for their human designers. Basically, the increasing autonomy of the robots could give rise to unpredictable and non predictable behaviours.

Without necessarily imagining some Sci-Fi scenarios, in a few years we are going to be cohabiting with robots endowed with self knowledge and autonomy – in the engineering meaning of these words. This means, for instance, that we could have to impose limits – up to a certain extent –

on the autonomy of the robots, especially in those circumstances in which robots could be harmful to human beings.

In our roboetics studies, we have taken into consideration – from the point of view of the ethical issue connected to Robotics – a time range of a decade, a frame in which it could reasonably be located and inferred – on the basis of the current State-of-the-Art in Robotics – certain foreseeable developments in the field. Moreover, for the above mentioned reason, we have considered it premature to deal with problems inherent in the purely hypothetic emergence of human functions in the robot: like consciousness, freewill, self-consciousness, sense of dignity, emotions, and so on. Consequently, this is why the Roadmap does not examine problems like the need not to consider robots as our slaves, or the need to guarantee them the same respect, rights and dignity we owe to humans. I am convinced that, before discussing robot's rights, we have to ensure human rights to all the human beings on earth.

For instance, we have felt that problems like those connected to the application of robotics within the military and the possible use of military robots against some populations not provided with this sophisticated technology, as well as problems of terrorism in robotics and problems connected with biorobotics, implantations and augmentation, were pressing and serious enough to deserve a focused and tailor-made investigation. It is clear that without a deep rooting of Roboethics in society, the premises for the implementation of artificial ethics in the robots' control systems will be missing.

How can you envisage the definition of a Roboethics guideline protocol, which has to be shared by different cultures?

GIANMARCO VERUGGIO

Roboethics is a work in progress, susceptible to further development and improvement, which will be defined by events in our techno-scientific-ethical future. We are convinced that the different components of society working in Robotics, interested people and the stakeholders should intervene in the process, in a grassroots science experimental case: the Parliaments, Academic Institutions, Research Labs, Public ethics committees, Professional Orders, Industry, Educational systems, the mass-media.

To achieve this goal we need an internationally open debate because, concerning the role of science and technology in law, politics, and the public policy in modern democracies, there are important differences between each of the

European, the American, and the - we could say - oriental approach. But we live in the Age of Globalization and robotics will have a global market, just like computers, videogames, cars or cameras.

In the United States, the general attitude is definitely more science-based than it is in Europe. In the former case, science is said to speak the truth, and the regulatory process is based more on objective scientific data than on ethical considerations. At the same time, the subjective point of view is taken up by the courts, which are now also intervening directly in areas such as risks in society and scientific knowledge, although the current conceptual tools of jurisprudence in the field of science&technology are still very limited. Nonetheless, in the Anglo Saxon culture, “law does not speak the language of science”.

On the other side, in Europe, in the frame of the ongoing process of the culture’s cohesion, the course of regulation and legislation of science and technology assume a character of the foundation of a new political community - the European Union, which is centred around the relationship between science and its applications, and the community formed by the scientists, the producers, and the citizens. We can safely assume that, given the common classical origin of jurisprudence, the latter process could be helpful in influencing other cultures, for instance, the moderate Arab world.

There is a third way to approach issues in science&society it could be called oriental. In fact, in Japan and in the Republic of South Korea, issues of robotics&society have been handled more smoothly and pragmatically than in Europe and in America. Due to the general confidence from their respective societies towards the products of science&technology, the robotics community and the ad hoc ethical committee inside these governments have started to draw up guidelines for the regulation of the use of robotic artefacts. This non-ideological, non-philosophical approach has its pros and cons, but it could encourage scientists and experts in Europe and the United States to adopt a more normative position.

This means that also Roboethics – which is applied ethics, not theoretical – is the daughter of our globalised world. An Ethics which could be shared by most of the cultures of the world, and capable of being translated into international laws that could be adopted by most of the nations of the world.

While we analyze the present and future role of robots in our societies, we shall be aware of the underlying principles and paradigms which influence social groups and single individuals in their relationship with

intelligent machines. Different cultures and religions regard differently the intervention on sensitive fields like human reproduction, neural therapies, implantations, and privacy. These differences originate from the cultural specificities towards the fundamental values regarding human life and death. In different cultures, ethnic groups and religions the very concept of life and human life differs, first of all concerning the immanence or transcendence of human life. While in some cultures women and children have fewer rights than adult males (not even the habeas corpus), in others the ethical debate ranges from the development of a post-human status to the rights of robots. Thus, the different approach in Roboethics concerning the rights in Diversity (gender, ethnicity, minorities), and the definition of human freedom and Animal welfare. From these concepts, other specificities derive such as privacy, and the border between privacy and traceability of actions.

Cultural differences also emerge in the realm of natural vs. artificial. Think of the attitude of different peoples towards the surgical implants or the organs implantation. How could human enhancement be viewed? Bioethics has opened important discussions How is the integrity of the person conceived? What is the perception of a human being?

Last, but not least, the very concept of intelligence, human and artificial, is subject to different interpretations. In the field of AI and Robotics alone, there is a terrain of dispute– let's imagine how harsh could it be outside of the circle of the inner experts.

Because we said that there are big differences in the way the human-robot relationship is considered in the various cultures and religions, only a large and lengthy international debate will be able to produce useful philosophical, technical and legal tools.

At a technical level we need a huge effort by the standard committees of the various international organizations, to achieve safety standards, just like for any other machine or appliance.

At a legal level we need a whole new set of laws, regulating for instance the mobility of robots in the place of work or in public spaces, setting clear rules about the liability and accountability of their operations.

At a philosophical and ethical level, we need to discuss in depth the serious problem of the lethality of robots. This means that humankind has to decide if the license to kill humans should be issued to robots, for instance in military applications.

This is precisely the mission that led us to start and to foster the Roboethics Programme, and to develop

the Roboethics Roadmap. The basic idea was to build the ethics of robotics in parallel with the construction of robotics itself.

Actually, the goal was not only to prevent problems or equip society with cultural tools with enough time to tackle them, but a much more ambitious aim. Indeed I feel that robotics' development is not so much driven by inexistent abstract laws of scientific/technical progress, but moreover by complex relations with the driving forces of the economic, political and social system. And therefore dealing with roboethics means influencing the route of Robotics.

It is certainly a great responsibility, which however cannot be avoided. As the American roboticist George Bekey says : <We roboticists must walk to the future with our eyes wide open>. Indeed in society there cannot be a "Non-choice" stance.

Abstention ultimately ends up favoring the strongest, and in our case, in the current world political, social and economic system, this means one thing only: a development policy driven by the interests of multinational corporations. And, as the French roboticist Philippe Coiffet says: <A development in conformity with a Humanist vision is possible but initiatives must be taken because "natural" development driven by the market does not match with the desired humanist project.>[7]

From the ethical point of view, which kind of approach have you selected in structuring the fundamentals of the ethical issues in robotic?

GIANMARCO VERUGGIO
Given the relative novelty of the ELS issues in Robotics, the recommended ethical methodological approach here is that of the Applied Socio-Ethics.

Lacking an existing body of ethical regulations related to ethical issues in Robotics, scholars in the field (Tamburrini, Capurro et al., 2007) have proposed to sort a high value selection of case-studies in the most intuitively sensitive field on robotics applications (learning robots and responsibility, military robotics, human-robot interaction, surgery robotics, robotics cleaning systems, biorobotics). These cases were analyzed from the following point of view:

a) a technoscientific analysis (risk assessment; stability, sustainability and predictability); dependability assessment;
b) Shared ethical assumptions: liberty, human dignity, personal identity, moral responsibility and freedom (European Charter of Fundamental Rights; UN Chart of Human Right and related documents);

c) General Cultural assumptions (the way we live in Europe, our shared values and future perspectives, the role of technology in our societies, the relationships of European citizenship to technology and robots; our shared notions of social responsibility, solidarity and justice). Successively, a cross- check analysis was carried out between techno-ethical issues and ethical regulations.

Let us look at one case. In the field of service robots, we have robot personal assistants, machines which perform tasks from cleaning to higher tasks like assisting elderly, babies, disabled people, students in their homework, to the entertainment robots. In this sector, ELS issues to be analyzed concern the potection of human rights in the field of human dignity, privacy, the position of humans in control hierarchy (non instrumentalization principle). The right to human dignity implies that no machine should be damaging a human, and it involves the general procedures related to dependability. From this point of view, robotics personal assistants could raise serious problems related to the reliability of the internal evaluation systems of the robots, and to the unpredictability of robots' behavior. Another aspect to be taken into account, in the case of autonomous robots, is the possibility that these were controlled by ill-intentioned people, who can modify the robot's behavior in a dangerous and fraudulent manner. Thus, designers should guarantee the traceability of evaluation/actions procedures, and the identification of robots.

On a different level, we have to tackle the psychological problems of people who are assisted by robots. Lack of human relationships where personal connections are very important (e.g. for elderly care or edutainment applications) and general confusion between natural and aryificial, plus technological addiction, and loss of touch with the real world – in case of kids – are some of the psychological problems involved.

FIORELLA OPERTO

We can underline other kinds of ethical issues involving personal robots. For instance: The emerging market of personal service robots is driving researchers to develop autonomous robots that are natural and intuitive for the average consumer who can interact with them, communicate, work and teach them. Human-Robot interaction is developing along the innovative field of the so-called "emotional" or "social" robots, capable of expressing and evoking emotions. These social robots (employed especially in education, edutainment, care, therapy, assistance or leisure) are produced for the average non-expert consumer, and are supposed to display "social" characteristics

and competencies, plus a certain level of autonomous decision-making ability. They are endowed with: a) natural verbal and non-verbal communication (facial expressions, gestures, mimicking); b) embodiment (that is, in our case, how the internal representations of the world are expressed by the robots' body) and social situatedness; and emotions.

In the process of modelling human schemes of emotions, facial expressions and body language are often used gender, race and class stereotypes drawn from the approach of the empiricist psychology school. From the point of view of ethcal issues in robotics, it should be considered, and possibly avoided, to adopt the discriminatory or impoverished stereotypes of, e.g., race, class, gender, personality, emotions, cognitive capabilities, and social interaction.

The Institut für Religion und Frieden – which is the Editor of this booklet – is promoting a survey on one of the main sensitive aspect of robotics' applications – and of Roboethics: Military robotics. I am aware that you have intervened several times on this issue?

GIANMARCO VERUGGIO

Military research in robotics is being extensively supported, both in the United States and in Europe. Ground and aerial robotic systems have been deployed in warfare scenarios. It is expected that an increasing variety and number of robotic systems will be produced and deployed for military purposes in many developed countries.

While the design and development of autonomous machines opens up new and never faced issues in many fields of human activity, be they service robots employed in caring people (robots companion), or robots used in health care, those autonomous machines employed in war theatres are going to raise new and dramatic issues.

In particular, military robotics opens up important issues of two categories:

a) Technological;
b) Ethical.

Concerning technological issues, these are managed under the so-called Dual Use. Dual Use goods and technologies are products and technologies which are normally used for civilian purposes but which may have military applications. The main legal basis for controls on Dual-Use Goods is the EU Dual-Use Regulation (also known as Council Regulation 1334/2000 to be repealed by Council Regulation 428/2009, adopted 5 May 2009 and published in the OJ of the EU on 29 May 2009, L 134.) (European Commission, External Trade).

In the case of robotics machines, their behavour is affected by issues regarding the uncertainty of the stability of robot sensory-motor processes and other uncertainty questions. For this reason, in robotic systems that are designed to interact with humans, stability and uncertainty issues should be systematically and carefully analyzed, assessing their impact on moral responsibility and liability ascription problems, on physical integrity, and on human autonomy and robotic system accountability issues.

Actually, in modern robots the algorithms governing their learning and behavioral evolution, associated with operational autonomy, give rise intrinsically to the inability to forecast with the needed degree of accuracy each and all the decisions that the robot should take, under the pressure of the operational scenario in which it is employed at that moment.

This window of unpredictability is a well-known issue appearing in every robotics application field; but it involves some dramatic implications when applied to military robotics.

In this field, in fact, we have not only important ethical and humanitarian considerations, but also questions of operational reliability and dependability.

The very same military milieus have several times underlined the danger implied by the lack of reliability of robotics systems in a war theatre, especially when the urgency of quick decisions and the lack of clear intelligence concerning the situation requires the maximum control over its own forces.

This is particularly evident when human-in-the-loop conditions jeopardize timely robotic responses, possibly leading on this account to violations of task constraints and increased risk conditions. In view of current limitations of robotic technologies, robots do not achieve human-level perceptual recognition performances that are crucial, e.g., to distinguish friends or by-standers from foes.

In shaping responsibility ascription policies one has to take into account the fact that robots and softbots – by combining learning with autonomy, pro-activity, reasoning, and planning – can enter cognitive interactions that human beings have not experienced with any other non-human system (Tamburrini, Marino, 2006)

The issue is worsened by the extraordinary complexity of the robot's artificial intelligence control system. This issue makes these intelligent machines vulnerable from the point of view of their software's reliability. We all know, in fact, that no program is free from bugs affecting its behavior. Now, it's one thing when a bug is affecting a word processor

program, but it is different when a program's bug on a robot endangers the human lives the robot is supposed to protect.

The other side of the issues – also stressed by military spokesmen – in military robotics is the high risk of information security gap. Autonomous robot employed in war theatres could be intruded, hacked, attacked by viruses of several types, and become an enemy's tools behind our back.

In some cases, a responsibility gap could also arise, when human adaptation to service robots could cause some phase displacements in human's behavior whose consequences should be carefully considered. The beneficial possibilities provided by robotics remotely and tele operations; by robots serving as human avatars in inaccessible and dangerous areas; the availability through robots to intervene in micro and nanometer ranges could induce in humans the rise of gaps in responsibility (because of the perceived shared responsibility between human and robot) which could lead to disengagement from ethical actions); a gap in knowledge (the so called "videogame syndrome", that is when an operator perceives reality like in a video game), and gaps in actuality and reality.

The second categories of issues are of ethical and social class.

Human life has so high a value to justify a war and to accept the sacrifice of one or more lives to protect a human community.

However, just for this reason, the extraordinary importance and seriousness of the issues has imposed that in civilized societies only and always human beings can decide on the destiny of other human beings, and not automatic mechanisms, as sophisticated as they might be.

Only human beings endowed with the power of reasoning and of free will are endowed with the power of moral responsibility.

Ethical reflection does not justify the exceptions rule that every individual robotic action be submitted to human supervision and approval before its execution.

It is recommended that in human-robot shared action control provisions be made for assigning humans the higher rank in the control hierarchy which is compatible with cost-benefit and risk analyses. Furthermore, it is recommended that robotic systems which are justifiably allowed to override human decisions or to act independently of direct human control or supervision be systematically evaluated from an ethical viewpoint. (Eticboths project, deliverable 5)

For all these consideration, although very briefly summarized, I am deeply

convinced that to attribute a "license to kill" to a robot is a decision of such an extreme gravity that no Nation or community alone can do it by itself. This question must be submitted to a deep and thorough international debate

The further development of a broad ethical framework as an enabling factor for the public to participate in discussions on dual use of robots is highly desirable, together with deliberative technology assessment procedures (for example consensus conferences) backed by technologically informed education initiatives. Suitable policies and actions fostering awareness about the dual use robots are highly recommended at the level of European society. Support of extensive initiatives in dual use problem dissemination and interdisciplinary techno-ethics community building is recommended too.

I am also deeply convinced that an "R" (robot) chapter should be added to the NBC treaties, discussing the guidelines for the use of robots in war theaters. As in the case of many new weapon systems, also in our case, military robotics, we will be witnessing many political, social, and philosophical stands. From the "ban the bomb" (there will be people fighting for "ban robot weaponry" or, "ban the killer robots") to all the nuances of military agreement's proposals, as we have had for the ABC weapons.

[1] Kuhn, Th. The Essential Tension. Tradition and Innovation in Scientific Research, The Third University of Utah Research Conference on the Identification of Scientific Talent, ed. C. W. Taylor Salt Lake City: University of Utah Press 1959.
[2] Kuhn Th., idem.
[3] Springer Handbook of Robotics, Siciliano, Bruno; Khatib, Oussama (Eds.), 2008.
[4] Paolo Rossi, Daedalus sive mehanicus: Humankind and machines, Lecture at the Euron Atelier on Roboethics, Genoa, Feb-Narch 2006. In: http://www.scuoladirobotica.it/lincei/docs/RossiAbstract.pdf.
[5] Ove Arup,1895-1988 http://www.arup.com/arup/policies.cfm?pageid=1259.
[6] http://www.aist.go.jp/MEL/soshiki/robot/biorobo/shibata/shibata.html.
[7] Ph. Coiffet, Conference' speech, International Symposium on Roboethics, 30th - 31st January, 2004, Villa Nobel, Sanremo, Italy, "Machines and Robots: a Questionable Invasion in Regard to Humankind Development".