ICRES 2018: International Conference on Robot
Ethics and Standards, New York, USA, 20-21 August 2018.
https://doi.org/10.13180/icres.2018.20-21.08.p04

FEAR OF ROBOTS: A ROBOTICIST PERSPECTIVE

RODOLPHE GÉLIN
Research Director,
Softbank Robotics, France

Among the general public, opinions diverge on their feelings about robots. Some people would love to have robots at home and are pushing the roboticists to work harder and faster, while others are much more reluctant and see robots as a threat to their job, to their freedom and even to humanity in general. The present paper addresses these topics from a roboticist perspective.

1. Fears about robotics

As it sometimes happens with scientific and technological endeavor, robotic development has been causing some adverse reactions in the public, mainly in the West. On the contrary, in Japan, the a priori opinion of people about robotics is rather positive, probably due to a distinct cultural framework where the boundaries dividing animate from presumably non-animate entities are not as rigid as they are in the West and also because, after WW2, robots have become heroes of the popular Japanese culture, mainly due to mangas and cartoons causing the image of friendly robots to populate the collective imagination.

Facing a demographic problem, since the 1980's, with an increasingly rate of aging population, low birth indexes and consequent shortage of labor power, Japan has been steadily promoting robotic research as a solution to many societal problems.

On the other hand, in occidental countries, fiction has always been much less positive about robots with their potential negative aspects grounding in ancestral fears associated to figures as Golem or Frankenstein's creature and fictioned and anticipated by writers and filmmakers long before they really existed as a technological artifact. In the same way that Steven Spiel-berg traumatized generations of swimmers with Jaws (when the probability of a shark attack is much smaller than being killed by a falling coconut or by a car when crossing the road), James Cameron and his first Terminator instilled the doubt about the long-term goal of the robots and their artificial intelligence [1]. In spite of the "nice" Terminator of the sequels, Terminator represents Godwin's law of robotics: discussion about robotics often converges to it and "terminates" the communication. On its hand, the science-fiction novel Robocalypse describes a self-aware artificial intelligence system that plans the elimination of human civilization. Steven Spielberg had envisaged adapting the book but postponed it. If this movie had the same impact on audiences "Jaws" did, it would probably be hard for normal consumers to accept having robots at home.

But thrill and anxiety creation is part of the job of science-fiction authors. It should not be the job of scientists. Nevertheless, three luminaries in science and technology have recently asserted terrible statements about artificial intelligence (often associated with robots). Professor Stephen Hawking said that efforts to create thinking machines would pose a threat to our very existence: "The development of full artificial intelligence could spell the end of the human race" while Elon Musk warned that artificial intelligence is a "fundamental existential risk for human civilization ". On his hand, Bill Gates wrote: "First the machines will do a lot of jobs for us and

not be super intelligent. That should be positive if we manage it well. A few decades after that though the intelligence is strong enough to be a concern. I agree with Elon Musk and some others on this and don't understand why some people are not concerned." Despite the fact that Professor Hawkins was an astrophysicist and not an expert in AI (but an early adopter of it with the machine that allowed him to speak and write), his position had a very strong negative impact on the general feeling about these new technological tools such as machine learning, deep learning, reinforcement learning etc. The position of Elon Musk and, particularly, of Bill Gates is more serious, since they are close to the people who drive re-search in AI. By displaying this concern, it means that even they see this evolution as doomed. Elon Musk had, however, a more constructive standpoint co-founding OpenAI, an organization "discovering and enacting the path to safe artificial general intelligence". This is, in fact, the right way for a scientist to act: not just claiming that technological development is dangerous but preventing risks. AI researchers look sometimes like a car manufacturer saying: "Have you seen how fast is my latest car? If you hit a wall at full speed, nobody will be able to recognize your body". No car manufacturer would speak that way. No responsible AI researcher should say that what he is developing is uncontrollable.

As we can see from the statements of those three prominent scientific figures, the fear of AI and robots taking over humanity is a fact. It is generally based upon the hypothesis that, by becoming very intelligent, robots will turn to be conscious and being conscious and intelligent, they will refuse to obey humans and then strive to destroy them. This reasoning implicitly tends to say that as soon as you are very intelligent, you disobey and you can cause harm to humanity. That is a very sad understanding of the concept of intelligence.

In the shorter term, the current and immediate application scenarios for robots raise two main fears: unemployment caused by professional applications of robots and dehumanization of relationships because of companion robots. As roboticists, we should not deny these risks. Of course, robots have already replaced workers in factories, mostly for tedious and repetitive tasks and in the future, robots, and mainly AI, will be able to execute more tasks that were previously assigned to humans. According to previsions, in the near future, 30% to 80% of the existing jobs will be performed by AI and robots. But what is the responsibility of roboticists in this?

Researchers keep developing new skills for robots that make them capable of performing more complex tasks replacing an increasing number of people in different functions. The first natural reaction would be to stop robotic research to avoid this situation. But stopping the progress in science is not a natural disposition for the researcher. A better reaction would be to understand what technology will be able to do in the future and what it won't be able to do. Education of future generations should focus on the areas where humans will remain irreplaceable: managing unexpected situations, bringing physical and psychological assistance, configuring and setting robotic systems in new occupations, fixing robots and many other jobs that we can’t even imagine today.

To accompany the development of professional applications of robotics, the role of roboticists is indirect: they can just provide information and assist political stakeholders to prepare the society to the arrival of robots. In the field of companion robots, they can be more active to prevent the risks that can generate fears in the general public. This use case is mainly illustrated by robotic assistance to elderly people. The robot in this application shares the home environment with an elderly person, who is alone and usually in a situation of loss of autonomy. The robot is a way to ensure the safety of the person and a certain level of comfort with-out having to be at a retirement home. The worries generated by this type of application are numerous and have been presented in [2], but risks exist, namely (i) hacking- a malicious person

removes control of the robot; (ii) over-attachment (the elderly person considers robot as her only friend and cuts herself off from all human companionship), (iii) manipulation (using the emotional bond that has been created with the older person, application developers man-age to make it act against the person's own interests). As described in [2], technical solutions exist against the two first kinds of risks and the robotic developer should implement them, but against the third one, only the ethics of the developer can prevent technology from progressing in this direction.

2. Temptation of overpromising

If science fiction can generate fears about robots there are also optimistic extrapolations that do not give a correct idea about the current technological state of the art. Let's mention for in-stance Sophia, the beautiful humanoid face developed by Hanson Robotics that seems able to have a very clever discussion with a human being. People are mostly impressed by the conversation Sophia holds when in fact the expertise of Hanson is rather in creating and animating a human-like face. The dialog is probably scripted and prepared in advance. The embedded artificial intelligence is absolutely not capable of having existential questions by itself as it is sometimes demonstrated on Sophia's famous videos. The other very well-known ex-ample is Atlas, from Boston Dynamics. Atlas is an impressive biped robot that can walk on uneven terrain, resist strong physical attacks and even do somersaults. For a robotic re-searcher, the videos of Boston Dynamics are always amazing. Our first reaction, as roboticists watching these videos, is to think that we can stop our research on biped locomotion and move into shark breeding or swimming suits sales. And the first reaction of the large audience is to think: "Here is it: Terminator exists". If robotic researchers can step back and real-ize that this is just a video depicting a very expensive prototype far away from what it is presently possible for mass production, the general public has rarely the opportunity to have in-formation about the truth behind these impressive videos.

In the field of companion robots, researchers are working, for instance, on emotion detection as it will be very useful for this type of robots being capable of adapting its behavior to the mood of its user. This is the birth of “the emotional robot”, as the CEO of Softbank introduced Pepper in 2014. Based on this emotional input, the robot will be able to better take care of an elderly person, detecting her anxiety, her sadness or her fatigue. Emotion detection is a promising area of research with first good results but, nowadays, it is still far away to work properly in realistic situations.

Let us say that roboticists like to embellish the performances of their creatures and video clips are the best friends of roboticists: it is a good way to demonstrate the ability of the robot with-out giving all the details of the experiment. Sometimes the experiment is not reproducible and the video clip on YouTube is the 100th test that finally succeeded. Boston Dynamics likes to present the "making of" where you can see Atlas falling after one of its impressive jumps. The video does not always show the complete experimental setup: the robot is connected to a server with unlimited computation power or use several external sensors to localize the ping pong ball. A demonstration of dialog can be completely fake: the human user and the robot are just playing a scripted scenario. Asimo, the wonderful humanoid of Honda, demonstrates incredible abilities (hoping, opening a bottle, climbing stairs) but is a priceless prototype con-trolled by a team of engineers. At SoftBank Robotics, we made a nice video of a NAO robot climbing a ladder, but we did not mention that the motion was completely scripted, NAO did not even see the ladder before starting the climbing behavior.

There are good reasons for roboticists to take some liberties with the exact scientific description of an experiment. First, videos often target the general public aiming to give an idea of what will be possible in the near future. They are supposed to be short, fast and impressive. A long introduction on the experimental setup would be boring. Details are generally available

in scientific publications accompanying the video. These impressive demonstrations are often a presentation of the vision of the research laboratory or the robotics company. They present the robot as it is anticipated more than the robot as it really is at the moment. This vision is also a way to prevent some of the fears that robots raise in the general audience. A robot that considers my emotion to adapt its behavior is less threatening than an automaton running blindly the same sequence. The last reason for overpromising we can mention is money. When researchers are looking for money for their research, they must make investors dream by promising features that were never seen before. Otherwise, they will not pay for something that YouTube already knows. It becomes, for instance, more difficult to get money for re-search on biped locomotion every time Boston Dynamics publishes a new video of Atlas.

3. Integrating robot companions in society: A personal experience

I had the chance to lead the Romeo project that was planning to develop a tall humanoid ro-bot to assist elderly people. This biped robot was supposed to walk all around the apartment to check if everything was fine, to monitor the mood of the person and call for help in case of trouble, to make some conversation and to learn the habits of the person in order to detect unusual behavior. This project was quite ambitious, but the state of the art and the quality of the partnership made the objective believable enough to be funded by the French govern-ment. At the end of the Romeo2 project (9 years later), we had developed a 1,40 high biped robot, cuter than many others but walking very slowly and far less stable than the Atlas robot. Anyway, the research on Romeo allowed Aldebaran (later SoftBank Robotics) to develop Pepper, the 1,20 m high wheeled robot. With Pepper, the Approche association (gathering about 20 rehabilitation centers working on the promotion of new technologies for elderly and handicapped people) ran experiments with 24 patients, in two rehabilitation centers, 24 hours a day, during a whole week. The robot shared the room of the patient during 7 days, providing simple services (agenda, weather forecast, music player, video player, time and date of the day,…). The result of the experiments was quite positive: elderly people accepted this robotic assistance in their private area and proposed new features that could be implemented on the robot. Nevertheless, the implemented features were far away from what we had envisaged 9 years earlier. And we faced very practical problems: in one center, the robot never succeeded in accessing the Internet, obstacle avoidance, based upon a laser range finder, was unable to detect some elevated obstacles (bed, chair,…), detection of people in a wheelchair was not robust because Pepper is mainly designed to interact with standing people. During the project, academic and industrial

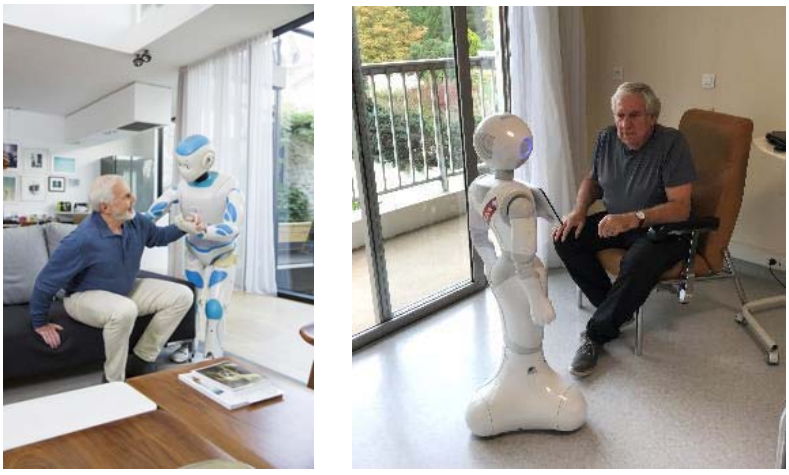

Figure 1. Romeo project: Vision (left), real experimentation (right)

partners could demonstrate much more advanced features but their integration, all together, on a single robot, running autonomously was extremely complex and could not be realized. It does not mean we lied when we proposed the project and when we described what we thought we

could achieve. It just means we were too optimistic, and reality is a ruthless judge for robotic systems.

The Darpa Robotics Challenge (DRC) was the opportunity for the general public to discover the real state of the art of humanoid robots. This challenge sought to address the problem of human-supervised robotic technology for disaster-response operations. Famous videos showed how difficult it was, in 2015, for humanoid robots to walk in the sand, to get out from a car, to climb a ladder or just to open a door.

Figure 2. Robots climbing ladders: Schaft's S1 winner of the DRC (left), SBR's NAO (right)

4. Conclusion

Roboticists have good reasons to display positive and embellished images of their robots. It is a convincing way to show that science progresses, to explain where they want to go with their developments and to reassure people who are afraid of these artificial creatures that this technology is meant to contribute to the well-being of humanity and the good of society Making all this information available to the general public is necessary on the other hand showcasing robotic advances contributes to raise the spirit of all the robotic community and to encourage their efforts. The whole robotics community is proud when an impressive demo of one of their members pop up on the Internet.

Nevertheless, we must be careful of possible misunderstandings. When Atlas is pushed with a hockey stick to demonstrate its stability, it can be interpreted as a defensive gesture against an aggressive huge robot. When a humanoid robot talks about the meaning of life, people can believe that AI does philosophy when it is just a scripted sentence. The risk of this kind of communication is to make believe that all this technology is available and will be soon in our homes. This could be considered as great news for some people, as a threat for some others and certainly a disappointment for most of them. Many people, visiting the website of the Ro-meo project, asked us if they could buy a Romeo for their elderly parents. We had to explain that, unfortunately, the technology was not yet ready to be used "in real life".

Roboticists must find a trade-off between making people understand what the positive future of robotics is and explaining how long the road to reach this future will be. Public benchmarks, like the Darpa Challenge or the RoboCup@home, should become mandatory to give a realis-tic status of the robotic research. But, even in their present state, there are so many wonderful things that robots will be able to do to assist people that it is not necessary to give false hopes or cause excessive fears.

References

1. M. Chita-Tegmark, *Terminator Robots and AI Risk.* Retrieved August 16, 2017.
2. M. I. A. Ferreira et al. A world with robots. *A world with robots*, International conference on robot ethics: ICRE. Vol. 84. 2015.