EUROP european robotics technology platform

Appendix to:

"Robotic Visions to 2020 and beyond – The Strategic Research Agenda for robotics in Europe, 07/2009"

Glossary of Ethical, Legal and Societal Issues of Robotics

Version: 1

Date: 07th July 2009

"Robotic Visions to 2020 and beyond – The Strategic Research Agenda for robotics in Europe, 07/2009" can be obtained from:

EUROP Secretariat
c/o EUnited Robotics
Diamant Building, Bd. A. Reyers 80
1030 Brussels, Belgium

Phone: +32 2706-8222
Fax: +32 2706-8223
Email: info@robotics-platform.eu
Internet: www.robotics-platform.eu

Supported by the "Coordination Action for Robotics in Europe", FP6-IST-045058

CARE
Coordination Action for Robotics in Europe

ELS Issues in robotics: Glossary and relevant terms

Accountability

It is responsibility of an individual or of an appointed department to perform a certain function. Accountability may be dictated or implied by law, regulation, or agreement. Professionals in any field can be called upon to justify their professional actions. The criteria against which they can be held accountable are those embodied in the normative standards of their particular profession. These standards are expressed in ethics guidelines or codes of conduct for each specialty area. They also may be embodied in law.
If any undesirable consequences result from a health professional’s work, every effort must be made to undo the damage, without dismissing or denying the problem. The definition of Accountability implies also the obligation by the person who is accountable to inform third parties about (past or future) actions and decisions, to justify them, and to suffer punishment in the case of eventual misconduct (Schedler A., 1999).

Anthropomorphisation of robots (as human emotional attitude and as a robots shape)

Anthropomorphic robots, with their morphology and sensing modalities evoking those of the humans, are regarded as especially useful as social interface; this is because the people’s mental model of autonomous robots are often anthropomorphic, in almost all the world cultures. According to several social roboticists, a human-like shape in robots could favor their integration in human’s environment and communities. Often, the robot’s shape is that of an infant, a woman-like or of a young boy (steward).
In this frame, and from the related ethical studies in robotics, the following issues should especially be taken into account. a) the misuses of infant, woman-like, or pet shape to attract consumer/users; b) the inherent perpetuation of longstanding stereotypes; c) women and others can be discouraged to use infant or woman-like shaped robots; illiterate people, or people with little familiarity with technology can be deceived by the shape of the robots believing them as social entities (Jutta Weber, *IRIE*, 2006).
On the other side, the design, shape and behavior of the robots has to confront the classic problem of the Uncanny Valley, whereby familiarity and acceptance rise as a robot becomes more human-like; however, at a certain point the behavior is too uncannily human-like but still far from being perfect, and the imperfections reverse the attraction and trust.

Applied Socio-Ethics

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 are to be 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 is to be carried out between tecno-ethical issues and ethical regulations.

Codes of Ethics and Standards

An ethical code (a.k.a. Professional ethics) is a collection of laws, or regulations; a written text that offers guidelines (rules, directives or principles for moral conduct) related to a particular field of activity.
Generally speaking, the guiding principles of any code of ethics are non-malfeasance and beneficence, indicating a systematic regard for the rights and interests of others in the full range of academic relationships and activities (engineering, health professionals, ect).

- Non-malfeasance is the principle of doing, or permitting, no official misconduct. It is the principle of doing no harm in the widest sense.
- Beneficence is the requirement to serve the interests and well-being of others, including respect for their rights. It is the principle of doing good in the widest sense (see IEEE Code of Ethics).

Although ISO suggest that parts of the current standards (OSO 1’218-2006) may also be applied in non-industrial robotics applications, there are no ISO regulations specifically for service and personal robots.

Cost benefit analysis (in robotics research and application)

In today’s economic frame, usually the cost benefit analysis of a technological device is left to the market. In the case of robots, their employment is judged relatively to their alternatives (as for instance in case of the industrial robots, or service robots). In many instances, the benefits are superior to the costs. However, the results of the comparison of the introduction of robots in medicine show that an overall cost benefit analysis for these techniques is not sufficient. Instead, especially in those cases in which the substitution of a human by a robot is taking place in the care of a ill, disabled and elderly person, it is necessary to determine the cost-benefit ratio in the concrete individual case, whereby the benefit has to be evaluated in the complete context. In particular, the question whether ethical issues are to be considered as part of a proper cost-benefit analysis has to be taken into account. This regards also the need for an equal world distribution of power and technologies.

Co-Workers (robotics)

Robots may eventually work with us or assist us under many different circumstances. They will interact with us directly, for example, to receive instruction or hand us a drink. This type of robot will need to be compatible with us and achieve safe and dependable operation, be it at the work place, in public, at home, or in space. They may be tele-operated or perform tasks or whole sequences of tasks autonomously. From this originate many ELS issues, which need to be taken into account. Of particular importance are the level of human-robot interaction, how to guarantee with regards to the transparency of robot’s learning processes, extension of human possibilities , robots may seem to human like (simulation of human traits), implications of multi agent decision making process (mixed teams), identification of autonomously acting robots , position of human in control hierarchy, the reduction of direct human contact resulting from robotic care assistants, and data protection issues related to data collection and processing through autonomous robots.

Dependability and Resilience

Advanced robotics covers many new areas of application which require safe and effective physical and cognitive interaction with humans. Key factors for the successful introduction of robots in our daily life is the availability of design paradigms and of enabling technologies that could minimize potential risks for end-users, avoid misuse and enhance overall acceptability of robotic artifacts (dependability). In many instances, robots will be performing their tasks in a human environment, not only surrounded by humans, but also in cooperation with them. That is why the other key quality of a service robot is the ability to maintain or recover a stable state when subject to disturbance (resilience). The field of dependability and resilience is a multidisciplinary research area that can greatly benefit from the advancements already made not only in the robotics research community, but also in other domains (dependability of software products and telematic network availability; respect to readiness for usage reliability; respect to continuity of service safety; respect to avoidance of catastrophic consequences on the environment security: prevention of unauthorized access and/or handling). (Giralt, Guglielmelli, 2008).
Another fundamental feature of robots cooperating with humans is the *Affordance*, that is the property of an object, or a feature of the immediate environment, that indicates how that object or feature can be interfaced with. (J. Gibson, 1966).

Dual Use Technology

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).

Gap (responsibility, knowledge, and actuality and reality)

In many instances, the human adaptation to service robots could cause some phase displacements in human's behavior whose consequences should be carefully considered. The beneficial possibilities provide 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 perceive reality like in a vide game), and gaps in actuality and reality.

Human Enhancing

In the field of professional service robots, prosthesis and bionic implants can be designed and applied to restore lost human functionalities. However, the thorough distinction between restoration and enhancement is problematic in many cases because of the unforeseeable side-effect of advanced technologies on the human bodies. Ethical issues can arise in the following areas: a) misuse of technologies (e.g, exoskeletons used by people with malicious intents); b) implant technologies could endanger people's rights to autonomy, freedom and privacy; c) restoration vs. perfectibility of the human body.

Human in control hierarchy and robot's transparency

Because of:

a) the growing complexity of robots;
b) b) of the difficulty of managing many unstructured areas of robots'operations;
c) c) of the unpredictability of learning machines' behaviour;
d) and of the ethical and legal necessities to assign liability in case of misbehaviours or crimes,

ethical and legal considerations have been invoked stressing the need that the authority principle of persons has to be maintained. The so-called "prohibition against instrumentalization", according to which persons may not be subjected to complete instrumentalization and never should considered as a means for someone or something else, would require that every individual robotic action be submitted to human supervision and approval before its execution.

From this point of view, robot's technical design of the control program should carefully consider the importance of human's decision authority. And, in order to allow humans to take responsibility for robot's behaviour, they should be controlled as for their transparency, forecast and influence.

Liability

In every case of liability ascription problems related to robot's malfunctioning or damages cause to humans and properties, learning robots should be distinguishable from non-learning robots, since by the use of learning algorithms, the liability for damage is affected between manufacturers and owners. The crucial decision is to acknowledge the distinction between liability or objective responsibility on the hand, and moral responsibility on the other hand. In principle, the subsequent decision is to apply or not to learning robots the rule about the inability of robots legal owners to prevent every possible damage caused by their devices.

More generally, the distinction between moral responsibility and liability is crucial to deal with responsibility problems in which one cannot systematically identify in a particular subject the sole or main origin of the causal chains leading to a damaging event. Producers of goods are held responsible in the absence of direct causal connections, on the basis of economic considerations that are aptly summarized in the Roman law principle *ubi commoda ibi incommoda* (where benefits, there liabilities) In these cases, expected producer profit is taken to provide an adequate basis for ascribing responsibility with regard to safety and health of workers or damages to consumers and society at large. (Marino, Tamburrini, *IRIE*, 2006).

In addressing and solving these responsibility ascription problems, one does not start from or rely uniquely on such things as the existence of a clear causal chain or the awareness of and control over the consequences of actions. The crucial decisions to be made concern the identification of possible damages and how compensation for these damages is to be determined and distributed (idem).

Liability issues arise also, in a very special way, in case of mixed teams of humans and robots.

Here the principle of designer responsibility is not readily applicable. New ideas are to be developed to understand newly emerging problems of allocating, monitoring, and diagnosing responsibilities.

Precautionary Principle (PP)

The Precautionary Principle (PP) was included in the 1992 Rio Declaration on Environment and Development ("Nations shall use the precautionary approach to protect the environment. Where there are threats of serious or irreversible damage, scientific uncertainty shall not be used to postpone cost-effective measures to prevent environmental degradation").

It has being increasingly invoked as a rationale for environmental health policy. However, it is a concept often difficult to define, and its implications to specific issues are not easily agreed upon

(von Schomberg 2006). Here below, the definition by UNESCO COMEST (World Commission on Ethics, 2005).

When human activities may lead to morally unacceptable harm that is scientifically plausible but uncertain, actions shall be taken to avoid or diminish that harm. Morally unacceptable harm refers to harm to humans or the environment that is
threatening to human life or health,

- or serious and effectively irreversible
- or inequitable to present or future generations
- or imposed without adequate consideration of the human rights of those affected.

The judgment of plausibility should be grounded in scientific analysis. Analysis should be ongoing so that chosen actions are subject to review. Uncertainty may apply to, but need not be limited to, causality or the bounds of the possible harm.
Actions are interventions that are undertaken before harm occurs that seek to avoid or diminish the harm. Actions should be chosen that are proportional to the seriousness of the potential harm, with consideration of their positive and negative consequences, and with an assessment of the moral implications of both action and inaction. The choice of action should be the result of a participatory process (..) To avoid misunderstandings and confusions, it is useful to elaborate on what the PP is not. The PP is not based on 'zero risks' but aims to achieve lower or more acceptable risks or hazards. It is not based on anxiety or emotion, but is a rational decision rule, based in ethics, that aims to use the best of the 'systems sciences' of complex processes to make wiser decisions. Finally, like any other principle, the PP in itself is not a decision algorithm and thus cannot guarantee consistency between cases. Just as in legal court cases, each case will be somewhat different, having its own facts, uncertainties, circumstances, and decision-makers, and the element of judgment cannot be eliminated" (The Precautionary Principle, World Commission on the Ethics of Scientific Knowledge and Technology, COMEST, 2005).

Socio-Ethics

Socio-ethics will be considered here as the ethics of sociality. This covers the relation of the individual with the group and with society, as the individual acquires the skills for social life with others and the conduct of 'normal responsible behaviour' that guides moral action. For a consideration of what it means to be socially skilled in everyday human interaction, and the ethical issues arising from the new conditions of interaction that come with the integration of intelligent interactive artifacts, we will sketch an analysis at multiple levels of these phenomena, drawing on variety of application domains (for example, healthcare and interactive media). (AAVV, Methodology for the identification and analysis of techno-ethical issues, European project "Ethicbots").

Stability and uncertainty issues

Robotics research and technological transfer efforts are extensively concerned with the stability of robot sensori-motor behaviour and related uncertainty .
It is recommended that in the ethical monitoring and evaluation of robotic systems that are designed to interact with humans, stability and uncertainty issues be systematically and carefully attended to, assessing their impact on moral responsibility and liability ascription problems, on physical integrity, and on human autonomy and robotic system accountability issues.

Stereotypes of Gender, Class and Race (adoption of)

The emerging market of personal service robots is driving researches to develop autonomous robots that are natural and intuitive for the average consumer who can interact with them, communicate, work and teach them (Breazeal, Takanishi, Kobayashi 2008). 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 express by the robots’ body) and social situatedness; and emotions.
In the process of modeling 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 (see also Ortony, Clore and Collins, 1988. The Ekman’s 6 fundamental emotions are Joy, Sadness, Anger, Surprise, Fear, and Disgust).
From the point of view of the ELS concerns, it should be considered, and possibly avoided, the adoption of discriminatory or impoverished stereotypes of, e.g., race, class, gender, personality, emotions, cognitive capabilities, and social interaction. (Weber 2007).

Usage of language in AI and Robotics

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