Available online at www.sciencedirect.com
ScienceDirect
Artificial Intelligence 170 (2006) 1251–1253
Artificial Intelligence
www.elsevier.com/locate/artint

Evolution of artificial intelligence

Lee Spector

Cognitive Science, Hampshire College, Amherst, MA 01060, USA

Available online 7 November 2006

The goal of creating non-biological intelligence has been with us for a long time, predating the nominal 1956 establishment of the field of artificial intelligence by centuries or, under some definitions, even by millennia. For much of this history it was reasonable to recast the goal of "creating" intelligence as that of "designing" intelligence. For example, it would have been reasonable in the 17th century, as Leibnitz was writing about reasoning as a form of calculation, to think that the process of creating artificial intelligence would have to be something like the process of creating a waterwheel or a pocket watch: first understand the principles, then use human intelligence to devise a design based on the principles, and finally build a system in accordance with the design.

At the dawn of the 19th century William Paley made such assumptions explicit, arguing that intelligent designers are necessary for the production of complex adaptive systems. And then, of course, Paley was soundly refuted by Charles Darwin in 1859. Darwin showed how complex and adaptive systems can arise naturally from a process of selection acting on random variation. That is, he showed that complex and adaptive design could be created without an intelligent designer. On the basis of evidence from paleontology, molecular biology, and evolutionary theory we now understand that nearly all of the interesting features of biological agents, including intelligence, have arisen through roughly Darwinian evolutionary processes (with a few important refinements, some of which are mentioned below).

But there are still some holdouts for the pre-Darwinian view. A recent survey in the United States found that 42% of respondents expressed a belief that "Life on Earth has existed in its present form since the beginning of time" [7], and these views are supported by powerful political forces including a stridently anti-science President. These shocking political realities are, however, beyond the scope of the present essay.

This essay addresses a more subtle form of pre-Darwinian thinking that occurs even among the scientifically literate, and indeed even among highly trained scientists conducting advanced AI research. Those who engage in this form of pre-Darwinian thinking accept the evidence for the evolution of terrestrial life but ignore or even explicitly deny the power of evolutionary processes to produce adaptive complexity in other contexts. Within the artificial intelligence research community those who engage in this form of thinking ignore or deny the power of evolutionary processes to create machine intelligence.

Before exploring this complaint further it is worth asking whether an evolved artificial intelligence would even serve the broader goals of AI as a field. Every AI text opens by defining the field, and some of the proffered definitions are explicitly oriented toward design—presumably design by intelligent humans. For example Dean et al. define AI as "the design and study of computer programs that behave intelligently" [2, p. 1]. Would the field, so defined, be served by the demonstration of an evolved artificial intelligence? It would insofar as we could study the evolved system and particularly if we could use our resulting understanding as the basis for future designs. So even the most design-oriented AI researchers should be interested in evolved artificial intelligence if it can in fact be created.

E-mail address: lspector@hampshire.edu (L. Spector).

0004-3702/$ – see front matter
doi:10.1016/j.artint.2006.10.009

Most AI texts nonetheless demonstrate, by their coverage, that their authors view AI as a set of design problems that human designers are expected to solve. Few describe evolutionary perspectives on the development of natural or artificial intelligence, and most describe existing evolutionary methods (usually just genetic algorithms) only as an odd form of search or learning.[1] But these same texts are generally less committed to human design in their introductory definitions. For example Winston states that AI is "the study of ideas that enable computers to be intelligent" [13, p. 1], while Charniak and McDermott say that AI is "the study of mental faculties through the use of computational methods" [1, p. 6].[2] Neither of these definitions seems particularly wedded to design. If they, and others like them, truly capture the over-arching objectives of the field then it would seem that evolved intelligence—the only kind of general intelligence of which we currently have examples—should be of central interest, as should be the conditions under which intelligence can evolve.

So AI researchers should be interested in evolved artificial intelligence in principle, but is there any reason to believe that evolution can produce artificial intelligence in practice? Traditional genetic algorithms have not done so, at least not in any general sense, but evolutionary computation is rapidly advancing, fueled in part by revolutions in our understanding of biological evolution.

Darwin, of course, didn't have the whole story on biological evolution, and neither did the developers of the first genetic algorithms. Darwin was ignorant not only of the relevant molecular biology but also of population and gene dynamics and of many other mechanisms and principles that are central to a modern understanding of evolution. We now know, for example, that biological evolution depends in important ways on genetic representation, on ecological interactions within and among populations, and on genetic control of gene expression, reproductive strategies, development, and learning. None of these features of biological evolution were incorporated into the first evolutionary computation systems—ideas for which, incidentally, date at least to Turing [12]—but serious work has been devoted to several of them in recent years.

The problem-solving performance of evolutionary algorithms has advanced significantly in the past decade or so, to the extent that human-competitive results have recently been achieved in several areas of science and engineering; these include evolved designs for antennas [4], photonic crystals [9], quantum computer algorithms [11], and even search heuristics [3].[3] Many of these results have been achieved using systems that incorporate at least a few of the insights from recent biological advances—for example, many involve the genetic representation of developmental processes instead of "adult" phenotypes—but there is still a long way to go if the computational work is to catch up with biology. Nonetheless, work in evolutionary computation is moving rapidly forward and it is doing so within an increasingly mature and stable research community. One indication of this progress is the recent establishment, by the Association for Computing Machinery, of a full-fledged Special Interest Group on Genetic and Evolutionary Computation (SIGEVO).[4]

None of this means that we should expect a simple extension of current evolutionary algorithms to produce general artificial intelligence any time soon. This is in part because the field of evolutionary computation has proceeded largely in isolation from developments in mainstream AI that will surely undergird future intelligent systems. But the shortest path to achieving AI's long-term goals will probably involve both human design and evolution.

It is not yet clear what sorts of interactions between design and evolution will prove to be most helpful or from which research communities they will emerge. It is possible that the long-established tradition of analyzing evolution as a form of search will continue to flower (as in [8]), to the extent that it drives a broader integration of evolutionary thinking and mainstream AI. Or perhaps the integration will be driven by successful evolutionary approaches to mainstream AI problems such as RoboCup [5]. Or maybe the integration will grow out of work in machine learning, an area that has traditionally provided formal models of evolution and which may also provide mechanisms to manage adaptation over multiple time scales (encompassing processes normally described as learning, development, and evolutionary change).

[1] Nilsson's 1998 text is a notable exception; it adopts an explicitly "evolutionary" approach and includes a full chapter on "Machine Evolution" [6].

[2] Several other definitions are surveyed and categorized by Russell and Norvig in the introduction to their popular text [10].

[3] The annual "Humies" awards recognize human-competitive results of genetic and evolutionary computation on the basis of objective criteria such as patents and publications. Information on past recipients and future competitions can be found at http://www.human-competitive.org/.

[4] http://sigevo.org/.

Whenever and wherever such integration occurs it may take a variety of forms. For example, future developers might use representations and algorithms that have been developed in mainstream AI as the ingredients in the "primordial ooze" that is sampled and recombined by evolution. Alternatively, they might use human-designed algorithms to intelligently compose components that have been devised and refined by evolutionary methods. Or they may combine design and evolution in ways that nobody has yet imagined.

In any event the prevailing assumption that general artificial intelligence can only be (or should only be) designed by intelligent human designers is flawed and should be rejected. That this pre-Darwinian assumption is indeed prevalent is demonstrated not only by the lack of evolutionary perspectives in AI texts but also by the niche treatment of evolution as a special purpose search or learning algorithm in most of the mainstream AI research literature.

A more appropriate post-Darwinian research strategy would consider evolutionary methods for any problem domain in which human design is difficult, and would expect evolutionary methods to be increasingly useful in increasingly general problem environments. Even when intelligent human design is appropriate, progress toward the field's over-arching goals would be served by situating the work within a larger evolutionary context.

The field has long acknowledged that intelligent behavior might best be achieved by agents that learn—agents that grow or redesign themselves to some limited extent as they confront their environments. But modern biology and technological advances both support a more radical offloading of design from humans to the intelligent systems themselves and to their ecologies. Not only should the AI systems of the future grow and learn, but their developmental and learning processes should be crafted by the most powerful designer of adaptive complexity known to science: natural selection.

In some senses AI has been moving in this direction since 1956. The field's historical trajectory from a focus on isolated and sophisticated mental faculties to a focus on the commonsense knowledge needed for everyday tasks, and more recently to a focus on the construction of complete, situated, and embodied agents, is a trajectory from *a priori* conceptions of intelligence toward theories derived from the natural forms of intelligence that we observe around us. The logical extension of this trend is to model not only the products of natural evolution but also its processes.

Acknowledgements

This material is based upon work supported by the United States National Science Foundation under Grant No. 0308540. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

[1] E. Charniak, D. McDermott, Introduction to Artificial Intelligence, Addison-Wesley, Reading, MA, 1985.
[2] T. Dean, J. Allen, Y. Aloimonos, Artificial Intelligence: Theory and Practice, Benjamin/Cummings, New York, 1995.
[3] A. Fukunaga, Automated discovery of composite SAT variable-selection heuristics, in: Proceedings of the Eighteenth National Conference on Artificial Intelligence, AAAI Press, Menlo Park, CA, 2002, pp. 641–648.
[4] J. Lohn, G.S. Hornby, D. Linden, An evolved antenna for deployment on NASA's space technology 5 mission, in: U.-M. O'Reilly, R.L. Riolo, T. Yu, B. Worzel (Eds.), Genetic Programming Theory and Practice II, Springer, New York, 2004.
[5] S. Luke, Genetic programming produced competitive soccer softbot teams for RoboCup97, in: Genetic Programming 1998: Proceedings of the Third Annual Conference, Morgan Kaufmann, San Francisco, CA, 1998, pp. 214–222.
[6] N.J. Nilsson, Artificial Intelligence: A New Synthesis, Morgan Kaufmann, San Francisco, CA, 1998.
[7] The Pew Research Center for the People & the Press, The Pew Forum on Religion & Public Life, Public Divided on Origins of Life, 2005.
[8] R. Poli, W.B. Langdon, Backward-chaining evolutionary algorithms, Artificial Intelligence 170 (11) (2006) 953–982.
[9] S. Preble, M. Lipson, H. Lipson, Two-dimensional photonic crystals designed by evolutionary algorithms, Applied Physics Letters 86 (2005).
[10] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, second ed., Prentice Hall, Upper Saddle River, NJ, 2003.
[11] L. Spector, Automatic Quantum Computer Programming: A Genetic Programming Approach, Springer, New York, 2004.
[12] A.M. Turing, Intelligent machinery, in: D.C. Ince (Ed.), Collected Works of A.M. Turing: Mechanical Intelligence, Elsevier Science Publishers, Amsterdam, 1992.
[13] P.H. Winston, Artificial Intelligence, second ed., Addison-Wesley, Reading, MA, 1984.