Robotica (1999) volume 17, pp. 229–235. Printed in the United Kingdom

Book Reviews

REINFORCEMENT LEARNING: AN INTRODUCTION by Richard S. Sutton and Andrew G. Barto, Adaptive Computation and Machine Learning series, MIT Press (Bradford Book), Cambridge, Mass., 1998, xviii + 322 pp, ISBN 0-262-19398-1, (hardback, £31.95).

This book introduces a new approach to the study of systems, living or artificial, that can learn from experience, and in so doing it indicates a new focus of activity within Artificial Intelligence and Robotics. The approach is new insofar as this is the first comprehensive account to appear, but references are given to earlier papers in which the ideas were developed, back to at least 1982 for the particular approach now given, and very much earlier for the origins of specific aspects.

The idea of machines that would learn from experience is, of course, quite old and there have been speculations by Shannon, Turing and Ashby, among others. Principles that underlie the present approach are used in the famous checker-playing program of Samuel, and there is reference to early discussions by Minsky, as well as to the pioneering work of a number of Russian workers including Tsetlin and Feldbaum. (Some ideas were given an airing at the 1960 IFAC Congress in Moscow.)

A capacity for learning is a feature of most schemes for artificial neural nets. Neural nets receive attention in the present book, but the main concern is overall principles of operation, with the comment that neural nets may or may not prove to be the most appropriate implementation. (The earlier papers by the authors are somewhat more strongly orientated towards neural-net implementations.)

It is refreshing to see that the general theory of learning systems is progressing, at a fundamental level. It can be seen as a refutation of the claim of some workers in *AI*, notably Schank (1985) that learning is only usefully studied within an expert-system, or knowledge-based framework. The aspects studied within this framework are also valid, and grist to the mill, but certainly not exclusively.

Previous studies of machine learning have tended to look to pattern classification as a source of tasks of a suitable level of difficulty, and the majority have simplified matters still further by assuming that learning would be 'with a teacher' rather than by interaction with a raw environment. In particular, most work on neural nets is directed at pattern classification in the 'with a teacher' situation. The limitations of this approach are discussed in the book, where instead learning is viewed as interaction between the learning agent and an incompletely-known environment, in such a way as to maximise, in the long term, a reward obtainable from the environment.

This view of learning has been adopted by many previous workers, and Selfridge coined the attractive term 'hedony' to denote the reward signal. It has been usual to see the reward as a simple quantitative feedback, even though it was sometimes acknowledged that it could be subject to unknown delay.

The term 'reinforcement learning' is now used to denote a more general approach, with the environment represented in finite-automation terms, so that an action applied to the environment in a particular state produces a (perhaps stochastic) immediate reward and a (perhaps stochastic) change of state. The long-term return from the action is the sum of the immediate reward and the return achievable from the new state. The return obtainable from each state, or 'value' of the state, is therefore evaluated by a form of bootstrapping, since the value of any one state depends on the values of others that can be reached from it.

The return achievable from a state, or 'value' of the state, depends on the policy according to which actions are selected as a function of state. The general idea is the basis of Dynamic Programming, but unlike it the learning procedure operates on an unknown environment and works to determine state values (or in some versions values associated with state-action pairs) and from these values to derive favourable policies. Since the state values are functions of policy, and policy improvement depends in turn on the values, policies and valuations have to evolve jointly. Algorithms are described which allow such joint improvement and for most of them the reader is assured of the existence of a convergence proof.

This approach, corresponding closely to Dynamic Programming, is contrasted with the alternative referred to as Monte-Carlo Method. This latter is only applicable to the learning of tasks that are episodic, in that they come to a natural end as does a play of a board game. The term is used in a rather unusual sense, as it is applied to observations on real-world events as well as to simulations depending on random or pseudo-random number generation. It essentially denotes derivation of state-values without bootstrapping.

A third form of operation is then introduced, having features of both the Dynamic Programming and Monte-Carlo versions. It is termed Temporal-Difference Learning and it is introduced by observing that data may become available during interaction with the environment that warrant revision of an earlier estimate of some quantity. The introductory example is the estimate of tbe time taken to travel home from work, which will be revised during the journey if one particular leg of the route takes longer (or shorter) than expected.

These considerations provide the basis for powerful methods of reinforcement learning. It is acknowledged that not all environments fit comfortably into the essentially finite-automation representation adopted. Often the number of states will be unmanageably large. In these cases some mathematical function can be used as an approximation, and the reinforcement-learning principles can be adapted. All of this receives attention in the book, as does means of modifying the model as experience is gained, distinct from operation to improve the fit of a given model.

The power of the method is illustrated by description of a number of case studies, in which Samuel's early work on checker (or draughts) playing is cited as exemplifying the general principles, specially of bootstrapping of state-values. Much more recently the methods have been used in constructing a powerful program to play backgammon. Another example described is their application to design of a dispatching system for lifts in a multistory building, and others refer to dynamic channel allocation in a communication network, and to job-shop scheduling. In all of these applications, impressive new results have been achieved, sometimes by taking substantial liberties with the basic methods.

A particularly interesting area of application is to robotics, and one case study is in this area. A reference is given (Connell and Mahadevan, 1993) to a publication in which robotics applications of reinforcement learning are treated more fully.

The treatment is accessible without great demands on the reader's knowledge of mathematics. At many points, more difficult

https://doi.org/10.1017/S026357479925117X Published online by Cambridge University Press

mathematical treatments are referred to, for example the convergence proofs mentioned earlier. The whole book is, in fact, admirably presented and is a magnificent piece of work. Attention is given to the historical development of the topic, and there is a large and valuable, though inevitably far from exhaustive, collection of references.

One minor criticism, with regard to the mathematics, is that although a useful summary of notation has been provided, it is tucked away between the references and the subject index and is easily overlooked. The symbols are, of course, defined when first used in the text and are mostly fairly obvious, but my initial perusal, when I had not noticed the summary of notation, was hampered by confusion between the curly-R representing the set of real numbers, and a rather similar-looking one (though with a subscript and superscript) indicating an expected immediate reward. Reference to the summary of notation would have clarified matters immediately.

This, of course, is a small point and this is definitely an extremely important and well-prepared work.

Alex M. Andrew,
95 Finch Road,
Earley,
Reading, RG6 7JX
(UK)

References

J. Connell and S. Mahadevan, *Robot Learning* (Kluwer, Boston, 1993).
R.C. Schank, "Looking at Learning" **In**: (L. Steels and J.A. Campbell, ed.) *Progress in Artificial Intelligence* (Ellis Horwood, Chichester, 1985) pp. 17–29.

IN THE MIND OF THE MACHINE: THE BREAKTHROUGH IN ARTIFICIAL INTELLIGENCE by Kevin Warwick, Arrow Books (Random House), London, 1998, x + 307 pp, ISBN 0-09-970301-7 (£7.99; Pbk).

This is a slightly revised version of the author's earlier *March of the Machines*, published by Century Books (also Random House). Kevin Warwick is Professor of Cybernetics at the University of Reading (UK), and his views, as well as the practical demonstrations of robots developed under his guidance, have received much publicity in the press and on BBC television and radio. As related in the book, the robot demonstrations have aroused wide international interest.

The central message is that the development of complex information processing machines, having a claim to be termed "intelligent", is at a stage where they may very soon surpass human intelligence, and since this new form of intelligence is unlikely to remain uniformly benign towards humans, this could spell the end of humans as the dominant species. In his second chapter, the author paints a fearsome "worst-case" scenario, according to which machines come to exploit humans essentially as humans have exploited farm and draught animals. In a later chapter he drives home his point even more strongly by suggesting that machines could come to hunt humans for sport and might stage gladiatorial combats.

The presentation is easily readable, in a chatty and entertaining style, free of mathematics and jargon. The topic invokes long-debated philisophical issues, and there are quotations from various classical writers, but on the whole the treatment is informal with a common-sense "homespun philosophy" character. I found myself generally sympathetic to the approach, and probably so will most people in the *AI* and computing field, though in a number of places the author seems to promise some profound revelation and then lets it fall rather flat. For example, he promises to deal with the thorny matter of consciousness in Chapter six and then has rather little to say about it.

One of the long-debated philosophical issues is the relationship of machine intelligence to the natural variety, and the extent to which the former can simulate the latter. Kevin Warwick ducks that one nicely by accepting that the two kinds of intelligence may continue to differ, but he insists that this gives no grounds for refusing to see the machine variety as powerful and threatening. He also claims that machines might have forms of consciousness and emotions that need not correspond to the human varieties. The threat to humans will come from machines that are given, or evolve, a goal of self-preservation, and it is easy to imagine such a system displaying, in a threatening situation, a flurry of activity that could be seen as a manifestation of the emotion of fear.

The terms "machine" and "robot" are used interchangeably, but it is made clear that neither is meant to imply restriction of attention to autonomous self-contained, preferably humanoid, agents. They are meant to include, for example, computer networks dealing directly on the stock market, and it is from systems of this sort that the greatest threat may come since their operation is goal-directed, and for the sake of speed is not subject to human checks, and the goals could easily come to include that of self-preservation. Other machines that are particularly likely to have a built-in regard for self-preservation are the various sophisticated autonomous weapon systems of which cruise missiles are the best-known example.

Where "intelligence" is viewed as broadly as it is here, it is difficult to find a basis for comparison. In most of the treatment, Kevin Warwick uses a criterion that has a flavour of tautology in the context, since he equates intelligence with the power to dominate. (In an analogous way, much of educational psychology makes sense if intelligence is defined as whatever it is that intelligence tests measure.) He is not totally consistent in his view of intelligence, and clearly assumes something rather different when he refers to "intelligent" interactions between humans and machines, and when he draws encouragement from the fact that the demonstration robots built in his department in Reading appeal to human viewers as behaving in a life-like way.

A rather curious inconsistency is that, although he is prepared to accept that the consciousness and emotions of machines are likely to be different from those of humans, he nevertheless lists musical composition and performance as areas in which machines may show creativity. If it is assumed that the music, or other aesthetic creation, will be judged by humans, this seems to imply that machines and humans have similar aesthetic preferences, a conclusion he would presumably not wish to defend. So long as aesthetic value is judged by human response, aesthetic creativity must always be one area where humans have an advantage over machines. There is of course no reason why machines should not start to produce alternative versions of music and other art forms that are pleasing to them.

Despite these criticisms of details, there is a message here that merits careful attention. This is far from being the first time the human race has been threatened by its own creations—to take a simple example, a variety of adverse influences on the environment have been attributed to the "tyranny of the motor car". The new threat is significantly different in character, in that it comes from machines that can be visualised as planning their campaign, and as continuing to operate independently of their human creators. Cars, on the other hand, although they frequently eliminate people, do not apparently premeditate such acts, and their own ultimate fate remains linked to ours.

A good deal of the book is devoted to an account of robot research in Reading. Some very good and ingenious work to aid handicapped subjects is briefly described, but the main focus is on groups of mobile robots that are referred to as the Seven Dwarfs. Several generations of these have been constructed and they show, in elementary ways, interaction with the environment so as to avoid obstacles and to look for a recharging station when necessary, and to improve their operation by learning. Later versions allow simple communication between robots and hence interesting group behaviour, as well as automatic improvement over a succession of generations as an analogue of natural evolution.

It is of course acknowledged that there is an enormous gulf between the capabilities of these robots and the complex machines that may challenge humans. At the same time there is a good case

https://doi.org/10.1017/S026357479925117X Published online by Cambridge University Press

for believing that certain insights are more likely to come from the building of robots in computer programs. It is all too easy to make unrealistic assumptions in constructing simulations, for example about the uniformity of objects and conditions encountered. Primitive though they are, the Seven Dwarfs give some feeling for the evolutionary processes that could possibly allow machines to pose a threat.

Although the warning in the book has to be taken seriously, it is impossible not to feel that the author underestimates the disparity between the flexibility and power of biological adaptation, and anything comparable yet shown by machines. Certainly, machines can show non-trivial learning behaviour, and impressive claims have been made for results of simulated evolution as a facet of "Artificial Life", but the relevance of these studies to real-world evolution is questionable. For the one thing, the assumed environments for simulated evolution must be even more tenuously related to real-world conditions than are those of the robot simulations suggested as alternatives to the Seven Dwarfs.

The matter of comparing machine and human intelligence is enormously complex and is certainly not settled by the simple observation that the brain has little opportunity for expansion, whereas machines, and machine networks, can expand indefinitely. The relevant comparison, of course, is not between a machine and an isolated brain but between a breakaway machine faction on the one hand, and human brains in collaboration with loyal or enslaved machines on the other. Using the comparison, machine takeover appears rather less likely.

A worrying further reflection, though, is that humans are well able to persist in a disastrous course even when nominally in control. The availability of intelligent machines can accelerate our lemming-like behaviour, even though if applied differently they could provide some partial solutions. It is easy to visualise several kinds of situation, short of a takeover, in which the consequences of involvement with machines could be serious.

One type of situation is, of course, that in which machines are enlisted in human conflicts (in which case, Asimov's proposed laws of robot behaviour are abrogated from the start). Another is that in which the activity of machines has undesirable side-effects, but the short-term benefits give an incentive to play these down. The alleged "tyranny of the motor car" is an obvious example and there are many others, some associated with industry and others with agriculture. Another danger is that we may allow ourselves to become over-dependent on apparently reliable machines, and then encounter an unforeseen "bug". The predicted breakdown of financial and other systems with the "millennium bug" is a prime example.

My own feeling is that Kevin Warwick is jumping the gun in visualising an imminent machine takeover as a deliberate orchestrated event. On the other hand, there are ways that the participation of machine intelligence could speed us towards any of the variety of disaster situations, and his book is thought-provoking and informative as well as making good reading.

Alex M. Andrew
95 Finch Road,
Earley,
Reading, RG6 7JX
(UK)

ARTIFICIAL INTELLIGENCE AND MOBILE ROBOTS: CASE STUDIES OF SUCCESSFUL ROBOT SYSTEMS

edited by David Kortenkamp, R. Peter Bonasso and Robin Murphy, AAAI Press (American Association for Artificial Intelligence), Menlo Park, CA, and MIT Press, 1998, x + 390 pp, ISBN 0-262-61137-6 (£29.95; Pbk).

Mobile robots are of particular interest from a number of points of view. They have a special appeal as the traditional popular image of a robot, and there are many purposes that can be served only by a mobile device. They provide a rigorous "proof of the pudding" testing-ground for the relevant *AI* techniques. It is also an area where useful comparisons can be made between animal and machine performance.

The thirteen chapters of this book, following a Preface and a useful Introduction, give descriptions of the same number of successful mobile robot systems, chosen to represent the best available from leading universities and research laboratories. Many of the robots have distinguished themselves, usually with first or second place finishes, at indoor and outdoor mobile robot competitions. As the Editors point out in the Introduction: "These are not robots that simply work in the laboratory under constrained conditions. They have left the lab and been tested in natural and unknown environments. They have all demonstrated their robustness many times over."

In the Introduction the state of the art is reviewed in more detail. The performance of mobile robots has been greatly enhanced in recent years by a paradigm shift prompted mainly by a paper in 1986 which introduced 'subsumption architecture'. Prior to the shift, mobile robots generally reviewed their sensory data and formed a plan of action. The perception of a previously-unnoticed feature, such as a hole in the floor in the robot's path, could only be taken into account by tediously repeating the whole process of analysis and planning.

The newer paradigm follows the example of biological systems, particularly vertebrate motor systems relying on spinal reflex arcs. Means of obstacle avoidance are embodied, and can produce a deviation in the robot's path or actions when necessary, but without abandoning the overall plan. The higher levels subsume these lower ('spinal') levels, hence the name of the architecture.

In each of the thirteen chapters a separate complete system is described, so readers have the chance to compare whatever feature of the systems they choose. The principles of operation are described in a fair amount of detail, including sections of pseudo-code. Nevertheless the papers are grouped under three headings, according to the aspect to which they contribute most strongly. The three parts of the book are concerned with Mapping and Navigation (4 papers), Vision for Mobile Robots (3 papers), and Mobile Robot Architectures (6 papers). Each part has an introductory section by an Editor or Editors.

Of the four papers in Part one, three are explicitly aimed at robots to operate in an office environment, and the other describes a similar system, but without explicit reference to an office application. Such robots can act as messengers, or collect rubbish, or can be tour guides. Navigation can depend on maps of different kinds, one being a metric map, usually stored as a grid, and another being a topological map in the form of a graph in which nodes correspond to objects (obstacles or waymarkers) linked by arcs denoting relationships. The two forms of mapping may be combined and one of the schemes incorporates a means of deriving topological maps from the grid kind.

The robot described in the fourth chapter is said to roam the corridors of the Carnegie-Mellon University, and uses navigation architecture based on "partially observable Markov decision process models", or POMDPs. These have found applications in other areas in which decisions must be made on uncertain data, and standard algorithms are available.

In the introduction to Part two, on Vision for Mobile Robots, it is mentioned that for a long time designers of mobile robots were reluctant to include vision, with its large overhead of processing. Where the main requirement is the detection of obstacles to allow navigation among them, simple ultrasonic or infrared rangefinding devices are more convenient. Visual input is needed where scenes and objects have to be examined in detail, and is increasingly coming into use.

Of the three papers in Part two, one describes an early visually-guided robot, unfortunately no longer operating, which offered to guide visitors around a floor of the AI Laboratory of MIT. The second is an exposition of a principle of "action-oriented perception", according to which any perceptual process should be planned with regard to a perceptual context, which includes analysis of the information requirements of the overall task. The principle is illustrated by its application to an unmanned land vehicle. The remaining paper in this part is also concerned with automatic vehicle guidance, with an immediate practical application, namely the safe steering of a truck onto the hard shoulder of a motorway should the driver fall asleep.

In the papers in Part three the focus is on architecture. In the first two, the treament is in the context of highly versatile mobile robots including, in the second paper, the line of development that includes the well-known names of 'Shakey' and 'Flakey', both products of the Stanford Research Institute.

The applications considered in this part range widely. There is another paper on vehicle guidance, one on robots able to collaborate as a team, and another on robots to operate alongside humans. Another describes an autonomous inderwater vehicle, and 'virtual reality' simulators to allow it to be tested in the laboratory under conditions corresponding, as far as sensory input is concerned, to a great depth in the ocean.

A paper from the group in Oxford University describes advanced methods for the guidance of mobile robots used to convey goods in manufacturing industry. Robot trucks for this purpose are widely used, guided by cables buried in the floor, or by painted lines. The use of rigidly defined routes has disadvantages, since as with public roads there will certainly be occasions on which some activity requires blocking of the route. In the paper an alternative scheme for guiding the robot trucks is described in considerable detail. The autonomous vehicles scan with laser beams to locate reflecting beacons, and can thereby navigate to their destination, finding an alternative route when the previously-known one is found to be blocked, and remembering the new route for future trips.

There is a wealth of useful material in this collection, with all of it, as emphasised in the Preface, referring to systems that have been well tested outside the laboratory and shown to be robust in operation.

Alex M. Andrew
95 Finch Road,
Earley,
Reading, RG6 7JX
UK

BEHAVIOR-BASED ROBOTICS by Ronald C. Arkin, with a foreword by Michael Arbib, Intelligent Robots and Autonomous Agents series, MIT Press, Cambridge, Mass., 1998, xiv+491 pp, ISBN 0-262-01165-4 (£39.95; Hbk).

As the notes on the dustcover say, this is the first true survey of this robotics field. It covers a very wide range of topics, all treated very clearly and outlining and contrasting diverse approaches and viewpoints. The reference to behaviour (with or withoput the letter 'u') carries two implications—first that the concern is primarily with the 'brains' of robots rather than with their physical features, and secondly that there is frequent reference to relevant theories and models from biology and psychology. Of course, robotics is by its very nature a topic where the 'brains' cannot be considered in isolation from the physical interactions, so the primary concern with perceptual and control processes must be mainly a matter of emphasis.

The foreword by Michael Arbib is enthusiastic, not least about the way the presentation is embellished with apposite quotations from a wide variety of sources, and about the copious provision of illustrations. Arbib makes the slightly laboured but nonetheless apposite comment that although the primary concern is with robot brains, the book gives many opportunities to admire human wit and robot bodies.

The author explains that he realised the need for the book, and decided to write it, then he was unable to find a suitable text for a course he had taught for ten years. He had been forced to recommend to the students rather indigestible technical material and collections of original papers. The book's intended audience includes upper-level graduate students and graduate students studying *AI* and robotics, as well as those interested in learning more about robotics in general.

Developments in robotics are covered over a long time span. Attention is restricted to the electronics era, so there is no review of the various golems and such things as Vaucanson's rather offputting duck. There is a full description of Grey Walter's tortoise, and the story is continued right to the latest developments. Michael Arbib comments favourably on the inclusion of motor schema theories, and there are also references to subsumption architecture, and to reinforcement learning in its modern robotics context, among much else.

Although the book has been prepared primarily as teaching text, and will serve that purpose admirably, it is written with a good deal of flair and imagination. This is perhaps best illustrated by quoting the last three paragraphs of Arbib's Introduction:

"The book's chapter on social behavior presents what is itself a relatively recent chapter of robotics—sociorobotics. While in its infancy, we can see in the studies of robot teams, inter-robot communication, and social learning the beginnings not only of a powerful new technology, but also of a new science of experimental psychology.

"Finally, we are taken to that meeting place between science fiction, philosophy, and technology that attracted so many of us to wonder about robots in the first place. The final chapter, 'Fringe Robotics: Beyond Behavior' (a nod to the 1960s British review 'Beyond the Fringe'?), debates the issues of robot thought, consciousness, emotion and imagination, returns to Arkin's long-standing concern with the possible utility to robots of analogs of hormones and homeostasis, and closes with an all too brief glimpse of nanotechnology.

"In this way we are given a tour that impresses with the depth of its analysis of the schemas underlying robot behaviour, while continually illustrating the deep reciprocity between robotics and biology, psychology, sociology, and philosophy, and the important connection between robotics and many other areas of computer science. This is a subject whose fascination can only increase in the decades ahead as many researchers build on the framework so ably presented here."

The glowing commendation is fully deserved. There is a comprehensive subject index and name index as well as a valuable bibliography of no less than 32 pages. This is a valuable reference work as well as a thought-provoking and imaginative review, providing much more than would normally be expected of a teaching text.

Alex M. Andrew
95 Finch Road,
Earley,
Reading, RG6 7JX
(UK)

GEOMETRY AND INTERPOLATION OF CURVES AND SURFACES by Robin J. Y. Mcleod and M. Louisa Baart, Cambridge University Press, Cambridge, 1998, xiv + 414pp., index (£50; Hbk).

This book book is the outcome of many years of research into a large number of practical problems, including some that had baffled mathematicians for a long time.

Seven chapters are devoted to the topics, as follows: Interpolation, Conic Sections, Synthetic Geometry, Algebraic Projection Geometry, Algebraic Curves, Examples and Applications, and a last chapter containing bibliographic notes and exercises. The authors have displayed considerable erudition and ability in dealing with a difficult topic. Their practical, step-by-step approach to algebraic curves and surface interpolation is very commendable in view of the complexity of the subject. Insight is thus gained into practical applications in engineering analysis, approximation and curve-plotting problems. Many concrete examples are presented, including the numerical solution of trajectory problems, parabolic interpolation, geometric approximation, and an analysis of rational and polynomial cubics. While the algebraic approach is the main theme of the book, there are also details of a synthetic approach.

This compilation is a sound and well presented work of interest and use to academics and professionals in engineering design and analysis, applied mathematics, graphics, solid modelling and robotics.

J.A. Prescott
(c/o editor)

INTRODUCTION TO ARTIFICIAL LIFE by Christoph Adami, TELOS, The Electronic Library of Science, Santa Clara, an imprint of Springer-Verlag, New York, 1998, xviii + 374 pp, ISBN 0-387-94646-2 (Hardback, with CD ROM, DM 118).

The Preface begins with the question: "What makes living systems alive?" and goes on to point out that our level of understanding of biological systems is such that it is difficult to answer even this fundamental question. In contrast, our understanding of non-living systems has advanced dramatically. The 'Artificial Life' approach is meant to clarify the basic principles of the biological sciences and to bring them more nearly into line with the physical sciences.

To clarify the essential characteristics of life, its origins are considered, probably in an RNA-world existing as a primeval soup before the emergence of cells. It has been shown that the spontaneous formation of amino acids is reasonably likely under conditions that can be assumed to have existed. The significant step is the appearance of molecules capable of replication. Research is mentioned that may reconstruct a molecule, termed RNA replicase, from which all life is descended, a molecule that has been extinct for three billion years.

Although the RNA replicase molecule is extinct in its original form, an important result of replication is the preservation of molecular structure, and its evolutionary enhancement from a survival point of view, over periods enormously greater than the time a particular molecule would survive. The latter may be a matter of hours or days, but the main features of the RNA replicase molecule have survived over these three billion years (and we trust they are not about to expire now). In a (slightly tentative) definition on page six of the book, this feature of persistence of information in the face of disruptive forces is postulated as the essential characteristic of life.

There is a review of computer programs aimed at simulating biological evolution, including one that produces designs of physical 'bodies' capable of swimming, etc. The possibility of replication was studied by von Neumann in the context of cellular automata, and Conway's game of 'Life' is an interesting example allowing a form of replication.

The main focus of the book, however, is on populations of computer programs capable of replication, and subject to some kind of mutation as well as to overall culling by a 'reaper' so that there is effective competition. Computer viruses are an example of self-replicating programs, which have evolved to more powerful forms during their period of existence. Their evolutionary development, however, is not a real model of biological evolution. It is not attributable to random mutation and selection but rather to the purposeful intervention of highly ingenious malicious human programmers.

Another example of a life-like situation in a machine is the game of 'Core Wars' in which the players write programs that compete for space in computer memory. To run on an existing computer, the programs are written in a specially-devised language and are restricted to a specified section of storage.

A system that convincingly simulates aspects of biological evolution is given the name *tiera*. In this, as in Core Wars, programs compete for space, but they are now subject to mutation. The instruction set differs from that used in Core Wars and is designed to be less 'brittle'. A 'brittle' program is one whose behaviour is likely to be drastically affected by a mutation, as compared to one which moves to a new mode that is different but not drastically so. A 'reaper' operates to limit the overall population.

Given the simple instruction set allowed in *tiera*, it was possible to find a minimum length of program allowing replication. It was found in operation, however, that shorter programs emerged, apparently with the capability of replication. Closer examination showed that these were effectively parasites whose replication depended on plugging-in to instruction sequences in other programs. Short programs, and hence the parasites, have a reproductive advantage, but only so long as there is also a population of the longer 'host' programs. At a later stage in the simulation the numbers of these parasite programs declined, and it was found that the longer programs had evolved a means of preventing parasitic use of their replication facility. The correspondence to biological interaction is compelling.

The book introduces a comprehensive facility for this sort of modelling, with the name of *avida*. The necessary software is provided on the CD ROM included. It differs from *tiera* in a number of respects, one of them being that a larger instruction set is provided. The set provided in *tiera* is insufficient to achieve universality of computational function, and it is felt that a set allowing this is preferable. The user of the *avida* software can specify of the size of the instruction set used in any experimental run, but the minimum set allowing universality is already considerably larger than that of *tiera*.

The exciting results obtained from *tiera* were facilitated by its small instruction set, which meant that the number of possible sequences of instructions, of sufficient length to be interesting as possible self-replicators, is large but not astronomical. For *avida* the space of possibilities to be sampled is much larger and experimental runs definitely have to be seeded with one or more 'mother' replicators. The instruction set can be referred to as the chemistry of the system, analogous to the amino acids to be assembled into gene sequences.

The *avida* system differs from *tiera* in its general arrangement, in that all of the putative programs in *tiera* share one linear address space, but in *avida* they are like distinct organisms, and can have a spatial distribution. This allows spatial phenomena in *avida*, such as diffusion and waves of change. The *avida* software allows the user to set up experiments with control of the instruction set, or 'chemistry', and of such things as the 'mother' replicators installed, the mutation rate and the type of mutations allowed. The viability of a program type is determined partly by its rate of reproduction, but is also influenced by a user-specified 'fitness landscape'. Needless to say, the package includes comprehensive facilities for collecting and processing data from the runs.

Besides introducing the *avida* system, and providing an Appendix that is a users' manual, the book deals with theoretical approaches that allow analysis of the results and comparison with those on true biological systems. One of these is Shannon's Information Theory, where it is emphasised that the optimal communication between programs is not perfect uncorrupted transmission but transmission introducing the right amount of entropy.

Other connections with thermodynamics are developed in another chapter, and yet another discusses the spatial concept of percolation through a regular array of elements of which a subset is tagged in some way. The problem is to assess the probability of finding a path through adjacent tagged elements that links opposite faces of the array. In one dimension the problem is trivial as the path is only formed if all elements are tagged, but there is a challenging problem when two or more dimensions are considered. The theory is shown to be relevant to analysis of evolutionary processes.

A type of behaviour that can be observed in evolving systems, living or artificial, is described as 'self-organisation to criticality', in which a succession of epochs of self-organisation reach a kind of stagnation and crumble away. In a later chapter it is argued that particularly interesting adaptation occurs when the population is close to an error catastrophe, which happens when a code string is likely to be mutated before it generates a copy of itself. These phenomena are related to phase transitions in physics and to analysis associated with the 'hypercycle' theories of Eigen and Schuster.

The software on the CD ROM allows *avida* to be installed under Windows 95 or NT, or on a UNIX system. The software can be reproduced and shared freely, provided it is not for profit, and it can also be obtained (even without buying the book) by downloading from a web page whose address is given on page 322. The version available from the web page will be updated.

It is easy to feel swamped by the deluge of persuasive theorising, but the book gives a good and readable introduction to this important topic, with the invitation to run the software and so to get hands-on experience and even contribute fresh findings.

Alex M. Andrew
95 Finch Road,
Earley,
Reading, RG6 7JX
(UK)

ROBOT LEARNING, edited by Jonathan H. Connell and Sridhar Mahadevan, Kluwer, Boston, 1993/1997, xii + 240 pp., ISBN 0-7923-9365-1 (Hardback, 218.00 Guilders, $120.00, £89.95).

This is a collection of eight papers, with a brief Preface. In the Preface it is pointed out that the topic of robot learning has recently experienced a resurgence of interest, with sessions devoted to it at conferences on *AI* and on Neural Information Processing. The present volume is not attributed to a specific meeting, and appears to have been assembled by invitation of the editors.

The first paper is by the two editors and gives a useful introduction to the topic and a brief survey of the other contributions, and of some background. Three main motivations for studying robot learning are listed. One is that the environment may be too complex for the necessary information to be hard-wired or programmed into the robot initially. Another is that the information may be unknown in advance, as when the aim is to explore a new planet, and another, related to this, is where the environment is changeable. Robot learning has characteristics that distinguish it from other computer learning, since it must operate with noisy sensory input, in a stochastic environment, and must react to inputs without undue delay. Learning has to be incremental, in that it enables the robot to improve its performance while actually performing, and in general training time is limited.

The feedback governing learning may be a simple scalar 'success' indication, or may be a control feedback as in the 'learning with a teacher' situation, or may be more complex still. The case of scalar feedback immediately suggests recent developments in Reinforcement Learning, and several of the papers in the volume make reference to this.

The second paper, however, is innovative in other directions. It is about autonomous vehicle guidance and in particular the ALVINN project at Carnegie Mellon University, where the acronym stands for Autonomous Land Vehicle In a Neural Network. The chapter can in fact be regarded as a synopsis of the book by Pomerleau,[1] which was presumably published later than the present volume since no reference is made to it. The project involves the use of neural nets in this practical context with various important new features prompted by its requirements.

The other papers are of more general applicability. The third is on reinforcement learning with multiple goals. In principle, reinforcement learning methods can accommodate multiple goals by forming a coposite reward criterion, but the learning process is then slow and cumbersome. A method is described in which modules are assigned to the separate goals, though with suitable coordination by an "arbiter". Reference is made to the multiple needs of an animal which must satisfy the goals of getting food, getting water, procreating, caring for young, and so on.

The fourth paper begins with the observation that the new incremental methods of reinforcement learning have fast real-time performance, whereas classical methods are slow but more accurate because they make full use of the observations. A technique called 'prioritised sweeping' is introduced with the aim of getting the best of both worlds. This concentrates computational effort on the most 'interesting' parts of the system. Results from simulations are given.

The remaining four papers all emphasise special features of the robot learning situation, as compared to machine learning in general. The fifth paper is another by the two editors and discusses a number of ways in which learning can be speeded up. The sixth treats the semantic hierarchy in robot learning, in which an abstract 'cognitive map' is at the highest level and this has to be related by stages to physical robot control. An important point receiving attention here is the transition from low-speed, friction-dominated control to higher-speed but subsequent operation is likely to be faster, and so to make different demands on sensory information and prediction.

The seventh paper is on the learning and updating of map information using uncertain data. The map need not be a topological representation as ordinarily understood but can be a graph of states and transitions in any environment. A conventional map used for spatial navigation is a convenient specific example. A total of twelve theorems is given, relating to convergence rates under each of the four permutations of deterministic/stochastic transitions with deterministic/stochastic observations.

The final paper is a perceptive review again emphasising the differences between the physical robot situation and the relatively tidy environment of theory and computer simulations. The conclusion is reached that: "The weaknesses of the existing learning techniques, and the versatility of the knowledge necessary to make a robot perform efficiently in the real world suggest that many concurrent, complementary, and redundant learning methods may be necessary". The comment brings to mind the remark quoted by Fischler and Firschein.[2] "I suspect that the representational system with which we think, if that's the right way to describe it, is so rich that if you think up any form of symbolism at all, it probably plays some role in thinking".

All the papers presented confirm the complexity of the problem and make significant contributions towards its solution. This is a useful and important book.

References

1. Dean A. Pomerleau, *Neural Network Perception for Mobile Robot Guidance* (Kluwer, Boston, 1993) (reviewed in *Robotica*, **vol. 14**, part 5, pp. 586–587, 1996)
2. M.A. Fischler and O. Firschein, *Intelligence: The Eye, the Brain and the Computer* (Addison-Wesley, Reading, Mass., 1987) p. 308.

REINFORCEMENT LEARNING, edited by Richard S. Sutton, Kluwer, Boston, 1992/1996, 172 pp., ISBN 0-7923-9234-5 (Hardback, 245 Guilders, $135.00, £101.25), reprinted from special issue of *Machine Learning*, **vol. 8**, nos. 3–4 (1992).

This is an exact reprint of the special issue, with the pages carrying double numbering – starting from one in the lower outer corners as appropriate for the book, and from 225 in the upper corners to allow reference to the journal. There is a short Introduction by the editor, followed by seven papers, all dealing with important aspects of the current "hot topic" of reinforcement learning.

In the Introduction it is stated that there is as yet no good tutorial textbook on reinforcement learning. This, of course, is written before the appearance of the very fine book by Sutton and Barto[1] recently reviewed in these pages. Even so, the present book is well worth perusing as it treats a number of important topics that are either omitted or not fully developed in the later text. Its editor has seen it filling the role of an introductory text, and in his Introduction he quotes a number of other papers, several of them having Andrew Barto as one of the authors, that should be read to supplement the present volume to get a comprehensive overview.

Rather surprisingly, the first of the seven papers treats reinforcement learning in the case where only an immediate return need be considered. It is on "Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning". The algorithms derived thus serve the same purpose as is served by the widely-used backpropagation method, and the author acknowledges compatibility between the two. However, the

https://doi.org/10.1017/S026357479925117X Published online by Cambridge University Press

new algorithms operate without the knowledge of network connectivity and of feed-forward transformations that is assumed in backpropagation. The network elements are seen as totally stochastic, with only parameters of the output distribution of each element controlled as a function of its inputs. A comprehensive mathematical analysis is given, though without the convergence proof that the author would clearly have liked to provide. The type of algorithm is indicated by the acronym REINFORCE, formed by taking letters from the equation: "REward Increment = Nonnegative Factor × Offset Reinforcement × Characteristic Eligibility".

The second paper focusses on delayed reward and gives a very thorough discussion of the difficulties that may be encountered in applying reinforcement learning in this case. The application of the Temporal Difference method in a program to learn to play backgammon is described in detail. The program has performed extremely well, and after developing its playing strategy by self-play, starting with no information except the rules of the game, it performed much better than existing commercial programs including one based on an elaborate analysis of human play.

Two other papers contribute to the basic theory by offering, in each case for the first time, convergence proofs for reinforcement learning methods applicable with delayed reward. One confers this accolade on the variant termed Q-learning. The other confers it on the Temporal Difference Method, for which a proof had been given earlier by Sutton in the TD(0) case, i.e. where the time horizon is just one step. A paper in this volume extends the result to the general TD(λ) case.

Two other papers acknowledge that the established methods of reinforcement learning with delayed reward, while effective, can be slow to converge in complicated tasks. In a paper by Lin a number of composite techniques are discussed, including the combination of the methods with teaching-by-example and re-using past experience.

The other paper, by Singh, gives another approach based on the decomposition of long sequential tasks into smaller units. As he points out, everyday activity of a person involves repetition of many sub-tasks such as lifting an object, opening a door, sitting down, walking, and so on. The sub-tasks have features in common and in Singh's method the learning process is merged usefully across them. His illustrative examples refer to robot navigation on a grid, and so the component tasks are less varied than the example of everyday activities suggests, but a general method is developed. Both of these papers receive favourable comment in the Introduction, where that of Lin is said to be one of the largest systematic comparisons of machine learning methods, and that of Singh is said to open up important new directions in which to extend reinforcement learning methods.

In the final paper in the book, reinforcement learning methods are applied in a novel way to robot path finding, in what is termed a non-maze-like environment. The environment is a square area, unobstructed except for a number of circular obstacles, and the paths are traversed by a point robot. The system is able to form situation-action rules that allow the choice of efficient paths. Besides being interesting for its potential application to robotics, this study interestingly breaks new ground by being the first that the editor knows of that combines reinforcement learning with continuous actions.

There is a great deal in this book whose significance is undiminished by the appearance in the meantime of the introductory text of Sutton and Barto.[1] For one thing, convergence proofs are not given in the introductory text although their existence is mentioned. Apart from these, though, it is clear that this book contains many pointers to future developments of the subject area, as well as first steps towards achieving them.

Alex M. Andrew
95 Finch Road,
Earley, Reading RG6 7JX (UK)

Reference

1. R.S. Sutton and A.G. Barto, *Reinforcement Learning: An Introduction*, (MIT Press, Cambridge, Mass., 1998).

https://doi.org/10.1017/S026357479925117X Published online by Cambridge University Press