ICRA'07

2007 IEEE International Conference
on Robotics and Automation
10-14 April 2007, Roma, Italy

WORKSHOP on ROBOETHICS

The Roboethics Roadmap

Gianmarco Veruggio, Fiorella Operto
Scuola di Robotica
Genova, Italy

Scuola di Robotica

EURON

EURON (European Robotics Research Network) aims to promote excellence in robotics by creating resources and exchanging the knowledge we already have, and by looking to the future (http://www.euron.org).

The means to achieve this objective are fivefold:

1. Research Coordination
1. Joint Programme of Research
   - Prospective Research Projects
   - Topical Research Studies
   - **Research Ateliers**
3. Education & Training
4. Industrial Links
5. Dissemination

EURON Robotics Research Roadmap

One major product of EURON is a robotics research roadmap designed to clarify opportunities for developing and employing advanced robot technology over the next 20 years. The document provides a comprehensive review of state of the art robotics and identifies the major obstacles to progress.

The main goals of the roadmapping activity are to identify the current driving forces, objectives, bottlenecks and key challenges for robotics research, so as to develop a focus and a draft timetable for robotics research in the next 20 years.

The EURON Roboethics Atelier Project

In 2005, EURON funded the Research Atelier on Roboethics, with the aim of drawing the first Roboethics Roadmap.
The ultimate purpose of the project was to provide a systematic assessment of the ethical 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 EURON Roboethics Roadmap

The Roboethics Roadmap 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 and technologists, in many fields of investigations from sciences and humanities.
This study will hopefully be a useful aid in view of cultural, religious and ethical differences.

Disclaimer

Let's see firstly what the Roboethics Roadmap cannot be:

- It is not an exhaustive picture of the State-of-the-Art in Robotics, nor a guideline of ethics in science and technology. The reason is that Robotics is a new science still in the defining stage.

- It is not a list of Questions & Answers. Actually, there are no easy answers, and the complex fields require careful consideration.

- It is not a Declaration of Principles. The Euron Roboethics Atelier cannot be regarded as the institutional committee of scientists and experts entitled to draw a Declaration of Principles on Roboethics.

Scope: Near Future Urgency

In terms of scope, we have taken into consideration – from the point of view of the ethical issues connected to Robotics – a temporal range of a decade, in whose frame we could reasonably locate and infer – on the basis of the current State-of-the-Art in Robotics – certain foreseeable developments in the field.
For this reason, we consider premature – and have only hinted at – problems inherent in the possible emergence of human functions in the robot: like consciousness, free will, self-consciousness, sense of dignity, emotions, and so on. Consequently, this is why we have not examined problems –debated in literature – like the need not to consider robot as our slaves, or the need to guarantee them the same respect, rights and dignity we owe to human workers.

Target: Human Centred Ethics

Likewise, and for the same reasons, the target of this Roadmap is not the robot and its artificial ethics, but the human ethics of the robots' designers, manufacturers and users.
Although informed about the issues presented in some papers on the need and possibility to attribute moral values to robots' decisions, and about the chance that in the future robots might be moral entities like – if not more than– human beings, we have chosen, in this 1.0 release of the Roboethics Roadmap, to examine the ethical issues of the human beings involved in the design, manufacturing, and use of the robots.

Target: Human Centred Ethics (2)

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 urging and serious enough to deserve a focused and tailor-made investigation..

It is absolutely clear that without a deep rooting of Roboethics in society, the premises for the implementation of an artificial ethics in the robots' control systems will be missing.

Methodology: Open Work

The Roboethics Roadmap is an Open Work susceptible to further development and improvement which will be defined by events in our technoscientific-ethical future. We are convinced that the different components of society working in Robotics, and the stakeholders in Robotics should intervene in the process of building a Roboethics Roadmap, in a grassroots science experimental case:

- The Parliaments
- Academic institutions and Research Labs
- Public ethics committees
- Professional Orders
- Industry
- Educational systems
- The mass-media

Methodology: Open Work

The Roboethics Roadmap is an Open Work susceptible to further development and improvement which will be defined by events in our technoscientific-ethical future. We are convinced that the different components of society working in Robotics, and the stakeholders in Robotics should intervene in the process of building a Roboethics Roadmap, in a grassroots science experimental case:

- The Parliaments
- Academic institutions and Research Labs
- Public ethics committees
- Professional Orders
- Industry
- Educational systems
- The mass-media

Ethical Issues

Here below are some of the ethical issues connected to the Roboethics Roadmap which can differ, in their definition and application, according to cultures, religions and societies:

- Concepts of Immanentism / Transcendentalism;
- Perception of human being / Integrity of the person;
- Diversity (Gender, Ethnicity, Minorities);
- Human enhancement;
- What is human? post-human? Cyborg?
- Human life/artificial life;
- Human intelligence/artificial intelligence;
- Freedom;
- Privacy vs. traceability of actions;
- What is science/knowledge?

Principles to Be Followed in Roboethics

- Human Dignity and Human Rights
- Equality, Justice and Equity
- Benefit and Harm
- Respect for Cultural Diversity and Pluralism
- Non-Discrimination and Non-Stigmatization
- Autonomy and Individual Responsibility
- Informed Consent
- Privacy
- Confidentiality
- Solidarity and Cooperation
- Social Responsibility
- Sharing of Benefits
- Responsibility towards the Biosphere.

Ethical Issues in an ICT Society

Roboethics shares many "sensitive areas" with Computer Ethics and Information Ethics. But, before that, we have to take into account the global ethical problems derived from the Second and Third Industrial Revolutions, in the field of the relationship between Humans and Machines:

- Dual-use technology;
- Anthropomorphization of the Machines;
- Humanisation of the Human/Machine relationship;
- Technology Addiction;
- Digital Divide, socio-technological Gap;
- Fair access to technological resources;
- Effects of technology on the global distribution of wealth;
- Environmental impact of technology.

Computer and Information Ethics

From the Computer and Information Ethics we borrow the known Codes of Ethics called PAPA, acronym of: privacy, accuracy, intellectual property and access.

Privacy: What information about one's self or one's associations must a person reveal to others, under what conditions and with what safeguards? What things can people keep to themselves and not be forced to reveal to others?

Accuracy: Who is responsible for the authenticity, fidelity and accuracy of information? Similarly, who is to be held accountable for errors in information and how is the injured party to be made whole?

Property: Who owns information? What are the just and fair prices for its exchange? Who owns the channels, especially the airways, through which information is transmitted? How should access to this scarce resource be allocated?

Accessibility: What information does a person or an organization have a right or a privilege to obtain, under what conditions and with what safeguards?

Engineering Ethics

By Engineering Ethics are meant the Codes of Ethics bearing on the professional responsibilities of engineers, guiding to a responsible conduct in research and practice. In this context, Security and Reliability are the most important ethical codes of conduct. Furthermore:

· Hold paramount the safety, health and welfare of the public.
· Perform services only in areas of their competence.
· Issue public statements only in an objective and truthful manner.
· Act in professional matters for each client as faithful agents/trustees.
· Avoid improper solicitation of professional assignments.

(American Council of Engineering Companies Ethical Guidelines)

Ethics in Science and Technology

"What is science? Not the collection of facts but the establishment, through open debate, of new principles that command wide acceptance. The process of incorporating ethical concerns and recommendations in daily application." (John Polanyi, Nobel Laureate)

How can the ethical principles discussed in transdisciplinary assemblies; expressed by warnings or the public's concern; suggested by religious personalities, theologians, and moral leaders; and/or forwarded by a community of concerned scientists be incorporated in the current application of research and development?

Codes of Conducts

Here below the main social and institutionalized forms of codes of conducts:

- Oath & pledge
- Code & guideline
- Appeal
- Recommendation
- Manifesto
- Statement & declaration
- Resolution
- Convention
- Charter
- Law

Universally Adopted Ethical Principles

In roadmapping Roboethics, we refer to the General Ethical Principles adopted by most Nations, Cultures and People of the World.

Among them:

- United Nations Universal Declaration of Human Rights (1948)
- Unesco: Declaration on Science and the use of scientific knowledge (1999)
- Charter Of Fundamental Rights of the European Union (2000)

EU Charter Of Fundamental Rights

Articles of interest:

- Human dignity
- Right to life
- Right to the integrity of the person
- Respect for private and family life
- Protection of personal data
- Freedom of the arts and sciences
- Freedom to choose an occupation and right to engage in work
- The rights of the child
- The rights of the elderly
- Integration of persons with disabilities
- Environmental protection
- Consumer protection

Robotics and Ethics

Is Robotics a new science, or is a branch or a field of application of Engineering?

Actually Robotics is a discipline born from:

- Mechanics
- Physics/Mathematics
- Automation and Control
- Electronics
- Computer Science
- Cybernetics
- Artificial Intelligence

This shows that Robotics is a unique combination of many scientific disciplines, whose fields of applications are broadening more and more, according to the scientific and technological achievements.

Specificity of Robotics

It is the first time in history that humanity is approaching the threshold of replicating an intelligent and autonomous entity. This compels the scientific community to examine closely the very concept of intelligence – in humans, animals, and of the machines – from a cybernetic standpoint.

In fact, complex concepts like autonomy, learning, consciousness, evaluation, free will, decision making, freedom, emotions, and many others shall be analysed, taking into account that the same concept shall not have, in humans, animals, and machines, the same semantic meaning.

Robotics and the Two Cultures

From this standpoint, it can be seen as natural and necessary that Robotics draws on several other disciplines:

- Logic/Linguistics
- Neuroscience/Psychology
- Biology/Physiology
- Philosophy/Literature
- Natural History/Anthropology
- Art/Design

Robotics de facto unifies the so called two cultures, *Science* and *Humanities*. The effort to design Roboethics should make the unity of these two cultures a primary assumption. This means that experts shall view Robotics as a whole - in spite of the current early stage which recalls a melting pot – so they can achieve the vision of the Robotics' future.

The Rebellions of Automata

The theme of the relationship between humankind and autonomous machines – or, automata - appeared early in world literature, developed firstly through legends and myths, more recently by scientific and moral essays.
The topic of the rebellions of automata recurs in the classic European literature, as well as the misuse or the evil use of the product of ingenuity. It is not so in all the world cultures: for instance, the mythology of the Japanese cultures does not include such paradigm. On the contrary, machines (and, in general, human products) are always beneficial and friendly to humanity.
These cultural differences in attitudes towards machines are a subject the Roboethics Roadmap should take into account and analyse.

Common Questions

Some examples of common questions:

- Although farsighted and forewarning, could Asimov's three Laws become really the Ethics of Robots?
- Is Roboethics the ethics of robots or the ethics of robotic scientists?
- How far can we go in embodying ethics in a robot? And, which kind of "ethics" is the correct one for Robotics?
- How contradictory is, on one hand, the need to implement Roboethics in robots, and, on the other, the development of robot autonomy?
- Is it right that robots can exhibit a "personality"?
- Is it right that robots can express "emotion"?

What is a Robot?

Robotics scientists, researchers, and the general public have about robots different evaluations, which should taken into account in the Roboethics Roadmap:

- Robots are nothing but machines
- Robots have ethical dimensions
- Robots as moral agents
- Robots, evolution of a new specie

Robots Are Nothing but Machines

Many consider robots as mere machines - very sophisticated and helpful ones - but always machines. According to this view, robots do not have any hierarchically higher characteristics, nor will they be provided with consciousness, free will, or with the level of autonomy superior to that embodied by the designer. In this frame, Roboethics can be compared to an Engineering Applied Ethics.

Robots Have Ethical Dimensions

In this view, an ethical dimension is intrinsic within robots. This derives from a conception according to which technology is not an addition to man but is, in fact, one of the ways in which mankind distinguishes itself from animals. So that, like language and computers, but even more, humanoid robots are symbolic devices designed by humanity to extend, enhance, and improve our innate powers, and to act with charity and god intentions. (J. M. Galvan)

Robots as Moral Agents

Artificial agents, particularly but not only those in Cyberspace, extend the class of entities that can be involved in moral situations. For they can be conceived as moral patients (as entities that can be acted upon for good or evil) and also as moral agents (not necessarily exhibiting free will, mental states or responsibility, but as entities that can perform actions, again for good or evil).
This complements the more traditional approach, common at least since Montaigne and Descartes, which considers whether or not (artificial) agents have mental states, feelings, emotions and so on. By focusing directly on 'mind-less morality' we are able to avoid that question and also many of the concerns of Artificial Intelligence. (L. Floridi)

Robots, evolution of a new specie

According to this point of view, not only will our robotics machines have autonomy and consciences, but humanity will create machines that exceed us in the moral as well as the intellectual dimensions. Robots, with their rational mind and unshaken morality, will be the new species: Our machines will be better than us, and we will be better for having created them. (J. Storrs Hall)

Main Positions on Roboethics

Since the First International Symposium on Roboethics, three main ethical positions emerged from the robotics community (D. Cerqui):

- Not interested in ethics (This is the attitude of those who consider that their actions are strictly technical, and do not think they have a social or a moral responsibility in their work)
- Interested in short-term ethical questions (This is the attitude of those who express their ethical concern in terms of "good" or "bad," and who refer to some cultural values and social conventions)
- Interested in long-term ethical concerns (This is the attitude of those who express their ethical concern in terms of global, long-term questions)

Roboethics Taxonomy (1)

In the period of a year, the Euron Roboethics Atelier carried out a tour d'horizon of the field in Robotics: an overview of the state of the art in Robotics, and of the main ethical issues, driven by the most recent technoscientific developments, which can only just be glimpsed.

A taxonomy of Robotics is not a simple task, simply because the field is in a full bloom.

A classification of Robotics is a work in progress, done simultaneously with the development of the discipline itself.

Roboethics Taxonomy (2)

Aware of the classifications produced by the main Robotics organizations, which differ from one another on the basis of the approach – technological/applicational -, we have preferred, in the case of the Roboethics Roadmap, to collect the many Robotics fields from a typological standpoint, according to shared homogeneity of the problems of interface towards the society.

Instead of an encyclopaedic approach, we have followed - with few modifications - the classification of **EURON Robotics Research Roadmap**.
For every field, we have tried to analyse the current situation rather than the imaginable. Thus, we have decided to give priority to issues in applied ethics rather than to theoretical generality.

Roboethics Taxonomy (3)

- Humanoids
  Artificial Mind, Artificial Body
- Advanced production systems
  Industrial robotics
- Adaptive robot servants and intelligent homes
  Indoor Service Robots, Ubiquitous Robotics
- Network Robotics
  Internet Robotics, Robot ecology
- Outdoor Robotics
  Land, Sea, Air, Space
- Health Care and Life Quality
  Surgical Robotics, Bio-Robotics, Assistive Technology
- Military Robotics
  Intelligent Weapons, Robot Soldiers, Superhumans
- Edutainment
  Educational Robots, Robot Toys, Entertainment, Robotic Art

A Case: Humanoids

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.

“Essentially, it is expected that a robot will provide assistance in housework, for aged people and for entertainment to keep up the amenity of life and human environment in the next century. A type of human robot, a Humanoid is expected, to work together with human partners in our living environment, and it will share the same working space and will experience the same thinking and behaviour patterns as a human being. The robot will integrate information from sensors and show coordinated actions which realize a high level of communication with a human without any special training using multimedia such as speech, facial expression and body movement” (source, Waseda Humanoid Robotics Institute)

Humanoids Benefits

- Intelligent machines can assist humans to perform very difficult tasks, and behave like true and reliable companions in many ways.
- Humanoids are robots so adaptable and flexible that they will be rapidly used in many situations and circumstances.
- Their shape, and the sophisticated human-robot interaction, will be very useful for those situations where a human shape is needed.
- Faced with an aging population, the Japanese society foresees humanoid robots as one way to enable people to continue to lead an active and productive life in their old age, without being a burden to other people.
- Research carried out in humanoids laboratories over the world will have as a side effect the development of platforms to study the human body, for training, haptic test and trainings, with extraordinary outcomes on health care, education, edutainment, and so on.

Humanoids Problems (1)

- Reliability of the internal evaluation systems of robots.
- Unpredictability of robots' behaviour.
- Traceability of evaluation/actions procedures.
- Identification of robots.
- Safety. Wrong action can lead to dangerous situations for living beings and the environment.
- Security. In cases where the autonomy of the robot is controlled by ill-intentioned people, who can modify the robot's behaviour in dangerous and fraudulent ways.

Humanoids Problems (2)

Because humanoids incorporate almost all the characteristics of the whole spectrum of robots, their use implies the emergence of nearly all the problems we are examining below. In particular, their introduction in human environments, workplaces, homes, schools, hospitals, public places, offices, and so on, will deeply and dramatically change our society.

We have forecast problems connected to:

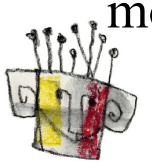
- Replacement of human beings (economic problems; human unemployment; reliability; dependability; and so on)
- Psychological problems (deviations in human emotions, problems of attachment, disorganization in children, fears, panic, confusion between real and artificial, feeling of subordination towards robots).
- Well before evolving to become conscious agents, humanoids can be an extraordinary tool used to control human beings.

Humanoids Recommendations

Activate working groups inside Standards Committees to study the possibility to define international technical/legal rules for commercial robots regarding:

- Safety. We should provide for systems for the control of robots autonomy. Operators should be able to limit robots autonomy when the correct robot behaviour is not guaranteed.
- Security: H/W and S/W keys to avoid inappropriate or illegal use of the robot
- Traceability: like in the case of sensitive systems, we should provide for systems like the aircraft's black box, to be able to register and document robot's behaviours.
- Identifiability: like cars and other vehicles, robots too should have identification numbers and serial numbers.
- Privacy: H/W and S/W systems to encrypt and password-protect sensitive data needed by the robot to perform its tasks or acquired during its activity.

Promote cross-cultural updates for engineering scientists that allow them to monitor the medium and long-term effects of applied robotics technologies.



The Roboethics Roadmap timetable

Ongoing workplan:

February 2006:

EURON Atelier on Roboethics;

May 2006:

Release 1.0 circulated among the Atelier participants;

September 2006:

Release 1.1 circulated among selected scholars & scientists;

January 2007:

Release 1.2 broadly circulated via Internet Forums;

April 2007:

Release 1.3 officially presented at ICRA2007 in Rome.