

Human-Level Artificial Intelligence? Be Serious!

Nils J. Nilsson

■ I claim that achieving real human-level artificial intelligence would necessarily imply that most of the tasks that humans perform for pay could be automated. Rather than work toward this goal of automation by building special-purpose systems, I argue for the development of general-purpose, educable systems that can learn and be taught to perform any of the thousands of jobs that humans can perform. Joining others who have made similar proposals, I advocate beginning with a system that has minimal, although extensive, built-in capabilities. These would have to include the ability to improve through learning along with many other abilities.

AI Goes to Work

The long-term scientific goal for many artificial intelligence (AI) researchers continues to be the mechanization of “human-level” intelligence—even though reaching that goal may be many years away.¹ Machines as intelligent as humans should be able to do most of the things humans can do. In this article, I propose a way of treating seriously this characterization of human-level AI. For those willing to continue toward this goal, in spite of its difficulty, I make some suggestions about ways to proceed.

Humans can “think,” and so also should any machine possessing human-level intelligence. Alan Turing claimed that it was too difficult to define thinking. Instead he proposed what has come to be called the “Turing test” (Turing 1950). To pass the test, a machine must be able, more often than not, to convince human participants in suitably staged anonymous dialogues that they are communicating with a human rather than with a machine. Many commenta-

tors feel, however, that the Turing test does not constitute an appropriate or useful criterion for human-level AI, and modifications, extensions, and restrictions of it have been proposed. (For a thorough discussion of the test and the controversies surrounding it, see the online article [Oppy and Dowe 2003] in the *Stanford Encyclopedia of Philosophy* [Zalta 2005].)

Machines exhibiting true human-level intelligence should be able to do many of the things humans are able to do. Among these activities are the tasks or “jobs” at which people are employed. I suggest we replace the Turing test by something I will call the “employment test.” To pass the employment test, AI programs must be able to perform the jobs ordinarily performed by humans. Progress toward human-level AI could then be measured by the fraction of these jobs that can be acceptably performed by machines.

Let me be explicit about the kinds of jobs I have in mind. Consider, for example, a list of job classifications from “America’s Job Bank.” A sample of some of them is given in figure 1:²

Just as objections have been raised to the Turing test, I can anticipate objections to this new, perhaps more stringent, test. Some of my AI colleagues, even those who strive for human-level AI, might say “the employment test is far too difficult—we’ll never be able to automate all of those jobs!” To them, I can only reply “Just what do you think human-level AI means? After all, humans do *all* of those things.”

Some people will counter that it would be achievement enough if those jobs requiring only “knowledge work” (such as a financial examiner) and not visual/motor coordination (such as a roofer’s helper) were automated. I agree, although there are those who claim that

vision, locomotion, and manipulation abilities are prerequisite even for knowledge work. Let's grant that point and include *all* jobs in the strong version of the employment test.

Others might say that achieving human-level intelligence is an inappropriate (perhaps even an unachievable) goal for AI. Many AI researchers prefer to work instead toward what has come to be called "weak AI"—which is focused more on building tools for *helping* humans in their work rather than on *replacing* them. To be sure, building machines that help humans is a laudable and important enterprise and motivates much excellent AI research. Still, achieving human-level AI (or "strong AI") remains the ultimate goal for some AI researchers.³

Finally, people might argue that complete automation of most human jobs won't ever happen even if it might someday be possible in principle. Dissenters might cite economic arguments (it would cost too much, and there are plenty of human workers anyway) or societal arguments (people need work to earn a living and to feel fulfilled).⁴ Furthermore, many of today's jobs will likely disappear—just as manufacturing buggy whips did. Nevertheless, the *scientific* goal of mechanizing true human-level intelligence implies at least the *potential* of complete automation of economically important jobs (even if for one reason or another the potential is never activated).

The Challenge

For the purposes of the employment test, we can finesse the matter of whether or not human jobs are *actually* automated. Instead, I suggest, we can test whether or not we have the *capability* to automate them. One way to decide that matter would be to have machines take examinations in vocational courses designed to teach and test the skills needed by the jobs in question.

There are correspondence schools offering online courses and exams matched to many jobs. One such school is "Education Direct." Its website (Education Direct 2005) lists more than 80 distance learning programs, each of which offers online exams to measure progress and competence in the skills taught. A passing grade requires a score of 70 percent or higher on each exam.

Passing a specified exam (or group of exams) by an AI program can be taken as evidence that the corresponding job can be automated.⁵ (Because many of the jobs require perceptual and motor skills, the exams would have to be augmented to measure these skills also.) As the number of tests passed by AI programs increases, so correspondingly, I would claim, does AI

Meeting and Convention Planner
Maid and Housekeeping Cleaner
Receptionist
Financial Examiner
Computer Programmer
Roofers Helper
Library Assistant
Procurement and Sales Engineer
Farm, Greenhouse, Nursery Worker
Dishwasher
Home Health Aide
Small Engine Repairer
Paralegal
Lodging Manager
Proofreader
Tour Guide and Escort
Geographer
Engine and Other Machine Assembler
Security Guard
Retail Salesperson
Marriage and Family Counselor
Hand Packer and Packager

Figure 1. List of Job Classifications from "America's Job Bank."

progress toward "human-level intelligence."

Of course, I am not advocating here the use of any particular correspondence school or of any particular list of exams. I cite the exams of Education Direct as only one example that could be used. Undoubtedly there are many others that might serve as well or better.

The brute-force way for AI to make progress toward automating human jobs would be to develop special-purpose programs matched to each job. Such an approach would require the invention of hundreds, perhaps thousands, of specialized "niche" systems. Computer science has already produced many such systems, and these have substantially automated many routine human jobs. But there are many more to go, and the remaining ones require increasingly sophisticated levels of human capabilities. Approaching the problem in this brute-force manner seems inefficient and likely to produce systems lacking the ability to adapt to changes (which is even a requirement for some of the jobs).

Habile Systems

Humans who perform these various jobs aren't born with the skills their jobs require. They have to learn them—perhaps by taking courses from correspondence schools or by on-the-job training. Ignoring for the moment the fact that humans are born with different aptitudes and different inherent abilities to acquire “intelligence,” we can nevertheless think of humans as being more-or-less general purpose—able to learn any of a wide variety of different skills. Analogously, I think AI should strive toward building a small number of general-purpose machines that are able to learn and to be taught skills (and to improve on them) rather than programming a much larger number of machines each possessing a different individual skill set from the start. In an earlier paper, I called such general-purpose machines “habile systems” (Nilsson 1995).⁶ Ed Feigenbaum (2003) called for something similar in his “Grand Challenge Number 2.” We would expect habile systems to be able, for example, to take and pass vocational correspondence courses from schools like Education Direct.

If we take this habile-system path toward the main goal of achieving human-level AI, we are immediately confronted with the subgoal of building AI programs that can learn job skills. The subgoal is no easier than the main goal, but at least we have to build only one such system (perhaps a few) rather than having to build a separate one for each of the thousands of different jobs. Working on this subgoal would involve a major shift in direction for long-term AI research. Instead of “applying AI” to the many, many problems for which some of its well-honed techniques are suited, some researchers, at least, would be focusing their efforts toward building habile systems.

Before discussing the prospects for building habile systems, it's worth reminding ourselves about the variety and difficulty of the things such systems must be able to learn. Education Direct's course on “Small Engine Repair,” for example, claims to teach its students how to troubleshoot and repair engine components and systems; work with engines, ignition systems, and electrical circuits; inspect and repair small engine fuel systems; and disassemble and rebuild all types of small engines, including two-stroke, four-stroke, outboard, and riding mower engines. Its course on “Paralegal Studies” claims to teach its students about legal terminology and thinking skills, interpersonal communications and ethics, legal writing, research, and civil and criminal litigation, family law, real estate law, and wills and estates. A tall order? Yes, but we said we were striving toward *human-level AI*, didn't we?

As I have defined them, habile systems should be able to be taught to perform various human jobs. I don't expect building habile systems to be easy or that they will be achievable in the next several years. They will take much research and experimentation. We will have to use our best intuitions about what pretutoring abilities they will need to have.

The fact that habile systems will need to learn to become specialists suggests that perhaps we should try the strategy of teaching somewhat less capable systems to become the habile systems we are seeking. Perhaps it will be easier to program the less able systems. Suppose, for example, that we were able to build a system that could take and complete high school or college courses. Thus educated, the resulting system could go on to take one of the specialized courses I mentioned earlier. Let's continue in this vein. If building systems capable of taking high school courses proves to be too difficult, as I suspect it will be, why not try to build systems that would be able to benefit from primary education?⁷ And so on, all the way back to teaching an infant.

I don't mean to imply that, after achieving a system at any one of these developmental levels, we should rely *only* on that system's experience, its learning abilities, and human tutoring to progress to the next level. AI has produced some very sophisticated and powerful techniques for perception, knowledge representation, language analysis, planning, search, learning, and reasoning. And, we can expect more to come. Such of this technical apparatus as researchers think appropriate and useful should be inserted at those stages of system development best able to utilize them. Indeed, researchers will want to examine the state of a system at any stage and debug and reprogram aspects of it as needed.

That is, we don't need to slavishly follow the developmental/tutoring approach. At any stage we can add, by programming, abilities we know will be important. We can skip steps. We can ignore abilities we think aren't needed for the ultimate goal. (We can backtrack to add these if we guess incorrectly.) Because the developing human is an existence proof of what we are trying to accomplish, we can always go back to it as reference.

The “Child Machine”

Human learning and skill acquisition begin at birth (perhaps before). The process feeds on itself. We learn how to learn, and we acquire skills that are employed in the service of acquiring more skills, and so on. Some of this learn-

ing is by imitation, some by trial and observation of consequences, some by being taught, and some by listening to and reading what others have to say. The more we know and the more skilled we are, the more we can learn.

Here is how I look at this process: The “intelligence” of a mature adult develops in layers, like those of a growing onion. At the core is a set of very sophisticated and powerful, let us say hard-wired, basic mechanisms provided by our evolutionary heritage. This core is able to build the next layer—providing the core is immersed in an appropriate environment. Together, they build the next layer, and so on.

If building “outer-layer” systems directly proves too difficult, perhaps building the core will be easier. First, we must decide what goes in it, and then we let this core loose in the world and teach and program it through the various stages I have already mentioned. This idea is not new with me. Ben Wegbreit,⁸ Rod Brooks (1996), John McCarthy (1999), Jeff Hawkins (Hawkins and Blakeslee 2004), and perhaps others, have suggested or discussed similar approaches. It is interesting to recall, that at the end of his famous 1950 paper, which described his eponymous test, Alan Turing proposed what has come to be called “the child project.” Here is what he said (1950, p. 456):

Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child’s? If this were then subjected to an appropriate course of education one would obtain the adult brain.... Our hope is that there is so little mechanism in the child-brain that something like it can be easily programmed. The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child.

We have thus divided our problem into two parts. The child-programme and the education process. These two remain very closely connected. We cannot expect to find a good child-machine at the first attempt. One must experiment with teaching one such machine and see how well it learns. One can then try another and see if it is better or worse.

Turing’s hope that something like the child-brain “can be easily programmed” was, obviously, overly optimistic. I believe the initial core will be quite complex indeed. Furthermore, the core and subsequent layering must be “subjected to an appropriate course of education.”

Turing was more prescient when he said “we cannot expect to find a good child-machine at the first attempt.” AI is in for many years of trials and experiments. Even so, I have concluded that coming up with a serviceable core, appropriate educational environments, and inspired programming provides the best path to hu-

man-level AI. I think the outer layers of the onion would be even harder for us to program directly than will be the core. To appreciate the scale of effort and difficulties we would face in these outer layers, we need only look at the CYC project (Lenat 1995)—one of the few attempts to program anything like a human-level commonsense knowledge base.

Core Suggestions

The core we are looking for must have many built-in capabilities. It will be the first task of this line of research to figure out what these should be and to begin trial implementations. John McCarthy’s thoughtful memo (McCarthy 1999) provides some foundational ideas. An article by Marvin Minsky, Push Singh, and Aaron Sloman (2004) reports on a symposium devoted to architectures for human-level intelligence. I do not pretend to have expert advice on this matter but will nevertheless offer some observations and cite some work that looks to me like it’s on the right path.

Sensory-Motor System

The core sensory-motor system should include, at least, capabilities similar to those possessed by a typical human infant or child. The inputs should include visual, tactile, audio, and haptic sensory mechanisms. A suite of basic perceptual programs, perhaps influenced by a suite of built-in goals, must be provided to process these inputs. Motor outputs should include manipulation, vocalization, visual display, and locomotion. A suite of basic activation programs—suitably responsive to the perceptual system and goals—will provide the initial means to control these outputs. Early experience in a suitable environment will provide opportunities for a learning system to add to these built-in abilities.

Some might ask, why not include additional, nonhuman, sensory modalities and motor capabilities, such as infrared, x-ray, laser beams, arc welding, and so on? After all, many human jobs require apparatus such as these. My answer is that, in the spirit of the enterprise, we need include only those basic abilities that will, with experience, practice, and elaboration, be sufficient to enable an agent to operate, as tools, any needed additional sensory and motor systems—much as humans do.

The sensory system must also be able to remember generalized versions of frequently occurring perceptual ensembles as “patterns” and to build programs that can recognize these patterns when it encounters them in the future. Human development of such pattern-recogni-

tion abilities, in its earliest stages at least, appears to involve building, rewiring, and perhaps ablating neural circuitry. Analogous processes may be useful in the development of habile systems.

Hierarchies of Perception, Representation, and Action

The core system should have the means to represent abstraction hierarchies of relational models (using a vocabulary of basic and learned perceptual predicates) and hierarchies of motor routines. I have in mind here something like the real-time control system (RCS) of James Albus (1997), blackboard systems (Englemore and Morgan 1988), and my triple-tower architecture (Nilsson 2001). What is needed is the ability to generate automatically such hierarchies based on experience and teaching. Gary Drescher (1991) proposed some interesting ideas for producing Piagetian hierarchies; Karl Pflieger (2002) has done work on generating perceptual hierarchies; and Jeff Hawkins (Hawkins and Blakeslee 2004) claims that his proposed model of the human neocortex can build perception/action hierarchies. I think work along these lines is quite promising.

Presumably much of the early experience of the core habile system will involve perception of spatial attributes (such as adjacent, above, inside, and so on) and motor actions in space (grasp, move forward, turn, and so on). Many of the needed higher-level perceptual predicates and motor routines might then best be represented in a way that exploits analogies to these spatially grounded basic experiences. (George Lakoff and Mark Johnson [1980] observe that sentences such as “We hit a peak last year, but it’s been downhill ever since” indicate that human abstract notions exploit such lower-level “orientational” metaphors, along with many others.)

Predicting and Planning

The core system needs the ability to make predictions of future perceived states that will result from proposed actions taken in present states. These predictions form the basis for making plans of actions to achieve goals. A system for learning to predict the effects of its actions and for making plans based on these predictions was developed by Scott Benson (1996). The ability to make predictions is a critical component of intelligence and is a key feature underlying Hawkins’s model of the cortex (Hawkins and Blakeslee 2004). Differences in the reward values of states predicted and states actually encountered are fundamental to temporal-difference learning methods for improv-

ing prediction accuracy both in machines (Sutton 1988) and in animals (Montague, Dayan, and Sejnowski 1996).

Regarding planning, McCarthy (1999) speculates that infants have a built-in capacity for recognizing that the preconditions for an action that achieves a goal should be pursued as subgoals. He calls this “goal regression.” Goal regression is fundamental to many AI planning systems.

Learning

The essence of a habile system is its ability to grow and change by learning, through imitation, experience, practice, and education—augmented by being reprogrammed when that is convenient. Turing (1950, p. 457) anticipated the need for what we now call reinforcement learning (Sutton and Barto 1998):

We normally associate punishments and rewards with the teaching process. Some simple child-machines can be constructed or programmed on this sort of principle. The machine has to be so constructed that events which shortly preceded the occurrence of a punishment-signal are unlikely to be repeated, whereas a reward-signal increased the probability of repetition of the events which led up to it.

At some stage in the development of a habile system, its reinforcement learning mechanisms must be able to work with the perceptual, representation, and action hierarchies mentioned earlier (Barto and Mahadevan 2003) and with relational descriptions of states (Dzeroski, De Raedt, and Driessens 2001). These remain important areas for future research.

Reasoning and Representation

Representation of knowledge and reasoning with that knowledge has engaged the attention of AI research since its beginnings. We know that sound and complete reasoning using logical methods is intractable if it is to proceed from the large collection of statements required to represent all of the knowledge a competent human must possess. Yet, many human tasks require effective reasoning. One strategy for dealing with an otherwise impracticably large knowledge base is to divide it into weakly interacting, semi-independent parts. Eyal Amir (2002) is one of those who has proposed ways to do this.

AI has had rather impressive success with reasoning systems limited to specialized areas of knowledge—areas that McCarthy calls “bounded informatic situations” (McCarthy 1996). McCarthy claims that in order to realize human-level artificial intelligence, reasoning systems will have to deal with the “common sense informatic situation—in which the phe-

nomena to be taken into account in achieving a goal are not fixed in advance.” Such systems, he says, must be capable of (among other things) nonmonotonic reasoning, introspection, and reasoning about contexts and with approximate concepts. Designers of a child machine will need to decide what seeds of these abilities must be built in.

Many of the conclusions humans draw depend on uncertain knowledge. AI has made major progress, again in limited domains, in reasoning with uncertain information. Perhaps further study of how humans represent knowledge and come to conclusions will provide useful hints for building the representation and reasoning capabilities required by habile systems.

Language

Linguistic abilities will be needed for several purposes: for communication among habile agents, for communication with humans, for learning by reading, and (perhaps) for internal representation and reasoning. And, of course, for taking the courses and the exams! Explicit information given by linguistic commands or suggestions gives more precise direction than that given implicitly by rewards and punishments. Turing (1950, p. 457) predicted:

The use of [a symbolic] language will diminish greatly the number of punishments and rewards required.

Just how much of the apparatus underlying human language ability is built-in and how much is learned is still debated. Researchers in natural language processing by computers similarly argue about the potential for machine-learning techniques in their work. Perhaps in the early period of habile systems research, an invented language with unambiguous syntax and semantics will suffice. Because our strategy depends on education by humans and by human-generated course materials, we will have to hope for progress by natural language processing researchers. Much of that progress might well depend on the success of habile system research.

Conclusions

I have proposed specific, testable criteria for human-level artificial intelligence. Systems with true human-level intelligence should be able to perform the tasks for which humans get paid. I don't see any reason, in principle, why such systems cannot be built, although they may be many years away. In any case, people will continue attempts to automate human jobs because hardware and software to perform them will ultimately be less expensive than human

salaries and benefits. Even if jobs are not automated for economic, social, or other reasons, we can nevertheless measure progress toward human-level AI by having candidate systems take performance tests that measure the requisite abilities for these tasks.

Considerable effort in AI and computer science has already been devoted to crafting specialized systems for automating some human jobs. Along the way, some very powerful and general AI technologies have been developed. Yet, chipping away in this manner seems an inefficient way to proceed toward automating all human jobs.

There is an alternative approach, namely the development of general-purpose systems that can be adapted to the different specialized tasks. My version of this alternative is to build “habile systems,” ones that can learn and be taught to perform the tasks that humans can perform. The plan would be to create the required specialized systems by subjecting a habile system to a training regime, much as a human learns needed skills by professional or vocational training.

Again, there are options. We could attempt to build such a habile system directly. At some point in the future, researchers may feel capable of doing that. But because the habile system must be educated anyway, why not educate “less mature” versions of a habile system—ones that need to be subjected to the equivalent of a high school or college education before being able to benefit from vocational training. Proceeding another step or two down this spectrum, how about building systems capable of elementary and preschool training?

Carrying this notion to its extreme, various researchers have suggested building a system that has the same built-in facilities for learning that human infants have and carrying on from there.⁹ My guess is that this so-called child machine will be the easiest starting place along this spectrum—although still *very* difficult.

Of course, the habile-system approach toward human-level AI can be entered at whatever level or levels researchers feel confident about. Perhaps multiple entry points should be tried. And, of course, the skills and knowledge gained by a system's education and experience, anywhere along the spectrum, can be augmented (or maybe even overshadowed) by whatever the then-current AI technology can provide. One can also hope that attempts to broaden the expertise of specialized systems beyond their “bounded informatic situation” will produce useful components of habile systems.

If the research program I have outlined is followed successfully, the performance of human

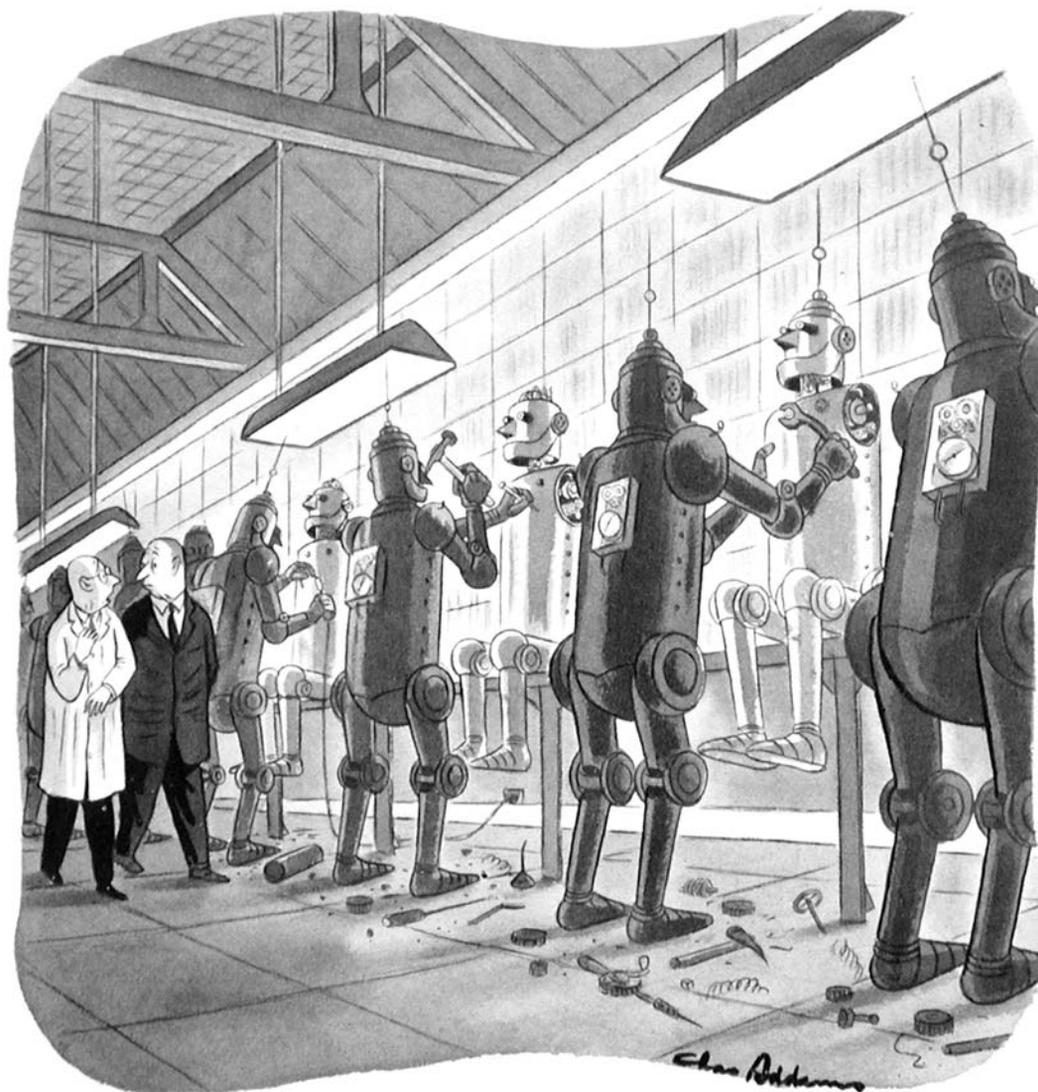


Figure 2. "Sometimes I Ask Myself, 'Where Will It Ever End?'"

Charles Addams, *The New Yorker*, 2/9/1946. © Tee and Charles Addams Foundation. Used with permission.

jobs by machines might not, in the end, be as important as what the results of this research tell us about our own intelligence. Do I want to see all human jobs automated? Not necessarily. I'm not as interested in manufacturing intelligence as I am in understanding it. But I think we'll need to manufacture it before we can understand it.

Turing's conclusion to his 1950 paper (1950, p. 460) seems equally fitting here:

We can only see a short distance ahead, but we can see plenty there that needs to be done.

Acknowledgements

I thank Mykel Kochenderfer, John McCarthy, Ed Feigenbaum, Tom Dietterich, Alon Halevy, David McAllester, Grace Abbott, and an anonymous referee for comments and for pointing out relevant citations.

Notes

1. Ed Feigenbaum concludes that "computational intelligence" is the "manifest destiny" of computer science (Feigenbaum 2003).
2. There are more than 1,500 jobs in the complete list, which is available at www.jobsearch.org/help/employer/SSONetJobCodeListbyCategory2.html. Doubtless there are many other such lists.
3. The terms *weak* (or *cautious*) and *strong AI* were first used (slightly differently) by John Searle: "According to weak AI, the principal value of the computer in the study of the mind is that it gives us a very powerful tool.... [According to strong AI] the appropriately programmed computer really is a mind." (Searle 1980).
4. For a discussion of some of these issues, see Nilsson (1984).
5. An anonymous referee noted that passing a vocational exam does not necessarily imply that the associated job can be performed by the person or ma-

chine taking the test. Acceptable job performance also requires commonsense, which vocational exams may not adequately measure—even though understanding and answering the exam questions, presumably stated in natural language, would require extensive commonsense knowledge. I acknowledge that point, but nevertheless maintain that the score on a vocational exam is a reasonable predictor of ability to perform the associated job.

For examples of AI programs taking and passing chemistry examinations, see Friedland et al. 2004.

6. The word *habile* means “generally able or adroit” and brings to mind one of the earliest tool users, *homo habilis* (“handy man”).

7. Online materials and examinations exist for these “preparatory” studies also. (See, for example, www.educationdirect.com/diploma/index.html and www.etap.org/home_files/how_to_teacher.htm#Elementary.)

8. B. Wegbreit, private communication, circa 1998.

9. There are even more “extreme” approaches. Instead of following the ontogenetic progression outlined here, various authors have proposed working along a phylogenetic spectrum—recapitulating the evolutionary steps toward intelligence. See, for example, Brooks (1997). Others propose simulating evolution itself. See, for example, Koza, et al. (2003).

References

- Albus, J. S. 1997. The NIST Real-Time Control System (RCS): An Approach to Intelligent Systems Research. *Journal of Experimental & Theoretical Artificial Intelligence* 9(2–3): 157–174, April 1.
- Amir, E. 2002. Dividing and Conquering Logic. Ph.D. dissertation, Department of Computer Science, Stanford University, Stanford, CA. (Available at www.cs.uiuc.edu/~eyal/papers/phd-thesis2001.pdf.)
- Barto, A. G., and Mahadevan, S. 2003. Recent Advances in Hierarchical Reinforcement Learning. *Discrete Event Dynamic Systems: Theory and Applications* 13: 41–77. (Available at www-anw.cs.umass.edu/pubs/2003/barto_m_DEDS03.pdf.)
- Benson, S. 1996. Learning Action Models for Reactive Autonomous Agents. Ph.D. dissertation, Department of Computer Science, Stanford University, Stanford, CA. (Available at www.robotics.stanford.edu/users/nilsson/trweb/bensonthesis.pdf.)
- Brooks, R. A. 1996. Prospects for Human Level Intelligence for Humanoid Robots. In *Proceedings of the First International Symposium on Humanoid Robots* (HURO-96). Waseda University, Tokyo, Japan, October 30–31. (Available at people.csail.mit.edu/u/b/brooks/public_html/papers/prospects.pdf.)
- Brooks, R. A. 1997. From Earwigs to Humans. *Robotics and Autonomous Systems* 20(2–4): 291–304, June. (Available at people.csail.mit.edu/u/b/brooks/public_html/papers/ascona.pdf.)
- Drescher, G. 1991. *Made-up Minds: A Constructivist Approach to Artificial Intelligence*. Cambridge, MA: MIT Press.
- Dzeroski, S.; De Raedt, L.; and Driessens, K. 2001. Relational Reinforcement Learning. *Machine Learning* 43(1–2): 7–52.
- Education Direct 2005. Education Direct Website. Stamford, CT: Thomson Learning, Inc. (Available at www.educationdirect.com/programs_cat.html.)
- Engelmore, R., and Morgan, T., eds. 1988. *Blackboard Systems*. Menlo Park, CA: Addison-Wesley Publishing Company.
- Feigenbaum, E. A. 2003. Some Challenges and Grand Challenges for Computational Intelligence. *Journal of the ACM*, 50(1): 32–40, January.
- Friedland, N. S., et al. 2004. Project Halo: Towards a Digital Aristotle. *AI Magazine* 25(4): 29–47, Winter.
- Hawkins, J., and Blakeslee, S. 2004. *On Intelligence*. New York: Times Books.
- Koza, John R., et al. 2003. *Genetic Programming IV: Routine Human-Competitive Machine Intelligence*. Norwell, MA: Kluwer Academic Publishers.
- Lakoff, G., and Johnson, M. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Lenat, D. B. 1995. CYC: A Large-Scale Investment in Knowledge Infrastructure. *Communications of the ACM* 38(11): 33–38, November. (Available at web.media.mit.edu/~lieber/Teaching/Common-Sense-Course-02/Lenat-CACM.pdf.)
- McCarthy, J. 1996. From Here to Human-Level Intelligence. Unpublished memo, Department of Computer Science, Stanford University, Stanford, CA. (Available at www-formal.stanford.edu/jmc/human.html.)
- McCarthy, J. 1999. The Well-Designed Child. Unpublished memo, Department of Computer Science, Stanford University, Stanford, CA. (Available at <http://www-formal.stanford.edu/jmc/child.pdf>.)
- Minsky, M.; Singh, P.; and Sloman, A. 2004. The St. Thomas Common Sense Symposium: Designing Architectures for Human-Level Intelligence. *AI Magazine* 25(2): 113–124, Summer.
- Montague, P. R.; Dayan, P.; and Sejnowski, T. J. 1996. A Framework for Mesencephalic Dopamine Systems Based on Predictive Hebbian Learning. *Journal of Neuroscience* 16(5): 1936–1947.
- Nilsson, N. 1984. Artificial Intelligence, Employment, and Income. *AI Magazine* 5(2), Summer. (Available at robotics.stanford.edu/users/nilsson/OnlinePubs-Nils/GeneralEssays/AIMag05-02-002.pdf.)
- Nilsson, N. 1995. Eye on the Prize. *AI Magazine* 16(2): 9–16, Summer. (Available at <http://robotics.stanford.edu/users/nilsson/OnlinePubs-Nils/GeneralEssays/AIMag16-02-002.pdf>.)
- Nilsson, N. 2001. Teleo-Reactive Programs and the Triple-Tower Architecture. *Electronic Transactions on Artificial Intelligence* 5B: 99–110. (Available at <http://www.ep.liu.se/ej/etai/2001/006/>.)
- Oppy, G., and Dowe, D. 2003. The Turing Test. In *The Stanford Encyclopedia of Philosophy*, ed E. Zalta (Summer 2003 Edition). (Available at <http://plato.stanford.edu/archives/sum2003/entries/turing-test/>.)
- Pfleger, K. 2002. On-Line Learning of Predictive Compositional Hierarchies. Ph.D. dissertation, Department of Computer Science, Stanford University, Stanford, CA. (Available at <http://www.ksl.stanford.edu/people/kpfleger/dissertation/download/>.)
- Searle, J. R. 1980. Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3(3): 417–57. (Available at <http://members.aol.com/NeoNoetics/MindsBrainsPrograms.html>. Also in *Mind Design II*, ed. J. Haugeland. Cambridge, MA: MIT Press, 1997.)
- Sutton, R. S. 1988. Learning to Predict by the Methods of Temporal Differences. *Machine Learning* 3(1): 9–44, August.
- Sutton, R. S., and Barto, A. G. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press.
- Turing, A. M. 1950. Computing Machinery and Intelligence. *Mind* LIX(236): 433–460, October. (Available at www.abelard.org/turpap/turpap.htm#N_3_.)
- Zalta, E. (ed.) 2005. *Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. (Available at plato.stanford.edu/.)

Nils J. Nilsson is Kumagai Professor of Engineering, Emeritus of Stanford University. He is the author or coauthor of several AI textbooks and many of the seminal papers in the field. His career in AI began at SRI International in 1961 and continued at Stanford University where he was the chair of the Department of Computer Science from 1985 to 1990. Nilsson is a Fellow and a past president of AAAI. He received the IJCAI award for Research Excellence in 2003 and the AAAI Distinguished Service award in 2005. He can be reached at nilsson@cs.stanford.edu.