Relazione in International Conference on Humanoid Robots, IEEE Robotics and Automation Society, Waseda University, Tokyo 22-24 Novembre 2001
(testo orale non pubblicato)

Technoethics: acceptability and social integration of artificial creatures

1. Introduction

The objective of this paper is to contribute to the development of a new concept, Technoethics (TE), that is an innovative subject matter in the university world. I think that this issue is vital for the immediate future of mankind.

In spite of my poor English, I hope to adequately transmit these ideas in which I fervently believe. As it is not possible to summarize in few minutes the basic concepts of TE, the logical nexus of the argument will have plenty of “gaps”, that I'll try to fill out in the paper for publication.

2. TE fundamental elements

2.1. Mankind is technical by nature

TE‘s central theme is that mankind is technical by its very nature. Technology is not an addition to man but is, in fact, one of the ways in which mankind distinguishes itself from

animals.

Two very clear examples illustrate this fact in the ambit of western culture. They are as follows:

First, the myth of Prometheus, who steals fire from the gods in order to enable man to use it. Therefore, man is already able to use the technology provided. Man is thus capable of producing fire, which in fact, is a kind of technology. Is an animal capable of doing the same thing? I don't think so. Animals are provided with "natural tools" in order to survive, but the human person is born devoid of natural tools to survive. He or she alone has the capacity of producing "artificial tools".

The second example is that of Adam in the book of the Genesis. Adam was put in the garden of Eden "ut operaretur", in other words, to work in the garden. Adam was meant to plant the garden and work in it in order to reap its fruits and to improve the garden.

In both examples, an "unfinished condition of mankind" is emphasized, so that human beings are forced to interact with the material cosmos in order to produce technology. This interaction, guided by reason, can be generically called "work".

This is the main difference between, let's say, a bee and a man. Man consciously proposes a finality to his or her work. Mankind uses a rational approach while animals act and produce on the basis of instinct, which is inherent in nature.

2.2. Technology final aim

The difference between one who works with instinct and one who works with reason, is that the former works merely in order to survive, while the one who works utilizing their intellectual capacity gives *added value* to their life. By *added value*, I don't mean a product, but I mean *added value* to the very essence of man, in which mankind is improved. For

example, a bee who works simply accumulates more honey (she doesn't become “more bee”), a man who works becomes more complete in his or her own essence of man or woman. This is because the aim of human work is not limited only to provide for specific needs, but also – and more widely – to tend to the last truth of reality, in order to use it in the progress of mankind.

At the moment that there is that specific *extra value*, then human actions are finalized in order to provide a better quality of life. The concept of Good is now introduced in order to differentiate the meaning of “better” quality of life.

If there is *added value* to life and the possibility of incrementing value through actions, then this implies the preexistence of the whole central concept of Good.

The result of this logic, is that man is able to perform good or evil actions. If the concept of good is not preexistent, good itself becomes a relative concept, not an absolute one; in this case, actions would not have any ethical or moral value. Actions would be indifferent, and then everything would be permitted and the end would justify the means. To quote Ivan Karamazov “If God doesn’t exist, then everything is permitted”.

2.3. Ethical dimension of technology

But not everything is permitted. An example to clarify this concept is that even Mengele created progress … the problem is … is it the kind of progress which enhances the well-being of mankind?

Perhaps one of the central points of what is called *post-modernity* is the central conviction of the incapability of mere progress to provide answers to the radical question of human existence. In fact, this “added value” is not only progress. It cannot be understood only as a result of an immanent process of history. The idea of progress is relative, not

teleologically determined. If *added value* is only progress, then man becomes an indefatigable ant.

At this point, we must pose the central questions about the *added value* of mankind. Where does truth lie? Is it inherent in the preexistence of good? What is the good?

Because human beings are naturally social, the search for Good never can be selfish fact. Aristotle says that the completion of man's being is to have friends and positive interpersonal relationships; hence truth and good consist in intersubjectivity which is the sharing of the intentional objectives of intellect (truth) and free will (good) with others.

In interpersonal relationships the sharing of intellect and free will is manifest by the sincere giving of oneself to another. In this giving act, man includes not only the spiritual dimension but also the physical one. Therefore this act of giving to each other also incorporates the capacity of man to interact with physical matter, which is technology.

2.4. Three TE theorems and the paradigm-shift

This is then Theorem 1: The objective of technology is to increment human relationships both with a physical dimension and a spiritual one.

Returning to the beginning of the thesis, we began by stating that the creation of tools points to ethics and that ethics points to intersubjectivity and that intersubjectivity is the fundamental dimension to the production of material tools.

So, I hold that every man is an engineer. By this I mean that every human being has the ability to make his or her material actions the subject of interpersonal dialogue. This then, makes engineering one of the most humanistic activities there is.

This stands in stark contrast to the commonly held view of the engineer in the 20th century. In the immediate past, an engineer was simply subjected to the dictates of

politicians and economists.

This reasoning leads us to the shift of paradigms that we are currently living through. The paradigm of the XXth century is the dominion of reality through the knowledge of its laws. This was based only on closed scientific reasoning without transcendental considerations. Science has been given a dominant position in our society. The icon of the XXth century was the atom.

Now the paradigm is relational. We have learned from experience that mankind's rapport with reality can not be based exclusively on its scientific domain.

One must now establish a mutual interconnectivity between mankind and the rest of the universe. In the past, knowing was considered as something exclusively objective. Now, knowing is relational. Web technology facilitate us this change. For this reason, technology is now assuming its original priority, typical of the Italian Renaissance, that of technology over science. The icon of the XXIst century is the net.

An example of the priority of technology over science: with Einstein's theory of relativity it is virtually impossible to act with charity and goodness; with the production of a robot, this is not only possible but desirable.

The second theorem is that "When experimental science becomes technology it then becomes spiritual".

So the humanization of science by technology means that the human person is highlighted. In spite of Stanley Kubrik's opinion, human beings and machines will always be in harmony. This can be summarized in this third an final theorem.

Theorem 3 is based on Antoine de St. Exupery's book "Wind, Sand and Stars". The main idea is that the more a machine is developed the more it is taken for granted. The classic example is that of electricity. When electricity was first introduced it was noticed and

marveled upon, now we don't even notice it … until it is no longer present.

3. Application to humanoid robots

3.1. The concept of "humanoid"

Now I will try to apply this principle to the specific question of the creation of humanoid robots.

The possibility of mechanically reproducing human life is one of the oldest dreams of man. As Prof. Paolo Dario previously said, various cultures throughout the world and throughout the ages have attempted to create a mechanical man either in reality or in fantasy.

Now, for the first time in history, it looks like this dream will soon become true. It appears that man is very close to producing a humanoid. In my opinion, current plans for the creation of humanoids have no relation to the historical dreams and plans of the creation of a mechanical being. These dreams were based on the conviction that man wanted to use his internal resources to fulfill his dream of emulating his creator.

It seems to me that research today has the sole objective of developing technical scientific resources.

In other words, I consider humanoids like mere machines. But what kind of machines?

There are three categories of machines, depending on the types of technical activities that man can perform, which are:

-creation of inanimate tools directly or indirectly guided by human intervention. The classic

example is the industrial robot painting automobiles.

-the next one is the creation of machines that artificially assist organic life. The second example is that of an artificial limb or neuron interface in bioengineering.

-the third is the creation of symbolic machines. This is a very classical concept.

Language is the artificial device that is necessary for a dialogue between two people. This is the most basic example, but perhaps the most important. A computer is also a symbolic machine, as is a book. They are symbolic tools.

3.2. The humanoid as a symbolic machine

The thesis, thus, is that humanoids are called to be the perfect symbolic machines. How and why? First I stated that the aim of technology was not limited to this or that specific need, but it was open to the whole sphere of reality. So, may be possible to dream (not only dream, but also to produce) a machine not limited by a specific function, but able "to do what man can do". In fact, every symbolic machine is unlimited in its "species", because, corresponding to the human symbolic capacity, it will have an indeterminate range of expression: a book is capable of saying everything that man can say, a computer can contain any information that man can develop.

But while other symbolic machines are limited by the nature of their significant capacity, humanoids will be machines capable of reproducing the complete symbolic spectrum of human beings, including all the aspects of the primordial symbolic device. Such a device is human language. Language includes not only oral language, but also body language. The body, in fact, is the primary symbolic expression that human beings have at their disposal.

3.3. Ethical dimension of humanoids

The symbolic character of human acts points to the archetypal dimension of which the acts are symbols. Every symbol requires an archetype.

The basic idea is that reproducing human symbolic functions is not reproducing human beings that are the archetype. 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 can not 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 of man. The symbolic capacity is not in the material condition of the language which a humanoid can reproduce, but specifically in the nexus with the significate archetype (free will) which is signified by the symbol.

All these points can be reassumed in the following scheme. I will illustrate both a symbolic use of humanoids and a non symbolic one.

When I listen to a musical composition played by a humanoid, it is through this material element that I am in dialogue with the composer and the engineer. In this case a humanoid is another element in the material realm of the work of art.

An example of an activity that can not be reduced to space and time coordinates is that of a caress. A caress is not simply a sophisticated movement of a hand accompanied by another sophisticated facial movement. A caress is a way of expressing love, and when the recipient is a human person then the act is not duplicable. A caress can not be repeated in exactly the same way, and a human being can receive human caresses forever as they are manifestations of love. How long would a human continue to appreciate a humanoid caress even though the movement is mechanically perfect?

The condition of the symbolic capacity of the humanoid is that it presupposes human free

will.

The humanoid then, is the most sophisticated thinking machine able to assist human beings in manifesting themselves, and this is ethically very good, as it suppose a radical increment of human symbolic capacity; humanoids will develop a lot of activities in order to increase the human quality of life and human intersubjectivity. But humanoids can never, and I repeat never, substitute human beings. If you use a humanoid to substitute a specific human action, which has its genesis in free will, then it is ethically incorrect. It is incorrect because the symbolical foundation in free will is lacking.

To conclude, everything that an anthropoid can perform is an extension of the human brain's capacity to support human relationships. When you look at the Sistine Chapel you are in dialogue with Michelangelo. When you shake the hand of a humanoid you are in contact with its creator, the engineer.