Mariella Scerri and Victor Grech

Artificial Intelligence in Medicine

Abstract

Various types of artificial intelligence programs are already available as consultants to physicians, and these help in medical diagnostics and treatment. At the time of writing, extant programs constitute "weak" AI—lacking in consciousness and intentionality. With AI currently making rapid progress in all domains, including those of healthcare, physicians face possible competitors—or worse, claims that doctors may become obsolete. We will explore the development of AI and robotics in medicine and examine the "Emergency Medical Hologram," a strong AI program (presented as being as fully "conscious" and motivated as the humans whose form he mimics) who serves as ship's doctor on *Star Trek: Voyager*. We will also briefly explore the issues pertaining to AI in the medical field and will show that weak AI not only should suffice in the demesne of healthcare but may actually be more desirable than strong AI.

What Do We Mean By AI

For the purposes of this discussion, we describe artificial intelligence as being based in accurate decision-making processes that can be carried out independently by a machine.

Fundamental to the ambitions of artificial intelligence (AI) is a thesis of computational sufficiency, the conjecture that any type of computation or decision-making can be modeled digitally (MunKim 77). MunKim contends that there is another tenet on which many AI researchers rely, the so called "AI thesis," that "As the intelligence of the machines evolves, its underlying mechanisms will gradually converge to the mechanisms underlying human intelligence" (71). From the beginning, AI researchers did not shy from making predictions of the anticipated coming successes in the development of strong AI. The following statement by Herbert Simon in 1957 is often quoted:

> It is not my aim to surprise or shock you—but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until—in a visible future—the range of problems they can handle will be co-extensive with the range to which the human mind has been applied (Simon 198).

Simon made concrete predictions that within ten years AI would have flourished. These predictions came true within 40 years, not ten. Russell and Norvig suggest that Simon's overconfidence was due to the promising performance of simple, early AI systems (Russell and Norvig 21).

However, with AI currently making vast and uninterrupted progress in the domain of medical health and elsewhere, physicians are now facing possible peers amidst claims that doctors may become obsolete. This paper aims to explore the juxtaposition of medicine and AI and the impact AI has and will continue to have on medicine. It will also seek to analyze the implications this will have on medical doctors and healthcare.

Artificial Intelligence in Medical Diagnosis and Interventions

The steady expansion of medical knowledge has made it progressively more difficult for the physician to remain abreast of medicine outside of a chosen specialty. Campbell alleges that this happens because knowledge is composed of a "continuous texture of narrow specialties ... a redundant piling up of highly similar specialties leaving interdisciplinary gaps" (Campbell 328).

By the early 1970s it became clear that conventional tools such as flow charts, pattern matching, and Bayes's theorem were unable to deal with more complex clinical problems (Gorry 50). Investigators thus began to study expert physicians in order to obtain detailed insights into the nature of clinical problem-solving (Kassirer and Gorry 250; Swanson *et al*. 160). The results derived from such studies have subsequently formed the basis for computational models of cognitive phenomena, and these models have further been converted into so-called artificial intelligence programs (Weiss *et al*. 150).

Many of the early efforts to apply artificial intelligence methods to real problems, including medical reasoning, have primarily used rule-based systems (Duda and Shortcliffe 263). Such programs are typically easy to create, since their knowledge is catalogued in the form of chains of simple "if-then" rules that reach a conclusion. However, most serious clinical problems are broad and complex such that straightforward attempts to chain together larger sets of rules often encounter major difficulties. Problems arise principally from the fact that rule-based programs do not embody a model of disease or clinical reasoning (ibid.).

Given the difficulties encountered with rule-based systems, more recent efforts to use AI in medicine have focused on programs organized around models of disease. Efforts to develop such programs have led to substantial progress in the conversion of various models into promising experimental programs. These improved representations of clinical knowledge and sophisticated problem-solving strategies have significantly advanced the field of artificial intelligence in medicine.

Any program designed to serve as a consultant to the physician must contain certain basic features. It must have a store of medical knowledge expressed as descriptions of possible diseases. Depending on the breadth of the clinical domain, the number of hypotheses in the database can range from a few to many thousands. In the simplest conceivable representation of such knowledge, each disease hypothesis identifies all the features that can occur in a particular disorder. In addition, the program must be able to match what is known about the patient within its store of information (Dilsizian and Siegel 442).

Medical diagnostic programs similar to the Virtual Space Station could serve as useful tools in today's overburdened healthcare system. For instance, IBM's Watson—best known for its remarkable performance on *Jeopardy!*—is now being used in healthcare applications. It provides many unique and transformative possibilities to resolve challenges associated with medical diagnosis and treatment. The Watson hardware costs approximately three million dollars. It can process 500 gigabytes of data per second, the equivalent of one million books (Pearson). Such medical applications may help physicians navigate a complex set of patient symptoms, laboratory data, and imaging results to come up with a set of most likely clinical diagnoses and treatment options. Software of this nature may ultimately improve patient outcomes and reduce health care costs (Dilsizian and Siegel 444). IBM initially utilized the *American College of Physicians Medical Knowledge Self-Assessment Study Guide* and subsequently improved on its performance by adding textbooks such as the *Merck Manual of Diagnosis and Therapy*, and additional medical journals not included in the original database. The developers then further improved on the performance for medical applications by fine-tuning the weighting associated with the various algorithms used by the application for these medical domain questions. The team also created a demonstration in which a patient's presenting symptoms are input into the software and a series of progressive questions are posed by a healthcare worker to personalize the diagnostic and therapeutic recommendations made by the software (Pearson). Two important features of the prototype software were the ability to provide multiple possible diagnoses and treatment options with relative confidence levels and the ability to trace the information used to make a recommendation (Dilsizian and Siegel 444).

AI is also being used in cardiology, including the determination of the most appropriate type of imaging study for a specific set of symptoms (Fornell 35). For example, the Imaging in FOCUS (Formation of Optimal Cardiovascular Utilization Strategies) quality improvement initiative of the American College of Cardiology was recently introduced to improve the use of diagnostic imaging through the arrangement of AI that tracks appropriate use criteria (Saifi et al 823). Fifty-five participating sites voluntarily completed the radionuclide imaging performance improvement module. Results showed that the proportion of inappropriate cases decreased by half—from ten percent to five percent. These preliminary data suggest that the use of self-directed quality improvement software and an interactive community may permit physicians to significantly decrease the proportion of tests not meeting appropriate use criteria (Saifi et al 825).

After images are acquired, additional AI tools may help physicians provide accurate interpretation of cardiac imaging studies (De Puey et al 1165). Artificial neural networks are thus excellent example of ways in which current AI systems that emulate human decision-making can be successfully applied in cardiac imaging (Itchhaporia *el al.* 517). These networks have been utilized in the diagnosis and treatment of coronary artery disease and myocardial infarction, the interpretation of electrocardiographic studies, and the detection of cardiac arrhythmias, such as ventricular fibrillation (Clayton et al 219).

These expert systems are costly to develop and difficult to maintain and require a perfect match between input data and existing rule forms. Software such as that which underlies Watson, on the other hand, uses natural language processing and a variety of search techniques to create hypotheses, making it more flexible, scalable, easy to maintain, and cost-effective. This new approach makes it much easier to keep up with ever-changing information in imaging, medicine, and surgery.

In a future clinical image interpretation scenario using AI, a requested test would first be evaluated for appropriateness based on the patient history and previous examination. The examination would, if deemed appropriate, then be "protocoled" with regard to the way in which it should be performed. For example, a magnetic resonance scan would be fine-tuned as to imaging sequences and amount and type of radiopharmaceutical and/or contrast material to optimize efficacy, safety, and efficiency of the investigation (Clayton *el al.* 221).

Technology is also revolutionizing the medical field through the creation of robotic devices and complex imaging. Robotically assisted surgery was developed to overcome the limitations of preexisting, minimally invasive surgical procedures and to enhance the capabilities of surgeons performing open surgery.

The first documented use of a robot-assisted surgical procedure was in 1985 when the PUMA 560 robotic surgical arm was used in a delicate neurosurgical biopsy, an open and therefore non-laparoscopic procedure (Kwoh et al 155). The system allowed successful robotic surgery and the potential for greater precision when used in minimally invasive surgeries. This led to the first laparoscopic procedure in 1987 involving a robotic system to carry out a cholecystectomy (gall bladder removal) (Jones and Jones 120). The following year, the same PUMA system was used to perform a robotic surgical transurethral prostate resection. In 1990, the AESOP system produced by Computer Motion became the first system approved by the Food and Drug Administration (FDA) for endoscopic procedures (Jones and Jones 123).

In the case of robotically assisted, minimally invasive surgery, instead of directly moving the instruments, the surgeon uses one of two methods to control the instruments—either a direct tele-manipulator or through computer control. A tele-manipulator is a remote manipulator that allows the surgeon to perform the normal movements associated with the surgery while the robotic arms mimic the same movements to operate on the patient. In computer-controlled systems the surgeon uses a computer to control the robotic arm and its end-effectors, though these systems can still use tele-manipulators for their input. One advantage of using the computerized method is that the surgeon does not have to be present, making remote surgery possible. One minimally invasive option is the *da Vinci* surgery system which has brought minimally invasive surgery to more than three million patients worldwide (Samadi). In 2000, the *da Vinci* system broke new ground by becoming the first robotic surgery system approved by the FDA for general laparoscopic surgery (Samadi). This was the first time the FDA approved an encompassing system of surgical instruments and camera-scopic utensils. The *da Vinci* system's three-dimensional magnification screen allows the surgeon to view the operative area with high resolution clarity. The one-centimeter-diameter surgical arms represent a significant advancement in robotic surgery from the early, large-armed systems such as the PUMA 560 (Samadi). This advance allows for less contact between exposed interior tissue and the surgical device, greatly reducing the risk of infection. The "Endo-Wrist" features of the operating arms precisely replicate the skilled movements of the surgeon at the controls, improving accuracy in small operating spaces (Samadi).

Robotic surgery is at the cutting edge of precision and miniaturization in the realm of surgery, thus the possible applications are as extensive as the uses of minimally invasive surgery. Robotic surgery has already become a successful option in neurological, gynecological, cardiothoracic, and numerous general surgical procedures (Samadi).

It is reasonable to assume that the current advantages of robotic surgery systems will be expanded with the next generation of medical robotics. Removing human contact during surgery may be taken to the next level with robotic surgery systems capable of functioning at greater distances between surgeons' control consoles and the patient side table robotics. This would allow robotic surgery to be conducted with patients in a nearby "clean room," reducing the risk of intraoperative infection. Major strides are also being made in creating robotic surgery systems which are more capable of replicating the tactile feel and sensation which a surgeon experiences during more invasive traditional procedures, allowing the operator the precision and advantages of minimally invasive procedures without losing the sensory information that is so helpful in making judgment calls during robotic surgery.

Weak AI and Strong AI

There are two major ways to think about the current utilization and power of artificial intelligence. The weak AI hypothesis states that a machine running a program is at most only capable of simulating real human behavior and consciousness. Artificial intelligence such as that being used in medical diagnosis and certain interventions are examples of weak AI because they are focused on one narrow task. Existing weak AI is consistent with the idea that a running AI program is only a simulation of a cognitive process but is not itself a cognitive process. The strong AI hypothesis, on the other hand, claims that an (as yet to be written) program running on an (as yet to be designed) machine can actually be a mind—that there is no essential difference between a piece of software exactly emulating the actions of the brain and the actions of a human being including their understanding and consciousness.

The debate between the opposing camps is intense with scientists and philosophers of both camps foregrounding claims to support their hypotheses. John Searle's familiar Chinese Room argument is a celebrated thought experiment designed to refute the hypothesis that "the appropriately programmed computer really is a mind" (Searle 420). To a Chinese speaker external to the room, by virtue of its question answering ability, the system passes the Turing test for machine intelligence, yet the system implemented by Searle-in-the-room is entirely without understanding simply because Searle does not understand Chinese (Searle 455).

Searle concludes that an AI program could give the impression of intelligence to an external observer but completely lack understanding. Searle and Ned Block contend that even if a system duplicates our behavior, it might be missing important "internal" aspects of mentality: consciousness, understanding, and intentionality (Searle; Block). Searle further purports that causal features of the brain are necessary and crucial for intentionality.

> "Could a machine think?" ... only very special kinds of machines, namely brains and machines with internal causal powers equivalent to those of brains. And that is why strong AI has little to tell us about thinking, since it is not about machines but about programs, and no program by itself is sufficient for thinking. (Searle 417)

Searle upholds "biological naturalism." He explains that human mental phenomena such as consciousness and understanding require a specific biological process that is found in brains of human beings, an important combination of physical and chemical properties. Thus a man-made, intelligent system that acted exactly like a human mind might still not be conscious (Searle 264). It is therefore Searle's contention that strong AI may share our behavioral and functional equivalence without being a conscious system since consciousness requires not only a computational organization but also a specific and unknown way in which functional organization is implemented in the biology of the organism. This is because Searle believes that the human brain is able to conjure "perceptual aboutness" (Natsoulas 76) and that:

> the brain's causal capacity to produce intentionality cannot consist in its instantiating a computer program, since for any program you like it is possible for something to instantiate that program and still not have any mental states. Whatever it is that the brain does to produce intentionality, it cannot consist in instantiating a program since no program, by itself, is sufficient for intentionality (Searle 424).

Colin McGinn further argues that the famous Chinese Room *gedankenexperiment* provides strong evidence that the debate revolves around whether a machine can be shown to actually be conscious as opposed to the theory that a machine can be conscious (McGinn 194). Searle, however, doubts that true consciousness in an android could ever be possible with our present state of knowledge. This is contrary to the tenets of strong AI—that computational states are functionally equivalent to mental states (Schank and Abelson; McGinn). DeDompablo Cordio opines that strong AI itself seems to be rooted in large part in science fiction (52).

Making the EMH

Strong AI is commonly depicted in popular culture, recently and famously in *Ex Machina* (Garland) and *Transcendence* (Pfister). Another exemplar of strong AI is the Emergency Medical Holographic program (EMH) created in *Star Trek*. The EMH is portrayed as a sophisticated hologram developed in the early 2370s and used on most Federation starships. It was designed to provide short-term advanced assistance during emergencies in sick bay to the extent of literally replacing a starship's medical officer. The EMH Mark 1 was first activated in 2371, and it was programmed with over five million possible treatments from the collective information of 2000 medical references and the experience of 47 individual medical officers. The EMH was also supplemented with contingency programs and adaptive programs to learn while serving as a supplement to a normal medical staff in cases of emergency. It contained fifty million "gigaquads" of computer memory, which is "considerably more than most highly developed humanoid brains" (Bole "Lifesigns").

Some might argue that the ideal of Strong AI would be to have a perfect replica of *Homo sapiens*, similar to the EMH, with identical physical and psychological attributes instantiated in a single system. However, the pressing point here is whether there will be an eventual desire or even necessity for strong AI to replace medical doctors. If an AI is programmed to possess two fundamental characteristics, flawless computation and indefatigability, so as to reach an accurate diagnosis treatment, AI would have reached its ultimate medical goal.

These attributes all point to the conclusion that ultimately, AI will function independently and will constitute a major part of the work force. Thus, this strongly suggests a major transition in the medical health field, where doctors, nurses, and the rest of the multidisciplinary team will be replaced with stronger intelligences. One obvious counterargument is that even a strong AI will be emotionless and lacking compassion—a paramount ingredient to care. Emotion is crucial and is inextricably tied to everything we say and do. Instantiating compassion modules as software-hardware components of such AI systems would make such systems far more acceptable. The idea of simulating emotion in computers is of great interest to AI research and psychology. While psychologists have known for some time that there are a great many physical correlates to emotion—voice changes, blushing, pupil dilation, etc.—reproducing them continues to prove difficult. Emotion provides us with motivation and drive and with a set of personal preferences, features that would be desirable in a sophisticated AI (Maciamo) and bordering on necessary for a truly effective physician surrogate AI such as the EMH.

The Sociable Machines Project has developed an expressive anthropomorphic robot called Kismet that engages people in natural and expressive face-to-face interaction. Inspired by infant social development, psychology, ethology, and evolution, this work integrates theories and concepts from these diverse viewpoints in order to enable Kismet to enter into a natural and intuitive social interaction with a human caregivers, and to learn from them (Breazeal 10). In order for Kismet to do so, it has the ability to perceive a variety of natural social cues from visual and auditory channels and to deliver social signals to the human caregiver through gaze direction, facial expression, body posture and vocal babbles. The robot has been designed to support several social cues and skills that could ultimately play an important role in socially situated learning with a human instructor (Breazeal 10).

Sociable humanoid robots pose a dramatic and intriguing shift in the way one considers the control of autonomous robots. Traditionally, autonomous robots are designed to operate as independently and as remotely as possible from a human. However, a new range of application domains (domestic, entertainment, health care) are driving the development of robots that can interact and cooperate with people, and thus play an active part in their lives (Breazeal

10). With this in view, nothing stands in the way of an AI medical program replacing the human medical doctor.

Conclusion

In James P. Hogan's 1979 novel, *The Two Faces of Tomorrow*, an artificial intelligence system solves an excavation problem on the moon in a brilliant and novel way but nearly kills a work crew in the process. Realizing that systems are becoming too sophisticated and complex to predict or manage, a scientific team sets out to teach a sophisticated computer network how to think more humanly. The story documents the rise of self-awareness in the computer system, the humans' loss of control and failed attempts to shut down the experiment as the computer desperately defends itself, and the computer intelligence reaching maturity. This novel illustrates the successes and pitfalls which the real world might eventually face when technology progresses from the current developments of weak AI to potential developments of strong AI.

However, weak AI would actually suffice for medical purposes. Such systems do not need to have intentionality and self-awareness to function well. The present AI programs available on the market come with a computational sufficiency and serve as useful tools in today's overburdened healthcare system. Medical aid programs are providing correct diagnoses within minutes if not seconds while robotic surgery is making huge advances in the field. This suggests that a strong AI might not even be necessary. If these systems are attainable now, the next obvious question is whether strong AI is truly necessary in the medical field. If professionals, especially those in the medical field, are served well by weak AI, how important is it to create a strong AI for these purposes?

Furthermore, a strong AI like the fallible EMH might not necessarily be better in this demesne. A weak AI would be controllable due to lack of intentionality and independence: a mind of its own. The lack of consciousness and intentionality in a weak AI makes it what it simply is—a programmable tool. On the other hand, a strong AI will be able to function independently, possessing the attributes of free will and moral conscience. The dilemma lies here: while a strong AI might be appealing for tomorrow's world because of its great potential, a weak AI is what human beings might actually need. *Homo sapiens* have reigned supreme in the animal kingdom since the beginning of their existence—relegating any portion of our power to a strong AI might not be on our agenda for many years to come, making us think twice about the creation of such machines.

Mariella Scerri teaches English in the Humanities, Medicine and Sciences Program at the University of Malta. Victor Grech is an associate professor of pediatrics at the University of Malta.

References

Films and Television

Ex Machina. Dir. Alex Garland. Universal Pictures. 2015.

"Lifesigns." Dir. Cliff Bole. *Star Trek: Voyager*. Paramount Pictures. February 1996.

Transcendence. Dir. Wally Pfister. Warner Bros. Pictures. 2014.

"The Ultimate Computer." Dir. John M. Lucas. *Star Trek: The Original Series*. Paramount Pictures. March 1968.

Primary Texts

Block, Ned. "Psychologism and behaviorism." *Philosophical Review* 90 (1981).

Breazeal, Cynthia. "Emotion and Sociable Humanoid Robots." *Elsevier Science*. (2002).

Campbell, Donald T. "Ethnocentrism of Disciplines and the Fish Scale Model of Omniscience." *Interdisciplinary Relationships in the Social Sciences*, edited by Muzafer Sherif and Carolyn W. Sherif. Chicago: Aldine Publishing Co., 1969.

Clayton, R, A. Murray, and RWF Campbell. "Recognition of Ventricular Fibrillation using Neural Networks." *Medical and Biology Engineering and Computing* 32 (1994).

DeDompablo Cordeo, Mariano. "Searle's Chinese Room Argument and its Replies: A Constructive Re-Warming and the Future of Artificial Intelligence." *Indiana Undergraduate Journal of Cognitive Sciences* 3 (2008).

Dilsizian, Steven E and Eliot L Siegel. "Artificial Intelligence in Medicine and Cardiac Imaging: Harnessing Big Data and Advanced Computing to Provide Personalized Medical Diagnosis and Treatment." *Nuclear Cardiology* 16 (2014).

Duda, R.O and E.H. Shortlife. "Expert Systems Research." *Science* 220 (1983).

Fornell, D. "An Introduction to Clinical Decision Support for Cardiology." *Diagnosis and Intervention Cardiology*. 2003. <www.dicardiology.com/article/introduction-clinical-decision-support-cardiology>. Accessed 21 February 2016.

Gorry, G.A. "Computer-assisted Clinical Decision Making." *Methods of Information in Medicine* 12 (1973).

Hogan, James P. *The Two Faces of Tomorrow*. New York: Del Rey Books, 1979.

Itchhaporia, D., P.B. Show, R.J. Almassy, and W.J. Oetgen. "Artificial Neural Networks: Current Status in Cardiovascular Medicine." *Journal American College of Cardiology* 28 (1996).

Jones, S.B and D.B. Jones. "Surgical Aspects and Future Developments in Laparoscopy." *Anesthesiology of Clinical North America* 19 (2001).

Kassirer, J.P and G.A. Gorr. "Clinical Problem Solving: A Behavioral Analysis." *Annals of Internal Medicine* 89 (1978).

Kwoh, Y.S, J. Hou, E.A. Jonckheere, and S. Hayall. "A Robot with Improved Absolute Positioning Accuracy for CT Guided Stereotactic Brain Surgery." *IEEE Transactions on Biomedical Engineering.* 35.2 (1988).

Maciamo, Hay. "Could a machine or an AI ever feel human-like emotions?" *Humanity Media* 2014. <www.hplusmagazine.com>. Accessed 29 February 2016.

McGinn, Colin. *The Mysterious Flame: Conscious Minds In A Material World: Conscious Minds in a Material World*. New

York: Basic Books, 2000.
Natsoulas, Thomas. "On Perceptual Aboutness." *Behaviorism* 5.1 (1977).
Pearson, T. "IBM Watson. How To Build Your Own Watson Jr In Your Basement." *IBM Developer Works Blog.* February 2011. <www.ibm.com/developerworks/mydeveloperworks/blogs>. Accessed 15 January 2016.
Russell, Stuart J. and Peter Norvig. *Artificial Intelligence: A Modern Approach*. USA: Prentice Hall. 2010.
Saifi, S., A.J. Taylor, H. Allen, and R. Hendel. "The Use of a Learning Community and Online Evaluation of Utilization for SPECT myocardial perfusion imaging." *Journal American College of Cardiology Imaging* 6 (2013).
Samadi, David B. *History of Robotic Surgery*. <www.roboticoncology.com/history/>. Accessed 22 February 2016.
Schank, Roger C. and Robert P. Abelson. "Scripts, Plans, Goals and Understanding; An Inquiry into Human Knowledge Structures." *The American Journal of Psychology* 92.1 (1979).
Searle, John. "Minds, Brains and Programs" and "Intrinsic Intentionality." *Behavioral and Brain Sciences* 3 (1980)
Simon, Herbert. *Models of Man: Social and Rational*. New York: John Wiley and Sons Inc. 1957.
Swanson, D.B., P.J. Feltovich, and P.E. Johnson. "Psychological Analysis of Physician Expertise: Implications for Design of Decision Support Systems." *Proceedings of the Second World Conference on Medical Informatics*, edited by Wolf H. Shires. New York: North Holland, 1977.
Weiss, S.M., C.A. Kulikowski, S. Amarel, and A. Safir. "A Model–Based Method for Computer-Aided Medical Decision Making." *Artificial Intelligence* 11 (1978).