On the influence of symbols and myths in the responsibility ascription problem in roboethics - A roboticist's perspective

Pierre-André Mudry and Sarah Degallier and Aude Billard

***Abstract*— Because of the increasing developments of humanoid robots, humans and robots are going to interact more and more often in the near future. Thus, the need for a well-defined ethical framework in which these interactions will take place is very acute. In this article, we will show why responsibility ascription is a key concept to understand today's and tomorrow's ethical issues related to human-robot interactions. By analyzing how the myths surrounding the figure of the robot in western societies have been built through centuries, we will be able to demonstrate that the question of responsibility ascription is biased in the sense that it assigns to autonomous robots a role that should be devoted to humans.**

I. Introduction

The development of some technologies, such as genetics or nuclear technology, has brought with it the specter of dangers so deadly that the ethical questions raised needed to be tackled for these new disciplines to survive. Robotics, an expanding science at the crossroads of several branches, is also arriving at a stage where its integration and impact on everyone's lives raises ethical questions like "Who is responsible for the actions of a robot" or "What could be the consequences of augmenting human capabilities by robotic means" ? To answer these interrogations, a new branch of ethics has been developing in recent years.

A. The birth of roboethics

Known as roboethics, this branch shares a number of features with computer ethics but is also at the junction of several disciplines [1] : artificial intelligence, philosophy, ethics, theology, biology, physiology, cognitive science, neuroscience, law, sociology, psychology and industrial design. Although it covers a wide domain, roboethics has not been considered as a field in itself until recently, which explains why this subject is specifically covered by only a few periodic publications.

However, since the first *Symposium on Roboethics*, held in Sanremo in 2004, the domain has become very dynamic. This meeting, which gathered philosophers, jurists, sociologists, anthropologists and roboticians can thus be considered as a real catalyst for the development of studies on the subject. Since then, many conferences have taken place and have led notably to the *Roboethics Roadmap* [1], a document that highlights the different ethical questions related to robotics.

Pierre-André Mudry and Sarah Degallier are with the Biologically Inspired Robotics Group, École Polytechnique Fédérale, Lausanne, Switzerland pierre-andre.mudry@epfl.ch, sarah.degallier@epfl.ch

Aude Billard is with the Learning Algorithms and Systems Laboratory, École Polytechnique Fédérale, Lausanne, Switzerland aude.billard@epfl.ch

As robots will come to interact more and more with naive human users, determining who can be held responsible in case of an accident is likely to be of increasing importance. We will thus focus in this article on the ethical question of ascribing responsibility to robots in the case of human-robot interactions (hereafter HRI).

Basing ourselves on the existing approaches to the responsibility ascription problem, we will discuss how the mythical imagery of robots present in the Western society influences our picturing of ethical problems. In this, we follow the approach taken in the roadmap and we consider that ethical questions touching to the emergence in robots of human functions such as consciousness, free will or emotions as premature, given the limitation of current robot control systems. Instead, we think that there is a strong need for an ethical reflection on the motivations behind humanoid robotics and on the ideal of society it implies, in the sense that scientists should feel responsible of their "creations" and their social implications (see for example [2], [3])

B. Plan

We will proceed as follows : in the first chapter, we will define responsibility from both a philosophical and a juridical point-of-view. This will lead us to consider previous works on the subject which offer different perspectives on how to answer the responsibility ascription problem for robots. Afterwards, we will analyze the influence of history on that question and, during this journey through the centuries, we will show how a specific image of mechanical beings was progressively built in western countries. We will then discuss how this mythology surrounding robots influences the responsibility ascription problem. Before concluding, we will briefly discuss the importance of the scientist's role in responsibility ascription and its relation with the predominant question of the precautionary principle.

II. Machines as responsible agents

As argued previously, we will focus on the responsibility ascription problem, which we view as a central question of HRI. To illustrate this problem, let us imagine the following situation (from [4, p. 179]) : an autonomous robot dog, able to modulate its walking gait to adapt to its environment, decides that the most appropriate means of crossing a carpet is running. During its run, it bumps into a child who, subsequently, falls and hurts herself. Who is responsible for this accident ? The robot constructor, because s/he did not include a control modality to let the robot stop when seeing an obstacle? Or is it the parents' responsibility to

evaluate the danger that robots can create for children and act preventively, as they do for all other appliances? These questions are within the scope of the responsibility ascription problem we are going to examine now.

A. The concept of responsibility

For starters, let us define what the term *responsibility* covers and how it should be understood in the context of ethics. For the *Petit dictionnaire d'éthique* (*Small Dictionary of Ethics*) [5], the definition of responsibility contains three main terms : a person to ascribe it to, a domain where it can be applied and an instance where it should be assumed (like a tribunal, a conscience or people concerned by a decision). Thanks to the notion of responsibility,

> "somebody becomes a law subject or a moral subject [...] that must face his acts and their consequences and that may be punished or rewarded [...] with social esteem, scorn or moral respect."[5, our translation]

In addition to this moral responsibility, there exists a juridical responsibility (including liability) which relates to one's capacity to be accountable for one's acts. The law considers also cases where a reduction of this responsibility is possible, such as when suffering from mental illness, actions performed under duress or when in a state of distress. It is thus easy to conceive that determining someone's responsibility is not straightforward, even in daily human-human interactions.

Moreover, from a legal perspective, responsibility implies that the agent has a conscious control of his/her behavior and of its consequences (see Fischer and Ravizza [6]). This implies that an agent, such as a robot, could be held responsible if, and only if, it can be demonstrated that the agent was free to take a decision and was able to choose among alternatives to this decision. When machines are involved in accidents, the law traditionally considers that the person who controls the machine is responsible, as long as the machine is functioning according to its specifications. When this is not the case, the maker of the machine assumes this role. It is thus possible to transfer responsibility from one agent to another and this is what the manufacturer of a product does through the user manual.

Keeping the above in mind, we will now examine the responsibility ascription problem when it comes to robots capable of adapting and learning new skills.

B. The responsibility ascription problem

For Matthias [4], the notion of control is primordial in responsibility ascription : as soon as the control on a machine is diminished, so is the responsibility of the controller. He takes as an example the hypothetical case of a robot sent to Mars for which there is a twenty-minute lag between the moment an order is given and its execution. Who bears the responsibility when the robot falls into a hole or can not be recovered because visibility was bad ? Is it the technician who was controlling the robot ? The answer given in his article is simply that nobody is responsible in that case and that is the reason why insurances exist, *i.e.* for sharing the costs of an accident among a group of people when nobody can be held responsible.

What happens then, when machines are capable of even a small amount of autonomy (in the robotics sense, *i.e.* independence of control) ? For Matthias, this depends on how the machines acquired this autonomy. When they first appeared, robots were programmed using "imperative" languages. All of their operations were then strictly sequenced and chosen carefully by the programmer. If an error occurred in the code, in almost every case (except for hardware failures), the programmer was sole responsible for this error. However, this kind of programming method does not allow unforeseeable situations to be dealt with easily, and thus newer methods were developed to compensate for the rigidity of the traditional programming techniques.

Neural networks, genetic algorithms and expert systems enabled the development of autonomous agents possessing a certain amount of flexibility and adaptability to their environment : with them, it became possible for these agents to acquire experience by themselves during their work. However, this flexibility comes with a price : it becomes almost impossible to test exhaustively every behavior of the system. In fact, how could all cases be examined when the purpose of such systems is to adapt to every situation, notably when they are not anticipated ? Autonomy implies a loss of control on the programmer's end and, thus, according to Matthias, a lessening of the programmer's responsibility. A responsibility gap is thus generated by the transformation of the traditional role of the programmer : if in the past the complete specification of the robot's behavior was given, the usage of higher-level control methods (more abstract in a way), where actions are hidden in a "layer of obscurity" [4, p. 182], implies for Matthias that nobody can be held for responsible in case of problem.

C. Critical answer

On the existence of such a "responsibility gap", Marino and Tamburrini [7] retort that the question of control was considered in that case in too wide a frame. For them, Matthias goes too far when he says that robot's creators are freed from their responsibility because they do not have total control over the causal chains implied in the construction of such robots :

> "Here the scope of CR [control requirement] is overstretched [..., CR] is not necessary for responsibility ascription, and the alleged responsibility gap depending on it concerns moral responsibility ascriptions only [...]. Traditional concepts of responsibility ascription exist for these problems and have been routinely applied in the exercise of justice."[7, p. 49]

These two authors propose the application of legal principles to fill this gap that, in their opinion, resembles the responsibility parents have to face towards minor children. Moreover, they explain that the situation is in reality simplified because the robots creators decided themselves to give a robot its independence. The nature of this autonomy as well as its

limitations are thus the result of a deliberate choice, and the robots' creators can not simply reject all responsibility. Note here that this is also the case with manufacturers : they keep a share of the liability, even after a user has read the manual, and can not reject all responsibility.

D. Responsibility ascription

As depicted by Asaro in [8], legal responsibility might be a good entry point to consider responsibility ascription for machines because, in contrast to moral responsibility, the justice system gives a precise framework to the question. Currently, moral responsibility suffers from the lack of a unified theory whereas the law provides more practical answers. This framework might seem restrictive to examine all the implications of the situation, however, observing the fact that legal responsibility can be considered as a subset of moral responsibility so to speak, it already provides an interesting perspective.

In his article, Asaro proposes, in the absence of a better solution, to use juridical principles to study robot ethics. For him, many aspects of the law are thus applicable : as long as robots remain simple, the principles defined for commercial products suffice, whereas with autonomous robots the diminished responsibility principle should be applied. Thus, in cases where responsibility can not be assumed completely, like it is the case with minors or people suffering from mental disabilities, the law considers exceptions and, thus, could also be applied to autonomous robots.

1) Penal law: The author goes further by examining the frame of penal law in which the questions of what is a moral agent and what punishment can be reasonably applied are raised. In order to apply the penal law to judge a given crime, one must find a moral agent to be held responsible otherwise the crime is considered an accident. In the hypothetical case of a robot committing a crime, could the robot be considered sufficiently autonomous to have done the action by itself ? Several authors propose different ideas to answer this question : Solumn [9] proposes the equivalent of a Turing test to determine if the agent can be held responsible in court. Harnard [10] gives a hierarchy of Turing tests and examines their advantages and disadvantages and highlights the most adequate test for robotics. More recently, Kahn *et al.* [11] propose different metrics permitting the evaluation of the resemblance between a robot and a human, which could be used for responsibility ascription. Coleman [12] identifies several characteristics that could be used for building a framework of virtue ethics, whereas Floridi and Sanders in [13] propose the concept of artificial agents for which they are able to separate the agent's morality and the agent's responsibility. However, to our knowledge, the question of how to punish a robot for a crime remains unanswered.

2) Moral agency and robots: Although we agree with Marino and Tamburrini that the law disposes of the necessary tools to cope with current situations in robotics, we also believe that it is important to consider the fact that, with current technology, the question of responsibility ascription is biased. In fact, even if machine learning is a very interesting way of implementing adaptability, robot's behavior is still only the result of the application of an algorithmic process. In other words, even if the computing process of a robot is based on genetic algorithms, neural networks or any other learning method, it still lacks free will.

Indeed, as it was shown by Kim and Sanders in [14], the more autonomous the robots are, the more they are considered as human and thus as responsible. However, robots that we are building today are not able of taking decisions (in the sense of making a free choice) and thus can not be held responsible for their behavior. Thus, when discussing responsibility ascription for robots, it is very important to keep in mind this tendency to anthropomorphize robots and our interactions with them. In the absence of free will, an essential component of moral responsibility, we think that the question of responsibility ascription remains pretty simple : it has to be shared between the user and the robot creator, exactly in the same way as for any machine. Ethically, the creator is responsible for the possibilities that he offers to the user, while the user is responsible for the use he makes of such possibilities.

3) A biased problem: For us, the question of robot responsibility ascription is biased because of the importance of history for the robot figure. To illustrate this, let us examine the following situation: in the case where a real dog causes damage or hurts somebody, the question of the dog's responsibility is never asked. If the dog is considered dangerous, it is put down and the question of the owner's responsibility is then examined.

However, it is fair to consider that a dog is an autonomous agent more highly evolved than anything robotics could produce in the next fifteen years. However, in cases like the one in our example, the question of the animal's responsibility is simply almost never questioned. Indeed, one could expect that the dog, as a living, complex system should be considered more responsible[1] than a pre-programmed robot. What is the fundamental difference that causes an ethical issue to be raised in one case and not in the other ?

To our understanding, the fact that the situation is reversed here can be explained by the fact that robots, although still relatively simple compared to animals, are designed to imitate human beings. Thus specific hopes, beliefs and fears are projected onto them : there is an identification phenomenon (anthropomorphization), which does not occur with animals, such as dogs.

III. The importance of history for the robot figure

Robots, more than any other machine, are still surrounded by an aura of mystery. By tracing the origins of this concomitant uneasiness and attraction people feel towards autonomous machines, we will show that the weight of the myths and symbols related to them is primordial in

[1]It is interesting to note that, in the past, animals were penal subjects and recordings of animal trials do exist [15]. A study of the emergence and disappearance of such a phenomenon may help to understand the evolution and the representation of the notion of responsibility.

roboethics. Indeed we will now show how they play a major role in the definition of the ethical issues related to HRI because they raise fundamental questions concerning the motivations, dangers and implications of creating an artificial human.

A. Antiquity at the origins of the myths

One of the first thoughts that comes to mind when robot figures or myths are mentioned are generally related to the literature or cinematography. Mary Shelley's *Frankenstein*, Fritz Lang's *Metropolis* or Asimov's law of robotics are, among others, well known and relatively recent western allusions to this idea, which all have in common the fact that they transmit stereotypical representations and symbols related to robots. However, the subject is older than these few examples from the last two centuries, and cultures as old as ancient Greece or the Egypt of the pharaohs was already preoccupied by this question.

For instance, one of the first allusions to autonomous artificial beings is found in Homer's Iliad (book XVIII), in which Hephaestus, the crippled god, builds golden servants and golden seats that are able to move by themselves. Hesiod's *Theogony* is even more interesting from the myth point-of-view because it relates to how the legendary Prometheus made, out of clay, a man to whom he brought the secret of fire so he could reproduce. However, because Prometheus stole Zeus' prerogative in the domain of giving life, he was condemned to have his liver eaten everyday by an eagle, a story which gave birth to what would be known as the Prometheus myth.

B. Autonomous statues

In addition to the legendary or mythological figures and machines we just mentioned, more or less complex automata have also been built in Greece and in Egypt during Antiquity. For example[2], statues that could move a part of their body (such as an arm linked to a string) or generate sounds using fire were created. These idols, made by priests, were supposed to be able to answer questions and to transmit the will of the gods. Most of the time, the priest himself asked the questions and "received" the answers by triggering the different mechanisms.

This practice holds a deeper meaning than the simple trick it appears to be. In fact, it comes to light with this ritual that the introduction of a certain form of autonomy (or life), even as simple as a head moving or a hand waving, seems to bring a moral dimension to a statue because the priests could "insert the soul in the effigy by a rite [...] to give it the power to speak and to do right or wrong" [16, p. 15, our translation].

Such religious rites were common and such statues have been found in various places such as Italy, India, China and throughout Africa. We think that the popularity of these rituals, in addition to the fact that they were certainly used by priests to convince credulous people, could play a non-negligible role in understanding the origins of the question of responsibility ascription. In fact, it shows that it is deeply anchored, in many traditions, that autonomy acts as a strong hint of the presence of discernment, a necessary building block for moral responsibility.

[2]The interested reader will find in [16] an interesting and deep historical review of automaton.

C. The Middle Ages

Scientists and alchemists from the Middle Ages wanted to do more than their predecessors by realizing a creature of flesh without a woman. Whilst the idea was not new (one of the first allusions of such a project appears in Aristotle), people from the Middle Ages developed this idea of biological reproduction in which the woman brought the raw matter and the man the soul. This concept of *homunculus* was partly developed by Paracelsus who even gave a recipe for the creation of embryo. The symbolical sense of this experience of the metal embryo lies in the desire to create a being, to play God, as in Gœthe's Faust. It explores the limits between artificial production and natural reproduction. Thus, if natural reproduction allows a share of uncertainty about the characteristics of the child to come, the artificial production of a human being is entirely controlled by the creator. Therefore, he can then try to iteratively generate a better creature based on the previous unperfect attempts.

This recurrence of Prometheus' myth through history is related to the fundamental question of what separates life from death, in other words the definition of life. It is very interesting to note that the questions highlighted by this myth also find a modern echo in bioethics.

D. Dualism and rationality

Rational theories that appeared since the XVI$^{\text{th}}$ century opposed this occultism. With his philosophy, Descartes radically separated body and soul. For him, because the body was only a machine made of pipes and pumps, animals had no consciousness and therefore did not differ much from automata except in their complexity.

This famous "Cartesian Dualism" also postulates that Man possesses something more : he is a machine with a soul. This conception is still somewhat present nowadays and conflicts with positivism, notably when it comes to determining if it is possible for a robot to have a soul or feel pain.

1) Automata in literature: Since the XVII$^{\text{th}}$century, mathematical and technical developments have allowed the creation of increasingly complex mechanisms. This period is marked by the appearance of automata like Vaucanson's duck or Pierre Jacquet-Droz's androids, which were able to imitate complex movements like writing or piano playing. This development of automation has marked the literature. From the Age of Enlightment, it began to interest writers : Rabelais' *Gargantua* occupies his leisure time with small toys that can move by themselves and Cervantès' *Don Quichotte* dreams of flying wooden horses. This era also sees the development of science-fiction : *Cyrano de Bergerac* of Edmond Rostand describes how insect-looking rockets are propelled using gun powder and Ernst Hoffmann's *Le*

Marchand de sable relates the story of a scientist whose daughter, Olympia, is in reality an automaton.

In addition to these few examples, the Prometheus' myth has been incorporated more clearly in a number of works. We will cite here only two of them : Mary Shelley's *Frankenstein* and Carlo Collodi's *Pinocchio*. In the first, the creature is made out of human organs and reborn with fatal consequences for its creator, whereas in Pinocchio, the puppet carved in wood acquires a human status after a journey of initiation. These two examples are very interesting in our opinion because they propose two opposing and complementary visions of the myth. Of course, they do not cover the complexity of the question and we refer the interested reader for example to [17] or [18] for a more complete panorama.

E. The threat of the artificial intelligence

After the development of realistic automata came the era of logical machines, such as Pascal's calculator, that threatened humanity's unique ability to think. Leibniz's hope of realizing algorithmic reasoning was first made a reality in the beginning of the $\mathrm{XIX}^{\mathrm{th}}$ century, at least partially, by Count Stanhope who realized a machine that, using logical notation, could use syllogisms to solve simple probability problems. After him, Charles Babbage built, around 1829, his *Analytical Machine* which, along with the one Jevons built forty years later, served as a model for the first computers. In fact, it is with the description of a method to compute Bernouilli's number with Babbage's machine that Ada Lovelace invented the notion of program. With her, computer science was born.

This new science was then developed thanks to Boole's algebra and the invention of electronics components like the relay that enabled the construction of increasingly complex machines from the beginning of the $\mathrm{XX}^{\mathrm{th}}$ century. It was in this context that Alan Turing formalized the concepts of algorithm and calculability. His article, *Computing Machinery and Intelligence* [19] was one of the first to ask the question of artificial intelligence.

F. New medias to perpetuate the myth

The $\mathrm{XX}^{\mathrm{th}}$ also saw the development of new medias that helped diffusing the myths surrounding robots. Science-fiction novels made robots one of their favorite themes and, to cite only two of them, authors like Isaac Asimov with his cycle of the robots or William Gibson's *Neuromancer* propose two different visions of a possible future with robots. Moreover, medias forms like the cinema have also been used since the 1920's to aliment the symbolical image surrounding robotics [20]. In parallel, robot companions (Nono in Ulysse 31 for instance) or super heroes (like Astroboy, Goldorak) present for example in japanese cartoons rise the ideal of a perfect, protective friend, made out of bolts but also able of feelings.

G. The two faces of the robots

In summary, in this section we have tried to emphasize two things, namely, (1) autonomy inevitably evokes life and existence, in the sense that we tend to consider as alive a machine able to behave in an autonomous way, and (2) this impression of life is both attractive and frightening, indeed, while the possibility of creating the robot of our dreams is attractive, our inability to forecast the behavior of such a creature is repulsive and make us fear this power of creation.

IV. Human responsibility

Thanks to this succinct historical exploration, we saw how a symbolic figure of the humanoid robot was built through centuries in Western countries. This mythical representation influences the way we perceive today's and tomorrow's robots. Because of this, we argue here that the mythical figure of the robot is accountable for a non negligible share of the arguments used in the discourse on responsibility ascription. Terms like autonomy, intelligence or decisions are often used for robots, however it must be kept in mind that they do not have the same meaning when used for humans. Moreover, because the mythology and symbolism related to the autonomous automata was historically and socially built, it should also noted that if the cultural references differ sufficiently, so will the the perception of robots as it was shown by Kitano [21] or Bartneck [22].

A. A human perspective

The discourses we have examined consider that autonomous robots, *i.e.* robots controlled using learning methods, are not to be taken only as plain machines but have to be held responsible for their behaviors because those latter cannot be forecasted (and thus the programmer is not responsible). Then, as robots cannot be punished (because they have no feelings and no notion of their own existence), the question is to know who to blame.

It is worth here to discuss the notion of emergent behavior : in most algorithms, the robot "chooses" its behavior relatively to rules and optimality criteria that have been pre-specified by the programmer. Thus, the outcome of the process is completely predefined mathematically[3], even if it has not been explicitly implemented and if it is not known *a priori* (only because we do not have an explicit knowledge of the solutions of the optimization problem). For us, the responsibility lies in the limitation of the possible behaviors (of the space of solutions of the problem), not in the choice of a particular behavior – and this limitation is totally up to the programmer. To illustrate this, let us take again the example of section II of an autonomous dog which selects the behavior "running" and bumps into a child. The important point here is not that it has selected this behavior, but rather that this potentially dangerous behavior was present in the space of solutions.

Having this in mind, we propose here another perspective : in our opinion, robots are, and will remain in the foreseeable future, plain machines without the necessary moral principles that allow one to ascribe them responsibility. This said, we still consider that the ethical question of responsibility in

[3]We postulate here that the algorithm used converges to the optimal solution

human-robot interactions remains primordial. However, we believe that for the robots existing today and in the near future, it can be addressed only from the human side of the interaction even if mythology tends to indicate differently.

For us, the issue of responsibility in robotics relates only to the robots' creators and users. If responsibility ascription for the user does not pose a particular problem from the ethical point-of-view at this stage, the question of the ethical position of the robot creator is more complex. In fact, it raises the question of the scientist's responsibility, not only as a representant of a company or academy but also as a member of a community.

B. The precautionary principle

Talking of the scientist's responsibility also raises the question of the interaction of technology and ethics, *i.e.* what should humanity do with the powers it possesses ? Hans Jonas, in his philosophical work on responsibility [23] analyzes why since Man acquired technologies powerful enough to destroy humanity, the scientist can not simply stipulate that his interventions will be necessarily counter-balanced by nature. He refuses the idea that technology will always find solutions to the problems it creates. This questioning is at the origin of the so-called *precautionary principle* which appears as a sort of answer on how to tackle the uncertain and unforeseeable consequences of human actions. In this framework, questions such as "Why building humanoid robots ?", "What kind of technology do we want to develop ?", "Are robots useful to society ? Why ?" are raised. Thus, it is important to consider that technics and science in general are not neutral but are influenced by the society.

V. Conclusion

The message of this article is that the ethical question of responsibility ascription is strongly related, in Western countries, to the mythical image of the robot. Indeed, there is a gap between what robots effectively are and can do and what people expect from them, basing them on their cultural representation of the robot rather than on existing robots.

It is for sure pertinent to forecast a future with intelligent robots and to question the related ethical questions, notably in terms of our motivations as roboticists and of our inherent representation of a desirable society. However, if we want to address the problem of roboethics in the upcoming years, it is also important to consider pragmatically robots as they are now and to get rid off all the mythological aura that still surrounds humanoid robots today. Indeed, through the question of responsibility, we have examined how the small capabilities that today's humanoid robots possess already sketches complications for a future where completely autonomous robots will be reality. It forecasts a huge development of roboethics to control the mutations that, like every technological revolution do, robotics will bring.

To complete our study, it would be interesting to analyze how an ethic built on a different cultural background would answer the responsibility ascription problem, an approach that would allow us to analyze what part of the results presented would be still valid in a different environment.

VI. Acknowledgments

We are grateful to Daniela Cerqui, Marc Attalah and Antonin Wiser for their insightful comments on the subject.

References

[1] G. Veruggio and F. Operto, "Roboethics: a Bottom-up Interdisciplinary Discourse in the Field of Applied Ethics in Robotics," *International Review of Information Ethics*, vol. 6, pp. 3–8, 2006.

[2] D. Cerqui Ducret and K. Warwick, "Re-designing humankind. The rise of cyborgs: a desirable goal?" *Designing: From philosophy to ethics, From engineering to architecture*, pp. 177–187, 2006.

[3] J. Wagner, D. Cannon, and H. V. der Loos, "Cross-cultural considerations in establishing roboethics for neuro-robot applications," in *Proceedings of the 9th Intern. Conference on Rehabilitation Robotics (ICORR'2005)*, July 2005, pp. 1–6.

[4] A. Matthias, "The responsibility gap: Ascribing responsibility for the actions of learning automata," *Ethics and Information Technology*, vol. 6, no. 3, pp. 175–183, 2004.

[5] O. Höffe, Ed., *Petit Dictionnaire d'Éthique*. Fribourg, Suisse: Éditions Universitaires Fribourg et Éditions du Cerf Paris, 1993.

[6] J. Fischer and M. Ravizza, *Responsibility and Control. A Theory of Moral Responsibility*. Cambridge: Cambridge University Press, 1998.

[7] D. Marino and G. Tamburrini, "Learning Robots and Human Responsability," *International Review of Information Ethics*, vol. 6, pp. 46–50, 2006.

[8] P. M. Asaro, "Robots and Responsibility from a Legal Perspective," in *Proc. of the IEEE 2007 International Conference on Robotics and Automation (ICRA'07)*, Rome, April 2007.

[9] L. Solumn, "Legal Personhood for Artificial Intelligences," *North Carolina Law Review*, pp. 1231–1287, April 1992.

[10] S. Harnad, "Minds, Machines and Turing," *Journal of Logic, Language and Information*, vol. 9, no. 4, pp. 425–445, 2000.

[11] P. H. Kahn, H. Ishiguro, B. Friedman, T. Kanda, N. G. Freier, R. L. Severson, and J. Miller, "What is a Human? – Toward Psychological Benchmarks in the Field of Human-Robot Interaction," *Interaction Studies*, vol. 8, no. 3, October 2007, to appear.

[12] K. G. Coleman, "Android arete : Toward a virtue ethic for computational agents," *Ethics and Information Technology*, no. 3, pp. 247–265, 2001.

[13] L. Floridi and J. Sanders, "On the Morality of Artificial Agents," *Minds and Machines*, vol. 14, no. 3, 2004.

[14] T. Kim and P. Hinds, "Who Should I Blame ? Effects of Autonomy and Transparency on Attributions in Human-Robot Interaction," in *Proc. of the International Symposium on Robot and Human Interactive Communication (RO-MAN'06)*, September 2006, pp. 80–85.

[15] E. Cohen, "Law, folklore and animal lore," *Past and Present*, no. 110, pp. 6–37, February 1986.

[16] J. Cohen, *Les robots humains dans le mythe et dans la science*. Paris: Librairie Philosophique J. Vrin, 1968.

[17] A. Chapuis, *Les automates dans les oeuvres d'imagination*. Neuchâtel: Éditions du Griffon, 1947.

[18] D. Ichbiah, *Genèse d'un peuple artificiel, Robots*. Genève: Editions Minerva, 2005.

[19] A. M. Turing, "Computing machinery and intelligence," *Mind*, vol. 49, 1950.

[20] W. Strauwen, "Le mécanoïde et l'androïde : deux faces du mythe futuriste dans le cinéma d'avant-garde des années vingt," *Cinémas*, vol. 12, no. 3, 2002.

[21] N. Kitano, "Roboethics; a Comparative Analysis of Social Acceptance of Robots Between the West and Japan," *The Waseda Journal of Social Sciences*, vol. 6, 2005.

[22] C. Bartneck, T. Nomura, T. Kanda, T. Suzuki, and K. Kennsuke, "Cultural differences in attitudes towards robots," in *Proceedings of the AISB Convention: Symposium on Robot Companions: Hard Problems And Open Challenges In Human-Robot Interaction*, Hatfield, 2005.

[23] H. Jonas, *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*. Chicago: The University of Chicago Press, 1984.