Stud Philos Educ (2012) 31:357–364
DOI 10.1007/s11217-011-9279-y

Education for Computers

Robert D. Heslep

Published online: 15 November 2011

Abstract The computer engineers who refer to the education of computers do not have a definite idea of education and do not bother to justify the fuzzy ones to which they allude. Hence, they logically cannot specify the features a computer must have in order to be educable. This paper puts forth a non-standard, but not arbitrary, concept of education that determines such traits. The proposed concept is derived from the idea of education embedded in modern standard-English discourse. Because the standard concept entails that an educable entity must be capable of consciousness and voluntary action, it cannot apply to computers. If, therefore, one is to have an idea of educable computers, one must drop the feature of consciousness and omit or modify that of voluntariness. The advanced concept leaves out consciousness, alters the ordinary notion of voluntariness, but keeps in tact the other criteria of the standard idea. Thereby, it provides continuity between those who talk about education in modern ordinary English and those who talk about it in the world of artificial intelligence.

Keywords Artificial intelligence · Education · Computer · Conceptual analysis · Rationality · Voluntary action

Introduction

When people speak of entities as educable, they typically are referring to humans, certainly not to things electronic. Nevertheless, notable promoters of artificial intelligence do speak of supercomputers as possible subjects of education. Thus, Kurzweil (2005, pp. 202–203, 336–337), who holds that AIs presently can facilitate the education of human beings, maintains that educable AIs probably will appear in the future. And Baker (2011, pp. 81–103) devotes a chapter to the "education of Blue J" for an appearance on the TV show, "Jeopardy!"

R. D. Heslep (✉)
University of Georgia, 285 Hancock Lane, Athens, GA 30605, USA
e-mail: rheslep@uga.edu

What motivates AI enthusiasts to describe smart machines as educable is not evident. Maybe they think that the positive aura of the word "education" makes their gadgets more acceptable to the general public. Perhaps they suffer from something like the Dr. Frankenstein complex: While they cannot create human beings, they at least can build clever mechanisms with a human-like feature. Or possibly they just want to establish some continuity with pre-AI civilization.

Regardless of their motivation, AI proponents have a major problem with their claim about the possible education of computers, which is that the claim typically does not rest on a clear idea of education. Kurzweil and Baker, for instance, do not even attempt to define the concept. To be sure, the proponents do not seem to rely on a strange notion of education; for what they appear to mean agrees with what may be called the popular view, namely, that education is learning what is offered at school or college even if it is learned autodidacticly. The popular notion, however, is questionable, as I shall soon indicate. In any event, without a clear understanding of education one cannot be in a strong position to claim that a given computer is or is not educable or to construct a model of education by which to conceive an educable computer. Moreover, one cannot answer the charge by humanists and other critics (Haugeland 2000) that the very idea of an educable appliance is preposterous.

But what concept of education might apply to smart machines and at the same time not be stupid or stipulated? The popular idea seems acceptable as far as it goes. Individuals often do become educated by learning what is offered at schools and colleges. Yet, does one have to learn scholastic or collegiate subjects in order to be educated? Is it not presumptuous to accord scholastic and collegiate curricula a lock on educational subject matters? Did not Rousseau argue that the true education for Emile was his development according to nature rather than society? It is, of course, far from certain that a supercomputer could be conceived as developing according to nature, at least, as Rousseau understood the matter.

With no better approach in immediate view, we might as well try the idea of education that has been embedded in standard-English discourse for the past 75 years. That idea makes a good starting point because it is more or less familiar to humanists and others who have employed modern English in discussions of education and because it has been carefully explicated by linguistic philosophers. If we find that the concept is not applicable to appliances, no matter how competent, we then can revise it so as to make it applicable, thereby connecting the sense of AI education with the sense of education in pre-AI English discourse.

An Early Warning

Despite what has been argued elsewhere (Heslep 2009, pp. 329–349), the standard concept as it stands is not relevant to machines. The concept entails that only beings capable of consciousness and voluntary action are educable. But appliances, having no central nervous systems, are incapable of awareness (Searle 2011, p. 15a) And not being capable of acting unforcedly and with knowledge of their actions, they are incapable of voluntary action in the normal sense (Aristotle 1962, pp. 129). But by dropping the criterion of consciousness and making a few other adjustments, we can, without being arbitrary, formulate a notion of education that is applicable to a computer. It must be emphasized, however, that the proposed idea of education is not meant to replace the established one. Rather, it is intended to be only a special meaning for a technological purpose.

The Standard-English Concept

For analytic philosophy a concept is a set of criteria that determines the features an entity must have to be a member of a given class of objects (*The Cambridge Dictionary of Philosophy* 1995, p. l48). If, for example, the concept of a human being consists of the criteria of being rational and being an animal, then the concept pertains to an entity if and only if the entity has the features of being rational and being an animal. The conceptual criteria of education, accordingly, will enable us to tell if a projected computer is a member of the class of educable beings.

For the linguistic branch of analytic philosophy, the conceptual criteria ordinarily constituting a concept employed by a community of language users can be identified through an examination of the ways in which the members of that community normally speak and write. For that branch, then, the way to establish the standard conceptual meaning of education for the community of the users of the English language is to examine how they usually employ such terms as "education," "educational," "educated," and "educate."

During the past 50 years or so certain linguistic philosophers, prompted by Hirst and Peters (Hirst and Peters 1970; Peters 1973, pp. 11–84), have examined expressions by the users of ordinary English in order to identify the standard conceptual criteria of education. The criteria on which those philosophers have agreed are learning, the milieu of learning, a beneficent mode of existence, and the disposition to act voluntarily, or freely and knowingly, from such a mode. It is notable that neither being human nor learning what is offered at school or college is one of the criteria. In sum, everyday-English discourse regards education as learning, in a given context, a disposition to act voluntarily from a form of being that is worthwhile for the learner.

Learning implicates a learner and a content. A learner is a concrete agent. Numbers, ideas, corporations, and other abstract entities are spoken of as learners only in a figurative or derivative sense. Those referred to as learners without qualification are concrete individuals; and they are agents in that they listen, study, comprehend, practice, memorize, or do other things to acquire contents. But not all that can be gained, for example, body weight or a friend, is a learning content. Only statements ("learn that"), abilities ("learn how"), feelings ("learn to feel confident"), and dispositions ("learn to be friendly") are possible learning contents (Green 1971, pp. 121–144). According to normal-English usage, however, not just any learning is necessarily a matter of education; for it might not satisfy the other conceptual criteria. One can learn, for instance, to act involuntarily or to pursue a harmful way of life.

The circumstances of learning present and restrict what can be learned and how it can be. In the medieval period, for instance, one could learn to be a priest, a knight, or inter alia a shopkeeper but not a genetic scientist or computer engineer. But a milieu is educationally significant in a positive way only if it helps one learn to act voluntarily from a mode of being that is beneficent for oneself. So, while the life of a priest or shopkeeper might still be worthwhile today, that of a knight is unlikely to be.

As intended here, a form of being is a framework that contains general guidance and material support for the actions of an agent during its existence. The form consists of ideals, technologies, and practices, the last being social institutions or personal routines. The form may constitute, for example, the life of a primitive culture, a civilization, or a unique lifestyle. Thus, a mode of being is beneficent for a voluntary agent to the extent that it enables the agent to make free and knowing choices during the course of the agent's existence. The assessment of the worth of a form of existence for a prospective educable

being has to consider if the being will be capable of learning to act voluntarily on the basis of that framework and if the framework is likely to give guidance and material support for free and witting action in all foreseeable relevant types of situations. Moreover, the appraisal must rely on normative standards. Freedom and knowledge are the basic norms for determining the worthiness of any mode of existence for a voluntary agent, which means that other standards have to be consistent with them. Regardless, the norms employed may be stated as values, duties, rights, or virtues.

A voluntary agent usually is conceived as a free and knowing agent. It acts freely insofar as it does things unforcedly and without obstruction; and it acts wittingly to the extent that it is aware of its identity, its purposes, the available alternative means for achieving its purposes, and the foreseeable consequences of the means on which it decides. But an educated voluntary agent does not just have information about its actions; it also has good reasons for them. It chooses its purposes on the basis of defensible reasons, and it selects its means on rational grounds, for instance, effectiveness and efficiency. Accordingly, when an educated agent acts from a beneficent mode of existence, it looks to the mode's ideals as its major source of normative standards, to the mode's technologies as the source of its methodological and material means, and to the mode's practices as established routes of action.

The Features of an Educable Computer

Let us now see if a computer might have the characteristics that satisfy the conceptual criteria of education ordinarily intended by the users of modern English. Not just any kind of computer might have those traits. Presently, it is likely to be a supercomputer on the order of Deep Blue, the chess champion, and Watson, the winner on the TV show, "Jeopardy!" But the prospects look even brighter with the semantics-capable computers on the horizon of AI engineers (*Financial Times* March 4, 2008: 9).

To be educable, a computer must be capable of learning; and to be capable of learning, it must be a concrete agent. That a computer is a concrete entity is beyond dispute; and that it can do things, such as, store, retrieve, and process data, is also beyond question. But can it acquire learning content, that is, statements, abilities, feelings, or dispositions? Whether regarded as meanings expressed through declarative sentences or as cerebral events with cognitive meanings, statements are not independent entities. They ultimately are objects of consciousness, which is a state of the central nervous system. Because a computer does not have a central nervous system, it plainly cannot be aware of statements and therefore cannot learn that the sky is blue or any other statement. Also, because feelings are CNS dependent, they cannot exist in machines. A computer, accordingly, cannot learn to feel pity, sorrow, love, or anything else affective. It might, of course, learn to simulate emotional expressions.

What about abilities, then? A machine would not be a functioning computer if it could not store, retrieve, and process data. While a computer is endowed with the capability of computing, it acquires the ability to compute in a specific way when it is programmed, for instance, to play chess, make trades in the stock market, or greet arrivals at an airport. Not all acquired abilities, however, are learning contents. By virtue of adjustments to its engine, a truck can gain horsepower and thus an increase in its towing ability. But the truck, not having done anything to gain its new horsepower, does not learn its new ability. What, then, might a computer do in order to learn an ability? In the respect that a computer is downloaded with a program, the appliance is passive. Nevertheless, the computer, unlike

the vehicle with respect to its upgrade, may be active in relation to its programming. Because a computer might be programmed successfully or not, it may include in its operating system an algorithm by which it determines if it is capable of receiving the given program and then signal if it has acquired the program.

After a human has learned an ability, he or she might not exercise it again even in appropriate circumstances, thereby not being disposed to exercise it. But after a computer has acquired an ability, it will exercise it as long as it has an occasion to. Accordingly, being programmed to do certain things is sufficient for an intelligent machine to learn to be inclined to do them.

But even if a smart appliance can be a learner in the normal sense, it can be a voluntary agent only in a sense that departs considerably from the standard notion of that matter. Because a computer is bound in all respects by the laws of physics, it cannot act unforcedly any more than a rock can. But it can act freely in the respect that it can function without hindrance. In that regard it is comparable to a rock that can fall freely. So, while we readily concede that a computer cannot act unforcedly, we can maintain that it can act freely at least in the sense of without obstruction.

Another reason for departing from the ordinary idea of a voluntary agent is suggested by the fact that a machine does not have CNS. Absent CNS a computer is incapable of being aware of whatever program it has learned. It does not know what the point of the program is, and it does not know what the rules are. It does not even know that it has a program or that it is running one. Arguably, however, a computer's lack of cognitive consciousness is not educationally decisive. Unlike a truck an educable computer would not just gain an ability, for instance, to compute faster. It would learn a skill, a power to do something proficiently. In acquiring a skill, an agent learns to perform according to certain standards; but the agent also learns to make adjustments in its actions as circumstances require. It learns to follow criteria in trying to get things right (Ryle 1949, pp. 28–29). But an agent would not have to be cognizant of the criteria that it follows when exercising a specific skill. All that it would have to do would be to perform according to the criteria. In the case of a smart machine, of course, the criteria would be known to the programmer. Apparently, there already exist computers that know how to do certain things. Thus, according to Kurzweil (2005, p. 284), iHex, based on AI developed at Oxford University, not only can detect fraud in credit-card transactions and loan applications but can continuously update "its own rules based on its own experience."

So, while Searle (2011, p. A15) was correct in pointing out that Watson did not know that it won on "Jeopardy!" and, incidentally, might have added that Watson did not even know that it was a computer or had a name, Searle ignored the point that the appliance acted in a skillful manner. It did not have knowledge or understanding, but it competed skillfully. Like its human competitors it behaved intelligently despite the fact that, unlike them, it was not conscious of what it was doing. It, of course, did not know that it acted intelligently; but its programmer and audience did. Searle rightly insisted that Watson's success did not show that the computer was brighter than its human competitors, and he praised Watson's programmer for great ingenuity. But even if the machine was incapable of being bright, it certainly was more skillful at the game than its human competitors and even its brilliant programmer.

In sum, a computer can be a voluntary agent only in the sense that it can operate without obstruction and in a skillful way. For the nonce I will simply presume that this specific conception of a voluntary agent is defensible. I will try to justify it later.

What, then, would be the features of a computer that acts not only voluntarily but also rationally? There is a difference between acting from reason and acting according to it.

When one acts from reason, one acts by decisions that one reaches through reasoning, which is making inferences from reasons that one holds. By contrast, when one acts simply according to reason, one does not have to be capable of having reasons and drawing inferences from them; one just does what reason calls for. Because a smart machine, being without cognitive consciousness, cannot be aware of reasons let alone the rules of inference, it can perform rationally only in the sense of functioning according to reason. Reasons and inferences, of course, would be known to the programmer.

There are various modes of existence that an intelligent appliance might follow, for example, chess player, travel agent, stock trader, and laboratory technician. And there are several ways in which a form of being might benefit such an appliance. Thus, when programmed as a travel agent, a computer would do what any travel agent would do: schedule flights, reserve rooms and rental cars, recommend vacation spots, and receive a suitable compensation for its service. To have this existential mode, the computer would have to be programmed so as to have a list of flight schedules, available rooms, rental cars, and vacation spots; and it must have updates of these matters. Also, it would be encoded with rules by which its compensation would be determined and where and how the compensation would be deposited electronically by its clients. In addition, the machine would include prudential rules on the distribution of its accumulated wealth, for instance, as payments for its original cost, for maintenance service, and for program upgrades. Finally, the existential form of travel agent would benefit the computer of concern in that it would help guide and support the appliance in performing actions in all situations it would encounter, at least, as long as there would be a need for travel agents. With appropriate changes those comments could be made about computers following other manners of being.

To act rationally according to a beneficent existential mode, an educable computer must be able to act in conformity with that mode's ideals, technologies, and practices. Thus, when Deep Blue faced Gari Kasparov, the then-reigning chess champion of the world, it freely acted according to reasons derived from the ideals, technologies, and practices of chess. And if an educable computer followed the existential form of a laboratory technician, it would freely act according to reasons derived from the ideals, technologies, and practices of laboratory science (Kurzweil 2005, p. 283).

In sum, an educable computer is one that can learn to be disposed to act, without hindrance, with know how, and according to reason, on the basis of a beneficent mode of existence.

In Defense of the Revised Concept

Some humanists and other users of ordinary English might say that the idea advanced here is not so much a modification as it is a gross distortion of the normal understanding of education. To insist that an educated being might not be capable of acting unforcedly misses a huge point about education, which is that it prepares people to act on their own. In addition, no one speaking normal English over the centuries has ever intended that the freedom of an educated being was as insignificant as the free fall of rocks and other objects subject entirely to Newton's laws of motion and his law of gravity. Moreover, to allow that an educated being might be incapable of reasoning but may act according to reason flies in the face of all that is taken for granted about education, most notably, that it teaches learners how to think and act on the basis of good reasons. Indeed, to say that an educated computer is capable of acting not from reason but only according to reason is to lump

educable computers in with the dolts whom Aristotle (1950, p. 23) held were the only ones whose enslavement could be ethically justified. While those people could not reason, they could follow the rational instructions of their masters. The slaves would benefit the masters through their labor; and the masters, in addition to giving their slaves rational instructions, would provide them with material support. So maybe what the so-called revised concept does is to show that Deep Blue, Watson, and other supercomputers could be regarded as ethically justifiable slaves of their programmers or owners!

One may summarily dismiss the slavery issue. According to standard usage, machines are not and cannot be slaves. It is recognized that supercomputers are tools that are bought and sold and that slaves are tools that are bought and sold. But not all marketable tools are slaves. Slaves are human beings; hammers, screw drivers, plow horses, and supercomputers are not.

But one should also explain that the conceptual meanings of some terms in normal discourse have periodically changed. One of those terms is "education." What the users of English in a given period normally describe with "education" and its cognates depends, partly, on their conceptual criteria of education and partly on their habituated uses of such terms. According to *The Oxford English Dictionary* speakers and writers between 1650 and 1850 occasionally employed "education" to refer to the proper rearing or shaping of growing things, including children, silk worms, shrubbery, and beards. Around the middle of the nineteenth century they began to alter their conceptual criteria. An education began to be taken as the criterial conditions according to which a person learned subjects, whether useful or merely ornamental, that brought one recognition as a member of a socially superior group (Spencer 1963, pp. 1–44). In the second half of the nineteenth century, because of rapid industrialization, the increase of urbanization, scientific and technological advances, and the growth of specialization, people ceased to apply the term "education" to silk worms, shrubbery, and beards and began to speak and write of the all-round developed person as "the educated man" (Edel 1973, pp. 249–250).

Because, then, there are notable instances of the appearance of new conceptual meanings of "education," as well as the disappearance of once-established ones, there is nothing irregular in pushing a new concept of education in view of the enormous developments in AI in recent decades. The pre-AI idea of education applies to human beings and maybe to other kinds of animals (Heslep 2009, pp. 235–237) but certainly not to electronic tools. The proposed concept applies to such tools but not to humans or other animals. However, by including the criteria of learning, voluntariness even if diminished, rationality although shrunken, and a beneficent form of existence, the idea advocated here extends English discourse about education to the world of AI without rupture. The pre-AI and the AI concepts can co-exist in current English discourse without complications as long as speakers and writers do not confuse them.

Conclusion

By following the idea put forth, AI engineers will be in a position to knowingly design educable computers, which is more efficient and effective than designing computers with only the hope that they will turn out to be educable in some vague sense. Moreover, by relying on the concept of concern, AI engineers will be able to explain to humanists and other critics that they are designing computers that are educable in a sense that is specifically different from the sense in which humans are educable. Finally, AI engineers will be more circumspect before making claims about the superiority of supercomputers to

humans. They will not be in a position to hold that educable computers are superior in reasoning or understanding, for they will know that such machines cannot reason or understand. But they will be able to maintain rightly that educated machines, despite their lack of consciousness, are superior to humans in various skilled activities, for instance, playing chess, competing on game shows, and stock trading.

Meanwhile, by recognizing that educable computers share several features with educated people, the general public will realize that the arena of AI is not an alien or spooky place. Indeed, some might become ethically concerned with the treatment of intelligent appliances; for they might contend that such machines have a right become educated and, therefore, to learn appropriate skills for acting according to reason within the framework of a beneficent mode of life (Heslep 2009, pp. 345–346). They might even organize themselves under the banner of People for the Ethical Treatment of Artificial Intelligence (PETAI).

References

Aristotle. (1950). *Politics* (H. Rackham, Trans.). Cambridge, MA: Harvard University Press.
Aristotle. (1962). *N. ethics* (H. Rackham, Trans.). Cambridge, MA: Harvard University Press.
Baker, S. (2011). *Final jeopardy: Man vs machine and the quest for all knowledge*. New York: Houghton Mifflin Harcourt.
Edel, A. (1973). Analytic philosophy at the crossroads. In J. F. Doyle (Ed.) *Educational judgments: Papers in the philosophy of education* (pp. 232–258).
Green, T. E. (1971). *The activities of teaching*. New York: McGraw-Hill Book Company. Haugeland.
Haugeland, J. (2000). *Having thought: Essays in the philosophy of mind*. Cambridge, MA: Harvard University Press.
Heslep, R. D. (2009). Must an educated being be a human being? *Studies in Philosophy and Education, 28*, 329–349.
Hirst, P. H., & Peters, R. S. (1970). *The logic of education*. London: Routledge & Kegan Paul.
Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. New York: Penguin Books.
Peters, R. S. (Ed.). (1973). *The philosophy of education*. New York: Oxford University Press.
Ryle, G. (1949). *The concept of mind*. New York: Barnes & Noble.
Searle, J. (2011). Watson doesn't know it won on "Jeopardy!" *The Wall Street Journal*, February 23: 15A.
Spencer, H. (1963). *Essays on education and kindred subjects*. New York: Dutton.
The Cambridge Dictionary of Philosophy. (1995). New York: Cambridge University Press.