CrossMark

The synthetization of human voices

Oliver Bendel[1]

Abstract The synthetization of voices, or speech synthesis, has been an object of interest for centuries. It is mostly realized with a text-to-speech system, an automaton that interprets and reads aloud. This system refers to text available for instance on a website or in a book, or entered via popup menu on the website. Today, just a few minutes of samples are enough to be able to imitate a speaker convincingly in all kinds of statements. This article abstracts from actual products and actual technological realization. Rather, after a short historical outline of the synthetization of voices, exemplary applications of this kind of technology are gathered for promoting the development, and potential applications are discussed critically to be able to limit them if necessary. The ethical and legal challenges should not be underestimated, in particular with regard to informational and personal autonomy and the trustworthiness of media.

Keywords Speech synthesis · Text-to-speech system · Artificial intelligence · Robotics · Information ethics · Machine ethics

1 Introduction

In November 2016, the project Adobe VoCo was presented on the Adobe Max in San Diego, and this prototype immediately won international attention (Beuth 2016). It could be called Photoshop for voices. 20 min of speech samples are enough, at least according to the promises, to imitate an individual voice in any kind of statement. The attendees and the testers certified the outstanding quality. They were much impressed but also concerned (Beuth 2016). In 2017, Lyrebird claimed that "it can recreate any voice using just one minute of sample audio" (Vincent 2017). This article abstracts from the actual products and from the actual technological realization. Rather, after a brief historical outline on the synthetization of sounds, tones and voices, to which repeated reference is made, exemplary applications of this technology are gathered for promoting the development, and potential applications are being discussed critically to be able to limit them if necessary, especially from an ethical perspective.

2 A short history of synthetization

For thousands of years, people have been fascinated by artificial creatures that serve us, support us and our friends, and eliminate our enemies. The works of Homer and Ovid are full of them (Bendel 2015) and the idea of them was popular in medieval times, renaissance, and baroque. Some of them lack speech, and incomprehension and muteness seem best suited to hint at the gap between humans and their creatures. However, there are some narrations of talking creatures, some of them even committed to truth, for instance the talking head by Virgil or by Gerbert of Aurillac (who was the Archbishop of Reims and later became Pope Sylvester II.). With regard to the Golem, a creature made of loam, Jacob Grimm at least noted that he would "understand most of what is spoken" (Grimm 1808).

✉ Oliver Bendel
oliver.bendel@fhnw.ch

[1] School of Business, University of Applied Sciences and Arts Northwestern Switzerland, Bahnhofstrasse 6, 5210 Windisch, Switzerland

The above was about history of the idea of speech synthesis. But what about the development history? The great age of automatons began in the late baroque, when the flute player and the mechanical duck by Jacques de Vaucanson as well as the androids—the draughtsman, the writer, and the musician—by the brothers Jaquet-Droz became very famous. The trio from 1774 today still is on display in a museum in Neuchâtel. The flute player and the musician both need airflow; the former is supplied by an artificial lung.

In 1779, Christian Kratzenstein built a "language organ" with freely swinging lingual pipes that produced five vowels. Wolfgang von Kempelen, best known for his chess-playing Turk, a pseudo automaton, began to design a talking machine in 1760, which he introduced in 1791 in an essay (Kempelen 1791). One of the elementary components was a single thudding lingual pipe capable of producing vowels as well as certain consonants, namely the plosives.

Charles Wheatstone constructed a speaking machine in 1837. 20 years later, Joseph Faber built the Euphonia, both followed the same principle (Klatt 1987). At the end of the 19th century, the tendency slowly but steadily led away from imitating human respiratory and speech organs towards simulating the acoustic space. Hermann von Helmholtz created vowels by means of tuning forks adjusted to the resonant frequencies (formants) of the vowel tract (Klatt 1987). Speech synthesis through the combination of formants was dominant well into the 1990s.

The Vocoder, a keyboard controlled speech synthesizer, was developed in the Bell Labs in the 1930s, reportedly with very comprehensible results (Klatt 1987). Homer Dudley improved this machine to the Voder, using electric oscillators to generate formant frequencies. The Voder was presented at the World Exhibition of 1939 in New York (Klatt 1987).

Since the 1950s, people have tried to teach computers how to speak. The first computer-based speech synthesis system was completed in the late 1950s, and the first full text-to-speech system in 1968 (Klatt 1987). Physicist John Larry Kelly, Jr developed a language synthesis in 1961 at Bell Labs with an IBM 704 and made it sing the folk song "Daisy Bell". Stanley Kubrick used it for his movie "2001: A Space Odyssey". The contemporary IBM Watson also features a text-to-speech engine that the user can program to make it speak his own text creations in different voices and different languages, while he controls pronunciation and accentuation via Speech Synthesis Markup Language (SSML).

In modern speech synthesis, two different concepts can be distinguished. At the one hand, the so-called signal modelling can refer to language recordings (samples). On the other hand, the signal can be generated fully on the computer through the so-called physiological (articulatory) modelling. Today, the first mentioned concept is predominant.

Through decades, speech samples have been made by professional speakers, mainly actors. New concepts have been developed recently. vocalID.org requests people to become donors, not of organs, but of their own voice. A database was furnished with thousands of voices and denoted "Voicebank". "By crowdsourcing the collection of voices, anyone can record from the comfort of their own home. Share your voice with others, or bank it for yourself." (Website vocalID.org)

Speech synthesis is mostly realized with a text-to-speech system (TTS), an automaton that interprets and reads aloud. This system refers to text available for instance on a website or in a book, or entered via popup menu on the website. Some systems, such as chatbots, can generate or aggregate text autonomously, and reproduce it.

3 Photoshop for human voices

There are some predecessors of VoCo and Co. and some experts question the novelty of such software (Plass-Fleßenkämper 2016). Still, a new level was probably reached in some regards with the samples concept; however, it was taken to extremes. Normally, one recorded as many words and phrases as possible. Now, very few samples seem to be enough to train VoCo to achieve the desired quality (Beuth 2016). According to Adobe, falsifications shall be made impossible by water-sign technology (Stark 2016), more precisely, by "acoustic water-signs" (Beuth 2016). As already mentioned, this article abstracts as much as possible from actual products and does not elaborate which safety measures could be developed and hacked.

Now that only a small quantity of individual speech material is necessary, and everyone is capable of making the synthesis, which new options will arise? Below, I list some, without claims for completeness, with the objective of gathering exemplary applications, and promoting the use and development in the field of synthetization. Own considerations are made, and reference is made to journalistic, popular, and scientific contributions, however scarce these are at the time being.

The literature frequently mentions the interview that never took place or the artificial citation generated by radio or TV stations from the original material (Stark 2016). In the last decades, the original tone was considered the last bastion in a media world characterized by the growth of falsifications and fakes—especially of visuals—threatening the integrity of newspapers and stations. These days seem to be over. Now, words can be put into the mouth of

anyone who has given at least one interview of some length, whose voice has been recorded, or broadcast live, for a certain time (and recorded afterwards). Fake news can be produced also by imitating the voice of the news speaker (Steinacker 2017). As in most others of the outlined cases, a text-to-speech system is the probable solution.

The voice can be used further to communicate with callers or visitors in a person's voice in the absence of the person. An interactive answerphone could be implemented that would not only transmit a defined message but also answer certain questions, for instance about the current place of sojourn of the person. Technologies as employed by the virtual assistants in Siri und Cortana could be used, or systems such as chatbots might be applied. Door inter-com systems could be expanded accordingly. Visitors could be rejected politely and seemingly personally. Per-sonified alternates in the context of communication and interaction have been an issue for a long time (Bendel and Gerhard 2004).

Those whose voice was recorded while alive could speak, or be spoken as would be proper to say, after their demise. Similar experiments have been made in the field of text. An AI expert recently chatted with her late friend by means of a suitable dialogue system (Nagels 2016). This will make some happy and others sad. Anyway it is, or will be, reality. This one case already shows that there is demand for it. Again, it has to be emphasized that this is not only about recorded voice only. 50 years ago, it was already possible to hear the voice of late beloved ones provided that their voice was recorded on tape. Here, the issue is a monologue of or a dialogue with a dead person.

Individual voices could also be implanted in robots. This would allow furnishing service robots in public with voices of prominent persons to attract interested persons and customers. At home, robots could acquire the voice of friends or partners. Not lastly sex robots, whose voices generally are very important, could be individualized (Bendel 2017b). During longer absences of a partner, they might substitute the partner also acoustically. Bordellos on the other hand could provide robots that speak like popular porn actors or famous music or movie stars to increase attractiveness.

Robots and, generally, partially autonomous and autonomous systems on principle are able to acquire voices on their own. As surveillance, information and entertain-ment robots (or as self-driving cars) and audio systems of all kinds, they could listen to their owners or to visitors, and after some time, they would be able to imitate them like parrots. An internal speech-to-text-to-speech system operates between microphones and loudspeakers. By means of artificial intelligence, the potential of applications could be expanded and tested. Further to entertainment applications, more serious purposes are also imaginable.

VoCo would also be interesting for producers of audi-bles, audio plays, podcasts, or radio programs (Beuth 2016). They would be able to improve failed voice recordings subsequently without necessity of participation and attendance of the bearer. Syllables or whole words, or even whole phrases, could be sampled anew. If the original material is very poor, one would depend on interpretations. "Mmmhs" and "aaahs" or a lump in the throat of the speaker could be eliminated more elegantly (Plass-Fleßenkämper 2016).

Frauds could use VoCo and Co. for biometric security processes and unlock doors, safes, vehicles, and so on, and enter or use them. With the voice of the customer, they could talk to the customer's bank or other institutions to gather sensitive data or to make critical or damaging transactions. All kinds of speech-based security systems could be hacked. It would also be possible to override human security systems, for instance a politician's assis-tant, to obtain confidential information or to stipulate abuse. If a voice is trusted, and this voice announces an attack, and for instance the president responds to this announcement, the consequences will be fatal.

Artificial voices are particularly effective in conjunction with video (Plass-Fleßenkämper 2016). The visual takes attention away from the auditive, which makes the speech imitation even more perfect, and image and sound work together for greater effect. Face2Face (Plass-Fleßenkämper 2016; Beuth 2016), a prototypical software for real-time face capture and reenactment of RGB videos (Thies et al. 2016), for example, could be used for this purpose.

Not lastly, animal voices could be recorded with this kind of tools. To produce sensible sequences, the cor-responding beings could understand; however, the lan-guage first has to be decoded in syntax and semantics. To date, decoding is available for very few species only, and for those only in parts. Still, this creates innovative possible applications, considering that higher developed species such as beluga whales or sperm whales recognize each other by their individual voices (Schulz et al. 2011).

4 A discussion of artificial voice

In the last chapter, exemplary applications have been gathered to view and promote the use and development in the field of synthetization. In this chapter, these will be discussed critically to enable responsible scientists and competent authorities to slow down or prevent the use and development from case to case. First, the special fields of ethics and machine ethics, the perspective of which is taken repeatedly, are explained briefly to be followed by the actual discussion.

4.1 Fields of applied ethics and machine ethics

Ethics is a sub-discipline of philosophy and originated two and a half millennia ago, initially mainly based on the works of Aristotle, which means following the Western scientific tradition. Applied ethics relates to delimitable topical fields and forms certain fields or specialties of ethics.

Morality (of and in) the information society is the object of information ethics (Bendel 2016b). It researches how we behave, or should behave, in moral terms when offering or using information and communication technologies (abbr. ICT), information systems and digital media. From a certain perspective comprising computer, network, and new media ethics, it is in the center of special ethics, and all of them have to deal with it, considering that all fields of applications are penetrated by ICT.

Technology ethics refers to moral questions of the application of technique and technology (Bendel 2016b). It can deal with automotive technology or arms technology as well as nanotechnology or nuclear technology. In the information society where more and more products contain computer technologies, technology ethics is particularly closely linked to information ethics and partly merges with it.

The object of media ethics is the morality of media and in media. Both the methods of mass media and the behaviors of the users of social media and their role as prosumers are of interest. Automatisms and manipulations by technologies move into the focus, linking it closely to information ethics.

The subject matter of machine ethics is the morality of machines, mostly that of partially autonomous and autonomous systems such as chatbots, certain robots, certain drones, and self-driving cars (Wallach and Allen 2009; Anderson and Anderson 2011; Bendel 2012). It can be assigned to information and technology ethics, or be understood as a pendant to human ethics, which means that it would not be a field of ethics but a new "core ethics". The term of morality is discussed quite controversially. However, it can be noted that autonomous systems more and more often have to make decisions of moral relevance, and these can be explicitly reasoned morally, for instance in annotated decision trees.

4.2 The theft of voice

Some native people reportedly were strongly averse to having their picture taken (Ingruber and Prutsch 2007). They understood photography as theft of their soul and feared disadvantages for their existence. In today's information society, we seem to be beyond such concepts. Still, many of us would probably feel strange if their voice took a

life of its own. In the 70s and 80s, many teens recorded their voice on tape and were astonished when they perceived the differences. As mentioned in the historical chapter, it has become possible to donate one's voice. Therefore, if one's voice speaks independently of one's self, the experience can be problematic, both in regards to the act of speaking and with regards to what is spoken, i.e., the contents, and it can be perceived as theft, not of something imaginary, but of something very real that finally and widely determines the identity in a great extent.

Here, ethics is required in many aspects. What is a human, what belongs to him or her, what can be separated from him or her, and what about the identity, and how to secure it? Should people determine while alive what shall happen to their voice after their death? Should they give orders for the data and information they put out during their lifetimes, as already happening with regards to social media (Lenke 2015), and shall these invoice their actions of speaking and other actions?

Information ethics can contribute its term of the digital identity. This identity is formed in the net, starting from a profile and documenting the activities of the user over a period of time. It is also highly relevant and a fix factor in the real space (for instance when job hunting or in the community of friends) (Bendel 2016a). Theft of the voice can influence and change the identity, and especially the digital identity. Not lastly, legal science is requested in this matter. Is it permissible to furnish robots with foreign voices without permission from the third party? Does one have a right in one's own voice, as one has a right in one's own image in many countries? Even if the contents are not critical, and it is immediately recognizable the tone is not original? Voice imitation is popular in satire, but when does it become a tool of deception?

4.3 The contents of speech

What is spoken, the contents of what was said, can be manipulated at will by those who have the necessary technology at their disposal. Therefore, one can put words in the mouth of a person, or let the person say something different or the opposite of what it really wanted to say, or make them say incredible things, or disclose personal things related to the person, its friends, or partners. This could discredit and compromise them, and link them to dubious statements unfolding a meaning in political or private contexts, this would be crucial especially for prominent persons (Stark 2016).

Especially at the beginning, one might still trust the original tone, and would not assume providers or stations had synthesized the voice, one would think any case that comes up was an exception. Maybe, one would also remember that the original tone was not always what it was

said to be. An interview can be falsified by abbreviation, or it can be furnished with an unsuitable context. Or the history of synthetization comes to mind, bringing to mind 18th century machines capable of talking. The manipulation might result in contents losing value, especially in later phases.

Information ethics researches cyberbullying and cyberstalking, both terms and phenomena unfold new meaning in this context. It is not really the original that causes friction, but a copy, however, the original is potentially damaged thereby. Next to information ethics, media ethics is requested to analyze the changed status of original tone and the modified attitude towards it. Both can work with terms such as "trust", "trustworthiness", and "reliability".

4.4 Contents are inflationary

If everything one says and hears can be produced automatically with little effort and can be untrue, the value of contents shrinks dramatically (Stark 2016). We have seen a similar deterioration of images. The first images were manipulated soon after photography came up, for instance by Roger Fenton, who placed some brothers and sisters next to the cannonballs in the "Valley of the Shadow of Death" for more dramatic effect (Lüpke 2014). It is known that Josef Stalin and Adolf Hitler had unwanted persons masked out of photos (Lüpke 2014).

Image editing programs such as Photoshop have long since enabled even lay persons to select image sections, modify image compositions, and replace image areas. Photoshop for voices now affects the spoken interview or statement as such, the value gets unclear or impaired, also in an economic sense. A freelance journalist who sells interviews and statements might not make the prices he or she needs, or might have to go to great lengths to prove the authenticity of his or her offers. He or she might have to use technical features or refer to certifying agencies, and the costs would have to split between all involved. On the other hand, the spoken word as such loses value, and becomes one of the many medial losers of the last centuries as well as a part of multimedia price dumping.

Media ethics is requested here as it can look back on decades of working with image and sound manipulations in the context of media. Information and technology ethics can deal with moral aspects related to production, use, and dissemination of information and (information) technologies.

Machine ethics also has to be involved in this matter. It has contributed studies on Munchausen machines (Bendel et al. 2016). Machines can be designed to lie systematically, but machines and their sources can also be secured in different ways, so that they could be called Kant machines (Bendel 2017a). Robots could be programmed to reproduce real voices under certain defined conditions only, for

instance only if the bearer gave his or her consent. Voice authentication cannot necessarily be trusted, so other or additional processes have to be found.

4.5 The breach with media

If voices can be reproduced, especially prominent persons, politicians and scientists should take care to whom they speak and whom they trust. This caution and care will be of little use, however, if a serious station brings a serious interview and this publicly broadcast sample could be reused as basis for manipulation. The only option would be to refrain from everything, to keep mum, but this is hardly feasible and would be detrimental to careers or to the diversity of opinion.

Not even a private person who does not want bots and robots speak with her or his own voice has sufficient control options. Auditive systems capable of eavesdropping are available in every household and at many public places. Every smartphone has an integrated recording function, and notebooks, tablets, and smart toys also can record voices. At this point, it has to be emphasized that systems and tools such as VoCo are available to a large number of users, and like with Photoshop, even untrained persons can produce acceptable results, at least after a certain training period.

Media ethics deals with topics like trust towards mass media, towards friends and their activities in social media, to which sound samples could be entered, for instance in video productions. Information ethics deals with topics like safeguarding and (re-)production of informational autonomy and the relevance of trust and trustworthiness in the information society, between humans, between human and machine, and in a figurative meaning between machines that produce and distribute contents.

4.6 An involuntary time warp

Especially providers of smart toys and smart TVs can gather voice not only at one point of time, but over a certain period of time. Parents, partners, and friends also can do so with the help of smartphones, etc. Further to the possible analysis of voice and content, this provides a lot of options for synthetization. Entire biographies can be invented; contradictions can be eliminated or added. A person's alleged progress, or regress, from childhood to old age can be fictionalized. Just as Facebook has childhood photos that might embarrass the teen or adult, childish speech can be sampled or invented.

Yet, this context has one particularity: A voice sounds different in childhood than in adulthood, this is true for both men (the voice change makes male voices sound one octave deeper) and for women (for whom the voice change

is not as grave). Voices continue to change after reaching maturity. This hints at certain limitations of VoCo. If a voice was recorded at one date, it might no longer be convincing 20 years later. Of course, processes could be found to make artificial voices age artificially. Then, there would be virtually no more limits for use and abuse.

Information and technology ethics can discuss, together with science ethics, what technology is allowed to do and should do. If there are so many opportunities for abuse, does research have to stop at a certain point? Or does it have to take place and legislative and judicative powers have to limit its application? It has to be remembered that almost every technique and technology can be alienated. Still a technique and technology, the purpose of which is not clearly defined and which still has to find its objective, can be dangerous, a plaything the use of which has serious consequences.

5 Summary and outlook

Synthetization of voices has been an object of interest for centuries. The brief historical outline showed both the enormous progress that took only a few centuries as well as explained different approaches to synthetization. Surely, there will be more leaps, and surely, artificial voices will be perfected. The growing availability of audio systems in households and in public as well as phenomena like voice donation will make sure every one of us soon will be represented with sufficient materials and thereby prone to artificial articulation and imitation. We seem to enter a new area of falsification. Whether the presently discussed technology is fully novel or not—through companies like Adobe, it could reach a mass audience, even if this is not its dedicated purpose.

It became obvious that there are many fields of application, and some of them will be profitable and attractive for companies, media, and private persons. Police and secret services could try to generate incriminating material to eliminate unwanted persons. It was not possible to present and discuss all of these and all other potential uses here, but it was also shown that the ethical and legal challenges should not be underestimated. Ethics and legal science, but also computer science, artificial intelligence, and robotics all have to deal with the problems of technology without killing its chances.

References

Anderson M, Anderson SL (eds) (2011) Machine ethics. Cambridge University Press, Cambridge

Bendel O (2012) Maschinenethik. Gabler Wirtschaftslexikon. Springer Gabler, Wiesbaden. http://wirtschaftslexikon.gabler.de/Definition/maschinenethik.html

Bendel O (2015) Surgical, Therapeutic, nursing and sex robots in machine and information ethics. In: Rysewyk SPV, Pontier M (eds) Machine medical ethics. Series: intelligent systems, control and automation: science and engineering. Springer, Berlin, pp 17–32

Bendel O (2016a) Cloud Computing aus Sicht von Verbraucherschutz und Informationsethik. In: Reinheimer S (ed) HMD—Praxis der Wirtschaftsinformatik, 25 July 2016 ("online first" article on SpringerLink)

Bendel O (2016b) 300 Keywords Informationsethik: Grundwissen aus Computer-, Netz- und Neue-Medien-Ethik sowie Maschinenethik. Springer Gabler, Wiesbaden

Bendel O (2017a) Towards Kant machines. In: The 2017 AAAI spring symposium series. AAAI Press, Palo Alto

Bendel O (2017b) Sex robots from the perspective of machine ethics. In: Cheok AD, Devlin K, Levy D (eds) Love and sex with robots. second international conference, LSR 2016, London, UK, December 19–20, 2016, Revised Selected Papers. Springer International Publishing, Cham, pp 1–10

Bendel O, Gerhard M (2004) Handy-Avatare—Möglichkeiten der mobilen Kommunikationsunterstützung. InfoWeek.ch, 12, pp 51–55

Bendel O, Schwegler K, Richards B (2016) The LIEBOT project. In: Machine ethics and machine law, Jagiellonian University. November 18–19, 2016, Cracow, Poland. E-Proceedings. Jagiellonian University, Cracow. http://machinelaw.philosophyinscience.com/technical-program/

Beuth P (2016) Tonaufnahmenfälschen leicht gemacht. ZEIT ONLINE, 4 November 2016. http://www.zeit.de/digital/internet/2016-11/adobe-project-voco-photoshop-audio-manipulation

Grimm J (1808) Entstehung der Verlagspoesie. Arnim LAv (ed) Zeitung für Einsiedler

Ingruber D, Prutsch U (2007) Imágenes—Bilder und Filme aus Lateinamerika. LIT, Münster

Kempelen WV (1791) Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine. J. B. Degen, Wien

Klatt D (1987) Review of text-to-speech conversion for English. J. Acous. Soc. Amer. 82:737–793

Lenke M (2015) Nutzerprofile nach dem Tod: So regeln Sie Ihren digitalen Nachlass. Focus Online, 12 February 2015. http://www.focus.de/digital/internet/sterben-2-0-virtuelle-grabpflege-so-regeln-sie-ihren-digitalen-nachlass_id_4224951.html

Lüpke MV (2014) Als die Fotos lügen lernten. Spiegel Online (Eines Tages), 13 October 2014. http://www.spiegel.de/einestages/bildmanipulation-falsche-fotos-vor-der-digital-aera-a-996453.html

Nagels P (2016) Wie eine Russin ihren toten Freund zum Leben erweckt. welt.de, 7 October 2016. https://www.welt.de/kmpkt/article158616017/Wie-eine-Russin-ihren-toten-Freund-zum-Leben-erweckt.html

Plass-Fleßenkämper B (2016) Dank Adobe können wir unseren Ohren nicht mehr trauen. WIRED Germany, 8 November 2016. https://www.wired.de/collection/tech/adobes-neues-tool-kann-sprache-imitieren

Schulz TM, Whitehead H, Gero S (2011) Individual vocal production in a sperm whale (Physeter macrocephalus) social unit. MARINE MAMMAL SCIENCE, 27(1), January 2011, pp 149–166

Stark J (2016) Adobe stellt Sprach-Software Voco vor. com! professional, 9 November 2016. http://www.com-magazin.de/news/adobe-systems/adobe-stellt-sprach-software-voco-1146967.html

Steinacker L (2017) Wirtschaftwoche, 20 January 2017. Mit falscher Stimme, p 52

Thies J, Zollhöfer M, Stamminger M, Theobalt C, Nießner M (2016) Face2Face: real-time face capture and reenactment of RGB videos. In: Proceedings computer vision and pattern recognition (CVPR), IEEE. http://www.graphics.stanford.edu/~niessner/papers/2016/1facetoface/thies2016face.pdf

Vincent J (2017) Lyrebird claims it can recreate any voice using just one minute of sample audio. The Verge, 24 April 2017. http://www.theverge.com/2017/4/24/15406882/ai-voice-synthesis-copy-human-speech-lyrebird

Wallach W, Allen C (2009) Moral machines: teaching robots right from wrong. Oxford University Press, Oxford