Meta-Research and Meta-Robotics*

Hans-Georg Stork
(h-gATcikon.de)

Avant-propos

I must begin with an apology. I am not a roboticist. That is to say, I have never been directly involved in activities that are in one way or another linked to the design and building of the kind of machines that this summer school is about. And I have not made the slightest contribution to the advancement of the science underlying the design and building of these machines. So I am not a robot (let alone rocket) scientist either. (I am aware of the ambiguity.)

So be prepared for a largely non-technical, non scientific interlude (as announced). It won't be quite as non-technical and certainly not as literary as Karel Čapek's famous 1920 play *Rosumovi Univerzální Roboti* (*Rossum's Universal Robots*)[1] which allegedly introduced not only the term robot to the world of Science Fiction but also made its robots reason in a real, noisy and dynamically changing world. In fact, the Czech word rozum, if I am not mis-informed, means just that: reason or common sense. Thus R.U.R. predates the "ReaRW task" by nearly one century and represents a fitting genius loci for this summer school.

As for me there is but one justification for speaking to you that I can claim: I have over the last eight years before my departure from the European civil service (two-and-a-half years ago) been involved in a kind of meta-research that was indeed strongly related to what you are doing or learning to do.

Introduction

As the term "*meta*" suggests this has been research about your research: finding out what the burning scientific and technical issues are, who is tackling

*Transcript of a lecture given at the Summer School – Reasoning in the Robot World, 2014 Prague, Czech Republic, 29-31 July, 2014. Hosted by the Czech Technical University in Prague, Center for Machine Perception (`http://summerschool2014.ciirc.cvut.cz/`)

[1] `http://en.wikipedia.org/wiki/R.U.R.`

these issues, what feasible approaches are taken, et cetera, and perhaps most importantly, to what end this research should be done and hence, financially supported.

I did this sort of research jointly with a fair number of more or less like-minded colleagues in my capacity as "Research Program Officer" working for the European Commission's Directorate General "Information Society (INFSO)" (which not long after my departure, has been renamed "Communications Networks, Content and Technology (CNECT)"). It is that department which is in charge of financially and otherwise supporting your projects.

Apart from this "*second-order*" research a "Research Program Officer" is engaged in, he or she has a number of more mundane, clerical, and bureaucratic tasks to attend to. As most of you probably know these include the preparation of Calls for Proposals, finding competent peers – not in cahoots with proposers – to assess and rank proposals, negotiating contracts (and associated work plans) with successful proposers, and last but not least, monitoring running projects and conducting periodic reviews. To the best of my knowledge, in the mid-term or perhaps even shortly, these tasks will be outsourced to an agency, especially set up for this purpose.

The meta-research part of our work usually boils down to short texts, called "Work Programmes" which loosely specify the content of research projects competing for European monies. This is why I like to refer to this part of our job as "*programming in the very large*".

But ever since Aristotle wrote his famous treatise on Metaphysics[2], so named because it was the book that came after his Physics, there is another customary meaning of the term "*meta*". It relates any subject to which it is prefixed to that which is or may be beyond that subject.

(We note in passing that there is at least one further use of "*meta*", as for instance in Metamathematics[3] and Metadata[4] ..., where it is formally the same thing X that is about X. However, this self-referential meaning of "*meta*" is of less concern in the present context.)

So we may, for the purpose of this talk, coin the term "meta-robotics". It is still missing on the very long list of "*metas*" on the respective Wikipedia page.[5]

Meta-robotics would probably comprise some of the issues that our meta-research addresses. For example the question: Why should the state fund robotics research and development (and not leave it to the market, the mantra of our times)?

[2] `http://en.wikipedia.org/wiki/Metaphysics_(Aristotle)`

[3] `http://en.wikipedia.org/wiki/Metamathematics`

[4] `http://en.wikipedia.org/wiki/Metadata`

[5] `http://en.wikipedia.org/wiki/Special:PrefixIndex/Meta`

But it comprises much more. There are obvious "*meta*"-questions: What are the potential consequences of this research? What impact will it have on our societies? On the economy, in the small and in the large? What impact will it have on us and our children and grand-children as individuals? And what about the dual use problematique as far as autonomous robots are concerned? Are there limits to what robots can or should do? Will there be a case for holding robots responsible for what they are doing? What about liability? And there are questions of a more journalistic and literary flavour, for instance: Will there be a case for treating robots as sentient beings, endowed with rights and to be treated with respect? Apparently taken seriously by many[6]. (Whether Karel Čapek took them seriously is an open question.)

You see, apart from doing (first-order) research with the aim of creating machines that operate – by virtue of their reasoning capabilities - autonomously and sensibly in the "Real World" a lot of human reason and reasoning may be called for in order to cope with the fruits of our joint and individual ingenuity.

Meta-research and meta-robotics demarcate the territories of this lecture. I will first briefly explain the why's and wherefore's of the funding programme you are benefitting from. I will then try to give you some idea of how this programme came about. Unfortunately, I cannot tell you much about its future as I have been, as mentioned before, since more than two years ago out of my office. Fortunately, this is likely to be a better position to speak on meta-robotics, in the final part of this talk.

Why research funding and for what?

Public funding of scientific research and technological development has a long history. With tongue in cheek we may say that it all started with Adam and Eve although they got severely punished as we know, by the higher powers-that-be for accepting funds from the devil. But of course we don't have to go that far back in time. For our purposes it may suffice to link the emergence of the idea of public S&T funding to the English philosopher and politician Francis Bacon who lived around the turn of the 16th to the 17th century. He too wrote a seminal text, entitled "*The New Atlantis*", describing a society that affords a publicly funded research facility called *Salomon's House* (also known as the *College of the Six Days Works*) "*where specially trained teams of investigators collect data, conduct experiments, and (most importantly from Bacon's point of view) apply the knowledge they gain to produce 'things of use and practice for man's life'* "[7].

[6] `http://en.wikipedia.org/wiki/Intentional_stance`

[7] `http://www.iep.utm.edu/bacon/#SH2b`

Adam Smith and his modern disciples had not yet been on our planet, so purely economic ends were not on Bacon's horizon. In fact, in the preface to his opus magnum *Instauratio Magna* he wrote: „*Lastly, I would address one general admonition to all; that they consider what are the true ends of knowledge, and that they seek it not either for pleasure of the mind, or for contention, or for superiority to others, or for profit, or fame, or power, or any of these inferior things; but for the benefit and use of life; and that they perfect and govern it in charity.*"[8]

Here Bacon may have wanted to counteract moods prevailing in Renaissance England. Yet he is usually credited with coining the famous adage "*knowledge is power*" ("*scientia potestas est*")[9]. He represents like no other that phase in European (and World) history when the giant wheel with the three spokes *political power – economic power – scientific/technical capacity* was set in motion.

We all know where this wheel led us to. In fact, robots may become the apogee of its path. So let us take a long leap forward to the years right after WW2 and meet a man who could be considered a modern successor of Francis Bacon's: Vannevar Bush[10], polymath, science policy advisor to US presidents (FD) Roosevelt and Truman, and administrator of the Manhattan Project that resulted in the first atomic bombs. In Summer 1945 he authored a report to the President under the heading "*Science the endless frontier*". In the letter of transmittal he wrote: „*Science offers a largely unexplored hinterland for the pioneer who has the tools for his task. The rewards of such exploration both for the Nation and the individual are great. Scientific progress is one essential key to our security as a nation, to our better health, to more jobs, to a higher standard of living, and to our cultural progress.*"[11]

In the core document he went on to suggest to set up a National Research Foundation that "*should develop and promote a national policy for scientific research and scientific education, should support basic research in nonprofit organizations, should develop scientific talent in American youth by means of scholarships and fellowships, and should by contract and otherwise support long-range research on military matters.*" The latter as we know, has in the meantime largely been taken over by DARPA, the funding agency of the US military. Vannevar Bush, by the way, also invented a hypothetical machine, called MEMEX[12], which somehow anticipated the later hypertext systems and thus the Worldwide Web.

From Vannevar Bush's proposal to our European RTD programmes, both national and on a European level, it is but a small step. Their rationale is not too different from what I just quoted. And indeed, the overarching objective of Eu-

[8] `http://www.bartleby.com/39/20.html`

[9] `http://en.wikipedia.org/wiki/Scientia_potentia_est`

[10] `http://en.wikipedia.org/wiki/Vannevar_Bush`

[11] `http://www.nsf.gov/od/lpa/nsf50/vbush1945.htm`

[12] `http://en.wikipedia.org/wiki/Memex`

ropean research funding (and presumably of public research funding anywhere in the world) is to boost economic growth through science-based innovation.

It had already been codified in the early treaties of the European Communities, most notably EURATOM. But it took until the early 80's of the 20th century before a full-fledged Europe-wide IT research programme was launched under the name ESPRIT[13]. This was partly in response to similar activities in the US and Japan. Since then we have the well-known successive multiannual "*Framework Programmes on Research and Technology Development (RTD)*" covering many areas of research and development. We are currently in the eighth cycle, somewhat less bureaucratically labeled "*Horizon 2020*"[14].

There are a number of general questions one may ask in relation to spending public money on RTD. First and foremost of course, there is the question "what is worth spending it for". Then: What balance should be kept between basic research, "applied" research, and systems development? What is the role of industry in publicly funded research? (After all, dishing out public monies to private companies could well be perceived as a market-distorting subsidy.)

Different answers to these questions have been given at different times. It would not make sense to go into all of them here and now. Only that much: There is a problem. At least from my perspective these answers – most importantly those given in terms of budget/resource allocation - have somehow emerged from more or less transparent discussions among elected (e.g., committees of the EU parliament) and non-elected bodies (e.g., departments, units of the European Commission). And the closer one gets to the bottom, to defining specific areas and specific issues that ought to be addressed, the less transparent it becomes – at least for non-specialists. The general problem I see is that of legitimacy: of who decides what taxpayers' money should be spent on and according to which criteria.

I believe this is a key problem if we accept that our modern societies are in so many ways shaped by science-based technologies. Given the complex interdependencies between science and society[15] it is a serious problem worth considering if we want to further our democratic ideals.[16] The conclusion may well be that whatever institutions and rules we invent in an attempt to democratise decision-making in complex societies there are limits that cannot be passed. (After all, no referenda have ever been held and no votes have been taken on whether or not we should drive automobiles, fly aeroplanes or use computers.) Given that political and economic players with vested interests can take advan-

[13] `http://cordis.europa.eu/esprit/home.html`

[14] `http://ec.europa.eu/programmes/horizon2020/`

[15] `http://ec.europa.eu/research/science-society`

[16] `http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1793&context=lhapapers`

tage of those limits - a familiar key word in this context is "lobbying" - we ought to be aware of their existence and potential impact.

Fortunately, some awareness of the need for principles guiding the public funding of RTD exists at the highest level of the European Commission. In a 2011 keynote contribution to a special issue (on "*Robotics: War and Peace*") of the journal "*Philosophy and Technology*" Neelie Kroes[17], the Commissioner in charge of the "Digital Agenda" wrote: "*But some questions remain. We cannot and must not curb scientific curiosity but we should ask: are there general principles that might guide public funding of research and the use of its results beyond innovation and competitiveness?*"[18]

In her answer she quotes the famous German playwright Bertolt Brecht and at the same time reminds us of Francis Bacon's *New Atlantis* and *Instauratio Magna*: "*Bertolt Brecht, in 'The Life of Galilei', had the great scientist say: 'I maintain that the only goal of science is to alleviate the drudgery of human life.' Sound advice indeed! We will continue to fund research whose results help create better living conditions for everyone on this planet and research that helps us to better understand ourselves and the world we live in.*"

And she concludes: "*Both go hand in hand — and robots should take their fair share in this ICT landscape*"; prompting me to move closer towards the subject matter of this seminar, at least partly guided by her wisdom.

Why robotics?

Of course, robots have been around for a long time. First and foremost in science fiction stories. (They are still there!) But from the late 1960s onwards also at product assembly lines, in space and on battle fields, to name but a few environments. When I say "robot" I assume that we all have a similar image before our mind's eye: that of an electro-mechanical device, designed and built to help people do jobs that are physically strenuous, potentially dangerous, repetitive and tiring, or simply impossible to do without suitable technical support. To qualify as a robot the device can be stationary or mobile; if stationary it should be able to handle and/or transport physical objects, large or small, heavy or light, depending on the kind of service it is supposed to deliver.

Given the persistent trend in industry to reduce the amount of manual labour in manufacturing goods for example, and keeping in mind the most general objective of research funding, it is easy to see and justify why Robotics was put on the ICT agenda. The specific aims of this research should be equally clear. "Traditional" robots are often nothing but more or less sophisticated

[17] `http://en.wikipedia.org/wiki/Neelie_Kroes`

[18] `http://link.springer.com/journal/13347/24/3/page/1`

machine-tools operating according to preset rules in strictly controlled environments (like an assembly line). To make robots fit for tasks in, say, open environments where remote control is not feasible or desirable, they ought to be endowed with capabilities that we normally find in ourselves but also in animals. In order to sensibly "*perform movements, manipulation, navigation, etc. in a real, noisy and dynamically changing world*" on their own (i.e., autonomously, the *ReaRW task*!) a robot should be able to correctly interpret what is going on in that world (yes, animals can do that). In other words, it should be an exemplar of an artificial "*Cognitive System*"[19] whose reasoning is informed by real-world inputs and results in real-world action.

Enormous sums have been disbursed with the intent to approach this goal. And our European programmes have contributed a substantial share. While topics broadly related to Artificial Intelligence (AI) have been part and parcel of European research programmes ever since they were first launched in the 1980s, Cognitive Systems became prominent as a specific item on the research agenda only in the late 90s when, under the heading Cognitive Vision, a cluster of eight projects was launched in response to a growing demand for more powerful computer vision systems that were able to interpret what they saw and sensibly to act upon it.

From 2002 onwards, this line of funding has been extended to cover both, Cognitive Systems in general and Robotics. It has been firmly established as a key chapter of the 6th and 7th multiannual Framework Programmes (FP6 from 2002-2006 and FP7 from 2007-2013 respectively), and codified in a series of usually biannual Work Programmes that underly the regularly published Calls for Proposals. By the time I left my office the European Commission, under this chapter, had spent more than half a billion Euros on nearly 140 projects and ancillary activities in the areas at issue. [20]

Meta-research on robotics - drafting a robotics research agenda[21]

We had asked the meta-research questions I mentioned at the beginning of my talk, with the understanding that the "first-order" research out there was still far from delivering fully operational systems that would satisfy criteria such as robustness, versatility, reliability, adaptability and last but not least, autonomy

[19] `http://www.vernon.eu/euCognition/definitions.htm`

[20] `http://cordis.europa.eu/fp7/ict/programme/challenge2_en.html`

[21] This section draws on a previous paper of mine: Towards a Scientific Foundation for Engineering Cognitive Systems - A European Research Agenda, its Rationale and Perspectives; in: Biologically Inspired Cognitive Architectures, Volume 1, July 2012, Pages 82–91. (online `http://dx.doi.org/10.1016/j.bica.2012.04.002`, preprint at `http://www.cikon.de/Papers.html`)

(i.e., to be free from outside control). Hence the explicit aim of our programmes became

... to strengthen the scientific foundation for engineering artificial cognitive systems - i.e., artificial systems that perceive and (inter-)act, in accordance with a suitable understanding of their environment;

and, in doing so ...

... to foster technologies that enable a variety of applications involving interaction within "real world" environments pertaining to, for instance, robotics, assistive technologies, and multimodal man-machine interaction.

Among the latter, robotics has undisputedly always been a major focus and most project work is indeed centred on robotic platforms.

A more detailed but sufficiently "liberal" research agenda was developed and at intervals revised after consulting representatives of different disciplines, disciplines that were believed to make relevant contributions to strengthening said scientific foundation. (By the way, this is a case in point illustrating what I alluded to before, regarding "legitimacy".)

For instance, given that "cognition" is first and foremost occuring in the living world one might ask: *What (if anything) do we need to understand about cognition as a biological phenomenon in order to specify, design and build artificial cognitive systems*? In light of the fact that natural cognitive agents (as individuals or species) are (up until now) practically the only entities that are capable of learning through acting on or interacting with complex dynamic environments, it seems evident that the engineering of artificial cognitive systems can be informed by studying natural processes related to cognition and control, including the role of the physical substrates of these processes. So it seemed a good idea to seek input from biologists and in particular, neuroscientists.

On the other hand, aircraft engineers do not draw on ornithology in order to design and build aeroplanes. Ornithology is simply not part of their scientific foundation. Likewise, although mainstream Artificial Intelligence (AI) research was more impressed with man's unique symbolic reasoning and planning capabilities than (for instance) with his gut feelings it managed to yield many interesting and useful results. But little did it contribute to creating the kind of systems aimed at under our programme. (By the same token, modern aeroplanes and even drones do lack some of the most outstanding avian faculties.)

Yet this would certainly not justify excluding traditional and more recent AI disciplines, such as Statistical Learning[22]. So one of the characteristics of our programme was its openness to multi-disciplinarity, inviting computer scientists, engineers, neuroscientists, psychologists, ethologists, mathematicians and possibly more to advise us and to team up in big and not so big projects. It was

[22] `http://en.wikipedia.org/wiki/Statistical_learning_theory`

also entirely agnostic as far as paradigms (e.g., computationalism, connectionism, enactivism) and different approaches to modelling were concerned.

But what about the utilitarian objective? About innovation and new markets? About the famous industry question? No, it has not been neglected. After all, research with a view to supporting engineering must not be confined to an ivory tower. Rather, it should be motivated by and cater to real needs, in line with the strategic goals of (public) European research funding. The FP6 and FP7 Cognitive Systems and Robotics programmes were definitely hospitable to commercial partners providing relevant scenarios in areas such as industrial and service robotics in all sorts of environments, scenarios where methods and solutions could be tested and validated. And if my interpretation is correct of what I hear on the grapevine, industry is given a much bigger part in the current robotics programme (under FP8 = Horizon 2020), hopefully not to the detriment of solving the still unsolved fundamental problems inherent in the "*ReaRW task*".

And hopefully not to the detriment either, of just sheer curiosity, of the desire to understand. Indeed, there is this other side to doing research which has often contributed more to "innovation" than targeted multimillion Euro/Dollar/Yen projects. Moreover, robotics as a science does have the potential of making us better understand our own nature, what "makes us tick" in our worlds, and how we make our worlds. Almost four centuries ago the Italian philosopher (of science) Giambattista Vico[23], regarded by some as one of the early ancestors of modern (radical) constructivism, expressed this in three words: "*verum ipsum factum*", or: "*The criterion and rule of the true is to have made it*."

Commissioner Kroes, in her short note, acknowledges this potential of robotics when she writes (I repeat): "*We will continue to fund research whose results help create better living conditions for everyone on this planet and research that helps us to better understand ourselves and the world we live in.*" So there is hope. You should take her at her word.

Do we need to know how the mind works (to build the ultimate robot) - can we know it?

But we should not get carried away. No, I do not mean with our hope to get more money for feeding our curiosity. I mean: let us not be too optimistic as far as understanding the human condition and the human mind are concerned. Arguing against the Cartesian idea of certain truth as something as clear and distinct as a geometry theorem, Giambattista Vico insists that "*our clear and di-*

[23] `http://plato.stanford.edu/entries/vico/`

stinct idea of the mind cannot be a criterion of the mind itself, still less of other truths. For while the mind perceives itself, it does not make itself."

Can the human mind make itself? Some people (e.g., those known as "transhumanists"[24] and "extropians"[25]) believe the answer is "yes" and postulate a future when human beings can achieve at least mental immortality (catchword: mind upload[26], a modern form of the dualist belief in an immortal soul). Some dream of phantastic scenarios where robots spread the intelligence evolved on our planet Earth to distant worlds in outer space, thus "conquering the universe". Some people seem to see no limits in what nature (of which we are part) can do. Others may dread a future when humans, as in Karol Čapek's play, become obsolete and are supplanted by their own superior creations. Some of you may remember an article published in Wired in early 2000, by Bill Joy, co-founder of Sun Microsystems, entitled "*The future does not need us*"[27], where he gives words to his concern about a somewhat casual view of some "visionaries" who made up a rather gloomy fate for mankind, apparently based on a very peculiar understanding (some may wish to call it misunderstanding) of what it means to be human.

I find such musings rather amusing. Indeed, if we make a mind it will not be in a human body and hence not be a human mind (unless we do it the traditional way that was invented by nature long before we could have had a say in it). Depending on how narrow or broad we take the concept of "mind" to be we may even say that we have already been creating minds galore; minds in different bodies for sure, but minds that greatly surpass our own, as far as "Algorithmic Intelligence" (another AI!) is concerned - but not more. (For example, just behold the laptop computer in front of you.) The "super-humans" are already there but of course they are not human. In fact they are about as super-human as a tractor is super-equus. Their minds are mere "shadows of our minds" (to recall the title of a 1994 book[28] by Roger Penrose[29], but without endorsing his ideas on quantum consciousness).

In this context it is interesting to note that one of the biggest and most expensive European Projects under "Horizon2020", the "Human Brain Project (HBP)"[30], a so called FET "Flagship"[31], is presently (July 2014) causing a ma-

[24] `http://en.wikipedia.org/wiki/Transhumanism`
[25] `http://en.wikipedia.org/wiki/Extropianism`
[26] `http://en.wikipedia.org/wiki/Mind_uploading`
[27] `http://archive.wired.com/wired/archive/8.04/joy.html`
[28] `http://en.wikipedia.org/wiki/Shadows_of_the_Mind`
[29] `http://en.wikipedia.org/wiki/Roger_Penrose`
[30] `https://www.humanbrainproject.eu`
[31] `http://ec.europa.eu/digital-agenda/en/fet-flagships`

jor controversy mainly among neuroscientists[32]. The HBP sets out to simulate the anatomy and physiology of (parts of) the human brain. Its "raison d'être" is (at least) twofold. Firstly, the expectation that this simulation will provide insights into the workings of real brains and thus helps to study brain diseases and to find pertinent remedies. Secondly, the intention to study "neural computation" more closely in order to create effective and more efficient neuromorphic hardware implementations of it.

Many neuroscientists now fear that the second rationale is overly gaining in weight. From my perspective, this is not surprising given that the HBP is funded under a technology and not a biology programme. Designing neuromorphic hardware[33] is certainly a laudable endeavour. But it seems to me that there is a hidden assumption, nurtured by the extropian claims – unproven - that a brain is fully simulable and that its mappings are Turing-realisable, and hence at least in principle replicable through technical artefacts.

One may not have to go as far as Roger Penrose and postulate non-deterministic quantum processes in microstructures of the brain[34], to be more than sceptical about these claims. Whether the mappings effectuated by brains are Turing-realisable is, to the best of my knowledge, simply an open question. (Here we may note in passing that the mappings effectuated by the members of a certain class of artificial neural networks – Analogue Recurrent Neural Networks - are provably super-Turing, a result obtained some 20 years ago by Hava Siegelmann.[35])

And as Anil Seth[36], computational Neuroscientist at the University of Sussex, in a recent (8 July 2014) op-ed article in The Guardian[37], points out: even if more detailed simulations of the brain could be achieved this would "*not inevitably lead to better understanding. Strikingly, we don't fully understand the brain of the tiny worm Caenorhabditis elegans even though it has only 302 neurons and the wiring diagram is known exactly. A perfectly accurate model of the brain may become as difficult to understand as the brain itself, as Jorge Luis Borges long ago noted when describing the tragic uselessness of the perfectly detailed map*." "Understanding", in this context, presumably means being able to falsifiably hypothesise links bet-

[32] http://www.neurofuture.eu/

[33] https://www.humanbrainproject.eu/de/neuromorphic-computing-platform
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3812737/
http://apt.cs.manchester.ac.uk/projects/SpiNNaker/

[34] http://www.quantumconsciousness.org/penrose-hameroff/quantumcomputation.html

[35] http://binds.cs.umass.edu/anna_cp.html
http://binds.cs.umass.edu/papers/1995_Siegelmann_Science.pdf

[36] http://www.sussex.ac.uk/Users/anils/index.html

[37] http://www.theguardian.com/commentisfree/2014/jul/08/
human-brain-project-missed-opportunity-simulating-neuron-activity

ween brain structures, functions and processes on the one hand, and observable behaviour on the other hand.

One also has to bear in mind the limitations inherent in models, regardless of whether digital or analogue, of non-manmade natural phenomena. In fact, discovering limits is sometimes more rewarding than assuming there are none and reaching one dead end after the other. Limits have been discovered in Metamathematics a long time ago, for instance to what the most paradigmatic computational model, the Turing Machine, can do[38]. Physics sets hard limits to what we can do given our and the rest of nature's nature. Of course one may ask: can the world be completely specified in formal, mathematical terms?[39] Again, some may believe the answer is "yes" and may even go one step further, to believing that emulating natural phenomena can fully capture the essence of these phenomena. But we know: simulations and emulations are always based on models which at best are homomorphic, but not isomorphic, images of the real thing.

(This may seem trivial but is often forgotten or ignored. It applies, by the way, also to social interaction between people. Which includes economics, a vast field of social interaction where it is perhaps most often forgotten. Instead there may be a tendency there to adapt the real thing to whatever model is *en vogue*.)

Meta-robotics - ethics

Let us get back down to earth, back from the lofty heights of brains, singularities and flagships, to the lowlands of the electro-mechanical devices called robots. Here is another verbatim quote from Mrs. Kroes's keynote commentary:

"Take for instance the concept of an autonomous machine. This could be a self-controlling road vehicle, which may become a reality sooner rather than later given the current speed of technological advancement. There are also various examples of military autonomous vehicles operating on land, at sea or in the air. Who is responsible for their actions? Who is liable in case of damage? Can it be considered that such machines operate on their own accord? The answer is a firm 'no' . Machines are designed, built and programmed so that they can render services. They are always owned and controlled by people. Machines — no matter how sophisticated — are as 'ethical' as the people who design, build, programme and use them. We humans, jointly and individually, have to take full responsibility for what we are doing, good or bad, constructive or destructive, through our own inventions and creations, to each other and our world at large."

[38] `http://en.wikipedia.org/wiki/Halting_problem`

[39] `http://www.idsia.ch/~pape/papers/pape2011agilong.pdf`

What our Commissioner addresses here are clearly issues pertaining to meta-robotics. What use should robots be put to, who is responsible for what they are doing, and what implications does using them have for the life of individuals, groups and entire societies? Questions that also underly a whole new scholarly debate on "*Robot-Ethics*". In fact Robot-Ethics has been the dominant theme of an EU supported action called ETHICBOTS[40]. And in October 2013 euCognition, another EU supported network of researchers interested in Cognitive Systems, organised a meeting solely dedicated to "*Social and Ethical Aspects of Cognitive Systems*"[41] including of course, robotics.

The Commissioner strongly denies endowing machines with any kind of responsibility. I can only agree. I would make it even more explicit and submit that man-made machines are categorically different from natural living, feeling, and thinking beings. The more we fancy machines to be human-like, ascribing them intentions, desires and beliefs (c.f., Dennett's intentional stance), the higher the risk of us becoming machine-like ourselves. The more we rely on machines to make decisions that only we can justifiably make, the more we deprive ourselves of our authority, independence and our essential human characteristics. Man-made machines – no matter how sophisticated - have no rights and should not be feared; we can switch them off, take them off line or, ultimately, dismantle them. (Joanna Bryson, University of Bath, in her Essay "*Robots should be slaves*", takes a very similar if not identical view.[42])

By the way, the danger inherent in relying on machines to make decisions in our stead not only concerns robots but technical systems in general. For instance, every bureaucrat knows how easy and convenient it is to hide his or her own incompetence, insecurity or ignorance behind the veil of whatever computerised workflow or transaction systems may have been imposed on him or her.

So what should robots be used for? Of course, they should do all the nice things that proposers of EU research projects like to put forward in order to justify, from a utilitarian perspective, the need for better machine vision, better robot navigation, better object manipulation, more autonomy, et cetera. Again Neelie Kroes: "*The ease of use, safety, and partial autonomy are essential if robotic devices are to leave the shop floor and strictly controlled environments and become truly useful and helpful for people, including those with special needs. This could include steering a wheelchair, driving a car, guiding a blind person, performing precision surgery, operating a leg amputee's prosthesis, or many of our everyday chores.*"

[40] `http://ethicbots.na.infn.it/`

[41] http://www.eucognition.org/index.php?page=2013-fourth-eucogiii-members-conference-gen-info

[42] `http://www.cs.bath.ac.uk/~jjb/ftp/Bryson-Slaves-Book09.html`

But she also pointed out that there are people who want your research to inform the engineering of devices that could - for example - replace a soldier on the battlefields of our times and thus make destructive and lethal military action (including full-fledged war) (even) more "acceptable". She did not challenge this kind of use, probably for good reasons of her own (being a member of the "political class"). And there are indeed many who take this kind of use very seriously. So seriously that they devote a considerable amount of effort to researching the possibility of making such battlefield robots "ethical", for instance by having them respect the rules of combat or, what is known as "*ius in bello*"[43]. A good reason for robot reasoning? Perhaps. I for my part believe the prospect of this kind of *dual use* is an even better reason for *human* reasoning, for thinking harder about "*ius ad bellum*", and the reasons for waging war in the first place.[44]

Meta-robotics - economics

Let us return to the traditional mainstay of robotics, to the assembly lines and shop floors. Here it is quite obvious what robots should do: free human workers from hard labour. Headlines such as "*The next generation of robotic assembly lines are emerging*"[45] have appeared only recently, and this is happening not least thanks to the kind of research you are doing or going to do. Foxconn, arguably the world's largest manufacturer of electronic devices, announced only this month (July 2014) to install 10000 "Foxbots" in its new iPhone6 plant, replacing thousands of workers and at the same time greatly increasing the factory output[46]. And Google, apart from its autonomous car venture has embarked on full-fledged robotics through the acquisition of various robot companies[47] including, by the way, Boston Dynamics[48], famous for its BigDog walking robots and its millions of DARPA R&D monies.

Freeing human workers from hard work, well, that is good news. But here again: questions remain. As I mentioned before, industry mechanisation and automation has been going on ever since industrialisation began. It led in some parts of our planet to unprecedented wealth and a growing "service economy"[49]. However, in many traditional industries managers and owners found it

[43] http://www.cc.gatech.edu/ai/robot-lab/online-publications/formalizationv35.pdf

[44] http://www.youtube.com/watch?v=XNpfeLhMT_Q
http://web.stanford.edu/~jacksonm/war-overview.pdf

[45] http://gigaom.com/2014/02/11/the-next-generation-of-robotic-assembly-lines-are-emerging/

[46] http://fortune.com/2014/07/07/apple-foxbot-iphone-6/

[47] http://www.popsci.com/article/technology/why-google-building-robot-army

[48] http://www.bostondynamics.com/

[49] http://en.wikipedia.org/wiki/Service_economy

more advantageous to employ human labour in parts of the world that had not previously benefitted from the blessings of industrialisation. This has over the last decades led to a veritable process of de-industrialisation[50] in Europe and North America.

Can innovation in robotics reverse this trend? Will robot-based re-industrialisation[51] create new jobs for "the masses"? Can the promise of more jobs through science-based innovation really be kept? Can we keep extrapolating from the past and trust that human labour taken over by robots will find new niches in other domains? Services again? More services? Different services? There will most likely be more jobs for highly qualified people like you.

However, given that service robotics has also become a strong RTD focus[52] there is already a force at work that in a way defeats the assumption that jobs lost in manufacturing will be made up for by expanding services. Recent announcements by Amazon for instance, to experiment with drones for the delivery of parcels to its clients[53], foreshadow future developments. And if we are to believe the proponents of service robotics even jobs in hospitals and old people's homes will be in jeopardy. A frightening, comforting, or amusing prospect, depending on one's point of view.

Would we really want to replace a human carer by a machine, if such an option were available? Or should we not use these machines to complete other tasks or roles and give people more time to care for and help each other?

Perhaps the job question should be rephrased: Can robotics, the current apogee of industry automation (see above), contribute to making our economies more effective and more equitable in terms of providing the means for everyone on this planet to lead a life in dignity and peace? I have grave doubts that this is possible given the present constitution of our economies, their underlying power structures.

What will happen to all those whose money will then be made by robots? After all, robots being owned by few (not by the workers!), are not likely to share the money they make with the workers they are laying off. The money they make belongs to their owners unless we change the law. They do not buy goods either, to keep the economy running[54]. Like all machinery before them they are

[50] http://en.wikipedia.org/wiki/Deindustrialization

[51] http://ec.europa.eu/commission_2010-2014/tajani/priorities/reindustrialisation/index_en.htm
http://ec.europa.eu/enterprise/policies/industrial-competitiveness/competitiveness-analysis/european-competitiveness-report/files/eu-2013-eur-comp-rep_en.pdf

[52] http://www.springer.com/engineering/robotics/journal/11370

[53] http://www.amazon.com/b?node=8037720011

[54] http://www.vqronline.org/essay/machines-dont-buy-goods

likely to increase the gap between the haves and the have-nots (as documented in the recently published bestseller "Capital in the 21st Century"[55] by Thomas Piketty[56])?

So, do we have to rephrase our question yet again? For instance: must our economies be restructured in such a way as to harness the enormous increase in productivity that is likely to be brought about by robots and robotic devices, for the common good and the benefit of all?[57] A big question and not an easy one to answer given the well-known doomed (for whatever reason) approaches to socio-economic reform that in various parts of the world have been taken since the days of the European Enlightenment.

The philosophers of the Enlightenment taught us that we can be the masters of our own fate, individually and collectively. And scientists and engineers know that they can provide the means to change the world to our advantage. It is up to all of us but in particular to our elected representatives and rulers, sometimes referred to as the "political class", appropriately to respond to the challenges - positive and negative - posed by new technologies and insights, and to adapt the political and socio-economic structures accordingly. The law is made by law makers and law makers can change the law. A great challenge to HUMAN REASON in a world changing through human reasoning - called science.

Be aware of the bigger picture

Well, today all these issues are not of your immediate concern. So I am not going to keep you much longer from doing what you came here to do: to learn how to make robots reason in the real world. But I do believe that everyone, but in particular scientists and engineers, should be aware of the bigger picture of what they are working on professionally, its contexts and ramifications. There are many more or less prominent role-models in this regard.

One of them is Joseph Weizenbaum[58]. When he was 13 years old he escaped with his parents from Nazi-Germany to the United States of America. There he eventually became a professor of Computer Science at MIT. Later in his life he returned to Berlin where he died in 2008, aged 85. Many computer scientists

http://www.theatlantic.com/business/archive/2011/10/why-workers-are-losing-the-war-against-machines/247278/

[55] http://www.lrb.co.uk/v36/n13/benjamin-kunkel/paupers-and-richlings

[56] http://piketty.pse.ens.fr/en

[57] http://www.cepr.net/index.php/blogs/beat-the-press/robots-dont-cost-jobs-bad-economic-policy-does

[58] http://en.wikipedia.org/wiki/Joseph_Weizenbaum

remember his ELIZA[59] program, written in the early sixties, that could mimick inter alia the conversational patterns of a so called Rogerian psychiatrist. What mainly qualifies him to be named in the present context is a book he published some forty years ago: *Computer Power and Human Reason - from Judgement to Calculation*[60]. What motivated him to write it was the fact that many people took his ELIZA program seriously, suggesting for instance that it could make up for the shortage of psychiatrists. (Today we can observe something similar around IBM's "Jeopardy"-winning WATSON[61], a greatly beefed-up version perhaps of ELIZA.) Much of what people are discussing today under the heading "*Robot ethics*" can be found in this book. There are of course new issues some of which I tried to point out. So maybe one day someone authors a sequel to Weizenbaum's treatise which might then bear the title "*Robot Power and Human Reason*". The subtitle may even remain unchanged.

The second name I would like to mention is Noam Chomsky's[62]: he is five years Weizenbaum's junior and hardly needs an introduction. Given the scope of publications[63], ranging from "*Syntactic Structures*" to "*Profit over People*" and more recently, "*On Western Terrorism: From Hiroshima to Drone Warfare*", I need not explain either why I am shortlisting him. He has become the proverbial public intellectual.

I conclude this list with Nikola Tesla[64], a name you may have heard before (perhaps in your physics class in high-school). Apart from his many inventions, Tesla's main claim to fame is perhaps his part in designing the electric power grid in the United States of America, in the late 19th century. In a way he did for the electricity networks then what Vinton Cerf[65] and his colleagues did for the information networks in the second half of the 20th century.

Tesla was a somewhat colourful character with a tendency to bragging and presenting himself as a kind of star. But he is also widely considered a father (if not the father) of the robot as a technical device. Thus he provides the ideal closing bracket to this lecture where Karel Čapek provided the opening one. In his own way Tesla was a polymath greatly interested in explaining to the general public the potential of the technologies of his time, a bit like a futurologist and science fiction writer. In 1935 he published a short article entitled "*A Machine*

[59] `http://en.wikipedia.org/wiki/ELIZA`
[60] `http://en.wikipedia.org/wiki/Computer_Power_and_Human_Reason`
[61] `http://en.wikipedia.org/wiki/Watson_(computer)`
[62] `http://en.wikipedia.org/wiki/Noam_Chomsky`
`http://www.chomsky.info/`
[63] `http://en.wikipedia.org/wiki/Noam_Chomsky_bibliography`
[64] `http://en.wikipedia.org/wiki/Nikola_Tesla`
[65] `http://internethalloffame.org/inductees/vint-cerf`

to End War"[66] in a popular magazine. Unfortunately that machine has not yet been realised. And he made other predictions which nowadays appear rather outlandish. But some of his predictions may be more realistic. It is still upon you to make them come true. One of them is right at the end of his article:

"*Today the robot is an accepted fact, but the principle has not been pushed far enough. In the twenty-first century the robot will take the place which slave labor occupied in ancient civilization. There is no reason at all why most of this should not come to pass in less than a century, freeing mankind to pursue its higher aspirations*."

What aspirations? Whose aspirations? Chomsky's? Weizenbaum's? Or whose? Given that this text was written in 1935 you still have 20 years to find out. Good luck and thanks for listening.

[66] `http://www.tfcbooks.com/tesla/1935-02-00.htm`