to appear in: *Proceedings of the IFIP Congress 94 – Volume 3*
to be held in Hamburg, Germany Aug. 28–Sept. 2, 1994.
K. Duncan and K. Krueger (eds.), Elsevier Science Publisher

From AI Technology Research to Applications

Armin B. Cremers, Sebastian Thrun, and Wolfram Burgard

Universität Bonn, Institut für Informatik III, Römerstr. 164, D-53117 Bonn, Germany
FAX: +49-228-550-382, Email: {abc,thrun,wolfram}@cs.bonn.edu

Abstract

Focusing on examples of knowledge systems and machine learning, this paper illustrates the transfer of AI technology from science to real-world applications. Decades of AI research precede a rather short but significant period, in which companies report the useful exploitation of AI technology. This paper illustrates the role played by science, and it argues that AI is just beginning to produce an ever increasingly variety of real world applications.

Keyword Codes: I.2.0; I.2.1; I.2.6
Keywords: Artificial Intelligence, General; Applications and Expert Systems; Learning

1. Introduction

Over the last 30 years, research on Artificial Intelligence (AI) has produced a rich variety of techniques for the acquisition, representation and processing of knowledge. In the early years, research on AI was centered around human intelligence, in particular reasoning and cognition. Today, AI has become a rather broad field, ranging from expert systems and theorem provers to evolutionary algorithms, fuzzy logic and artificial neural networks. Part of AI is still concerned with cognition, and solid progress has been made along this line. Other parts of AI are predominantly concerned with building smart computer systems that can recognize speech, can see, can predict the stock market, can drive robots and so on. A broad set of interests, methodologies and backgrounds has led to some of the most advanced technology for organizing and handling knowledge in computers.

This paper focuses on the transfer of academic research to industrial and commercial applications within AI. Artificial Intelligence is a rather young discipline, and we are just beginning to witness an increased impact of AI technology in the real world. The goal of this paper is to review some of the recent advances in AI, and their applications to real-world problems. In particular, this paper describes recent work on *knowledge systems* and *machine learning*, two major areas which were picked as representative examples for AI technology. As an example, we also describe briefly the BUT system, a knowledge system developed at the University of Bonn. We use this and other examples to elucidate the role of research in the development of AI technology, and its increasing impact on consumer products and industry.

2. Processing Knowledge in Computers: Knowledge Systems

2.1. A Brief History

In the very beginning of knowledge systems, the general research goal was to create computer programs with the power of general problem solvers. It quickly become apparent, however, that

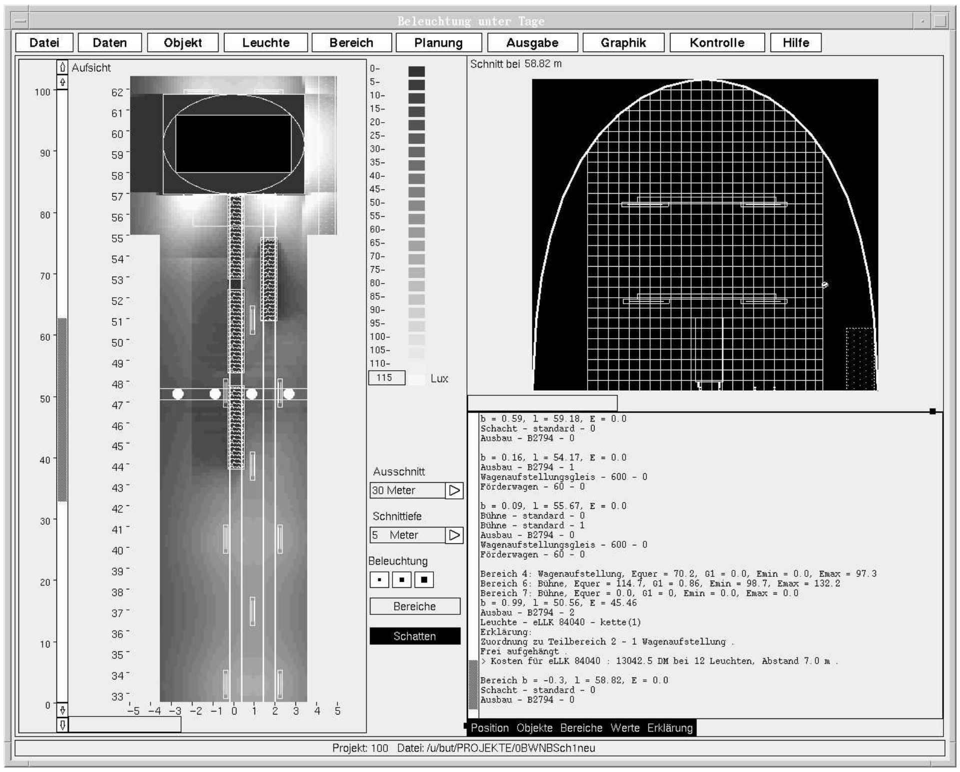

Figure 1. Output generated by the BUT system. Left side: Illumination plan generated by BUT. Right side: Front view of a shaft in a mine.

---

the development of such general-purpose programs was infeasible, as research projects lacked the expected results. Consequently, research begun to focus on rather specific and narrow application areas, in which knowledge systems were remarkably successful. Significant milestones of early knowledge systems research includes applications to medical diagnosis [27], chemistry [4], and the design of computer systems [17]. These and other applications are surveyed in [8,32].

The development of dedicated problems solving procedures for different tasks, such as specialized knowledge representation schemes or inference techniques, led to an exhaustive tool-box which facilitates the integration of knowledge intense problem solving methods into conventional software systems. However, such techniques were not always motivated by industrial applications. For example, the theory of fuzzy sets [37], which is now broadly applied to consumer products in Japan, was originally developed in 1965. Some 10 years later, in 1974, it was applied to control [14], which eventually triggered various industrial applications [28].

AI technology has become an integral component in industry and consumer products. In what follows we will briefly describe a particular knowledge system, BUT [5]. The BUT system is an example for the transfer of AI technology from scientific research to industry.

2.2. BUT: A Real World Application

BUT (short for: *"Beleuchtung unter Tage"*) is a knowledge system that optimizes underground illumination in hard-coal mines. This system was developed in the context of a cooperative research project between the Ruhrkohle AG, Essen and the University of Bonn. The goal of this project was the creation of an intelligent assistant for the cost-optimal layout of light sources in hard-coal mines.

BUT solves the following problem: Given a particular mine layout, which includes size of the mine, and a description of all installed objects and equipment therein, identify the tasks to be performed by the miners. Subsequently, determine which light intensities are required in different areas of the mine. Finally, select lights and lamps, determine appropriate mounting places and optimize the distances of the lights in chains. An acceptable solution is one that satisfies the ergonomic requirements, and an optimal solution does this with minimum costs.

The initial situation, when starting the BUT project, was prototypical for many applications of knowledge systems in industry. Research, carried out by the industrial partner, provided the domain-specific problem solving knowledge, such as heuristics for the derivation of working areas, good mounting places, and a domain-specific approach for computing light intensities. The university research provided the appropriate software technology for implementing and handling knowledge bases and problem solving strategies. For example, in the BUT approach a frame system facilitates the description of the mine layout. In addition, special techniques for geometric reasoning are employed to determine mounting places with an optimal illumination of the working areas. The BUT system, which is the result of this cooperation, integrates both the application-specific knowledge by the industrial partner, and the algorithmic and representational structures provided, ultimately, by the science of AI.

Before the development of BUT, the task was carried out by human experts without any computer support. The most important problem was that faults in the design could only be found *after* the illumination was installed. Deficiencies in the design thus resulted in an expensive change of the installation. For that purpose BUT provides a graphical simulation of the planned scenario, which is shown in Figure 1. The results of the planning process can now directly be checked by the designer in order to eliminate deficiencies and to improve the overall design. A second merit of BUT lies in the automatic minimization of costs. Optimizing costs is too complex a task to be performed efficiently by hand. This is because different types costs, such as the number of lights needed, installation, energy consumptions, maintenance, and cleaning effort would all have to be considered, resulting in an enormous complexity. BUT, which now replaces the previous method, was found to reduce the costs for an illumination by more than 20%, while reducing the actual design time from weeks to minutes.

Traditional knowledge systems require all knowledge to be hand-crafted by human experts, imposing certain limitations on the flexibility and development costs of such systems. Experience with the BUT system indicates that this is a severely limiting factor in applying knowledge systems. One of the goals of AI is to overcome these limitations. In what follows we will describe aspects of machine learning technology, elaborating on using neural network as an example.

3. Machine Learning, Neural Networks, and Their Application to the Real World

3.1. Why Learning?

A second example for the successful transfer of AI technology to the industry is the field of *machine learning*. Machine learning is a strongly interdisciplinary field. Some scientists seek an understanding of how learning in biological systems works, and what role, if any, it plays in intelligence and cognition. Others are more concerned with building advanced systems that can robustly accomplish complex tasks in hard-to-model and potentially time-varying environments. Machine learning has, over the last decades, recruited researchers with various scientific backgrounds, including psychologists, philosophers, engineers, computer scientists, physicists, and mathematicians.

The ultimate goal of machine learning is to build computers that can program themselves. Given a *performance goal*, specified by the user, the computer shall have the ability to gather knowledge through experimentation, such that it will eventually improve in performance. In order to do so, the computer is equipped with *sensors*, which allow it to query information about its environment (like a camera on a robot, or a keyboard on a computer). It is also equipped with *effectors*, which allow it to act and to influence its environment. None of the currently available learning systems truly allows a computer to program itself efficiently from scratch, and there is reason to question whether such programs will ever exist. However, research on machine learning has led to a variety of learning systems that allow programmers to leave certain "gaps" in their programs. These gaps are then, during run-time, filled systematically through observation and experimentation. Typically, gaps are inserted when the programmer faces complex problems which are too expensive to solve manually. This might be because the problem at hand is too complex to program efficiently, or simply because the programmer lacks the knowledge necessary to fill these gaps, knowledge which might be very hard to obtain and thus expensive.

To date, machine learning comprises a broad family of approaches that basically can generalize from observations. *Unsupervised learning techniques*, on the one hand, aim to find statistical regularities that characterize complex probability distributions. On the other hand, *supervised learning techniques* seek to fit unknown functions based on input-output examples. In recent years there has been an increased interest in learning techniques that are beyond these basic paradigms. For example, *hybrid learning techniques* have been developed that integrate knowledge systems, reasoning and learning from examples [10,19]. *Active learning techniques*, to name a second example, refer to learning approaches in which a learner can act, and explore its environment [1,30]. At the same time, remarkable progress has been made in the theoretical understanding of learning from examples [31]. In particular, researchers have related *what* a learner can learn and the amount of training instances required for learning. This trade-off, known as *bias-variance dilemma*, is fundamental in machine learning.

From an applied standpoint of view, there are two fundamental reasons that motivate the use of learning techniques in computer software. First, machine learning techniques often result in much shorter a design time while offering enhanced flexibility. Second, machine learning techniques sometimes outperform human programmers, and hence may be used to improve the overall performance of a system. Both aspects, short design time and superior performance, have been driving forces in applying machine learning technology in practice.

Recently machine learning technology has begun to make its way to real-world applications. It is now being used in a broad spectrum of consumer products and industrial applications. For example, Mangasarian and Street [15] report a learning system which allows to detect breast cancer in a non-invasive way. More specifically, the system is able to detect malign cancer cells in fine needle aspirates by visually inspecting images of cells with close to 100% accuracy. Their system is currently being used in at the University of Wisconsin hospitals, conceivably saving the lives of many women. Another example for a successful implementation may be found in Leech [12]. He reports that a company saves large amounts of money by controlling the processing of fuel using decision tree learning techniques. Mitchell and co-workers [20] have implemented a learning apprenticeship system that gives advice for scheduling meetings with a computerized calendar. The system, which uses hybrid learning techniques for inducing rules, has routinely been used by a secretary over several years. Surveys of industrial applications of machine learning approaches may be found in [3,6,11].

3.2. From Wetware to Software: Neural Networks

Many of to date's successful real-world applications of machine learning technology can be found in the field of *artificial neural networks*. Artificial neural networks, in rough analogy to the human brain, consist of a collection of simple and densely interconnected processing units, which process information in a massively parallel manner. This abstract description, in fact, fits what can been found in the brain, although none of the current approaches is powerful enough to account for the phenomenon of intelligence.

The long way of academic research to successful real-world application is best illustrated by elucidating the history of artificial neural network research. The roots of formal neural network models can be traced back to early work by McCulloch and Pitts [16] and Hebb [9] in the 40's, who established first models of neural processing and plasticity. Minsky [18], in the late 60's, showed some of the limitations of the early methods, which had quite a discouraging effect on the field. At this time there existed virtually no industrial or commercial applications, and the active research community was very small. Finally, in 1986, a book series [25] triggered an "explosion" of world-wide activities in this field, which also produced an impressive variety of industrial and commercial applications. The tremendous growth of this field can be attributed to the following three facts: First, the power of computer hardware improved tremendously, providing sufficient processing time and memory required for interesting simulations. Second, there has been a significant shift of the basic AI paradigm from pure symbolic reasoning towards perceptual issues and learning, which can partially be attributed to the failure of AI to achieve its original, rather ambitious goals. Third, neural computation is appealing, since it promises to exploit the computational power provided by massively parallel computers. Today the field of artificial neural network research covers a rather broad and interdisciplinary field, with roots in computer science, physics, psychology, philosophy, biology, and others.

Artificial neural network research has led to a more profound understanding of learning, both in the human brain and in technical systems, along with a variety of learning methods and applications. The single most widely used neural network algorithm, which has been employed in the vast majority of applications of neural network technology in industry, is the *backpropagation* algorithm [24,34,35]. The term backpropagation was originally introduced in 1986. Since then, a drastically increasing number of researchers has used this and related neural network learning algorithm for fitting functions of arbitrary types. The idea of this learning algorithm is simple enough to be sketched in this paper.

Backpropagation networks realize mathematical functions that map a multidimensional input space to a multidimensional output space. Inside these networks, each processing unit receives values from other processing units. These values are combined linearly, weighted by the so-called *weight values*, and an output value is determined which in turn is passed on to other processing units. The weights, which determine the influence of other units' activations in the process, play a special role in learning, as they are adjusted in the learning process. More specifically, during learning gradients are computed that allow to adapt the weights such as to fit the training data better. Starting with randomly initialized weight values, iterative application of the weight tuning procedure eventually leads to a network that typically closely matches the data. From an external point of view, an artificial neural network may be understood as a black-box, which can fit sets of input-output pairs (the *training data*) by adapting its internal degrees of freedom, the weights.

In the last years rapid progress has been made in the theoretical understanding of neural network learning. One of the key properties of backpropagation networks is that their capacity seems to increase with the magnitude of their weight values. This is because when weights are small, networks realize approximately linear functions, whereas with larger weights the function typically becomes gradually more non-linear. As a result, simple statistical or Bayesian techniques (like cross-validation or weight decay [13,33]) allow the same network to fit very simple functions as well as highly complex ones, yielding good generalization in either case. In the light of these findings, which have been empirically validated by various experimental results, the backpropagation algorithm has proven to be a fairly flexible tool for function fitting.

To date, as scientific research is still very active, we are witnessing an increasing use of backpropagation-style algorithms in industrial and commercial applications. Neural networks are being used to solve a large variety of real-world problems, including various aspects of pattern classification, prediction, control and optimization. One of the earliest applications of neural network was an explosive detection system used to check airport baggage [26]. A similar system learns to steer a vehicle from a human teacher based on camera images [21]. Both systems establish important demonstrations of the potential of machine learning to reduce production time while increasing flexibility. Rennie [22] reports results of neural networks applied to screening pap smears. His approach is able to spot cancerous cells much more reliably than human experts, which are currently employed to do this job. Among other neural network-based medical diagnostic tools, the system is commercially available on the marked and being used by the U.S. Food and Drug Administration. Siemens Corp. has successfully applied neural network technology to the control of steel rolling mills [23]. Their system produces much more homogeneous steel plates than previously used conventional systems. Tesauro reports a neural network-based backgammon playing system that currently leads the championship of computer backgammon programs [29]. Widrow and co-workers, in a survey paper, describe a colorful and impressive variety of applications in which neural network technology proved superior to conventional solutions [36]. They report that linear filters trained with the LMS learning rule [35], which bears close resemblance to the backpropagation algorithm, are used in today's modems to filter out noise in high-speed communication.

In Japan, which has a long history of applying advanced AI technology, many commercial applications have been designed in conjunction with fuzzy control. Asakawa and Takagi [2] report an impressive list of consumer products based on neural network technology, including washing machines, driers, vacuum cleaners, rice cookers, and many more. For example, Stork and co-workers at Ricoh [7] have implemented neural network learning and pruning in commercially available photocopying machines, in which networks are used to control the voltage of several critical components that influence the quality of the prints. Asakawa and Takagi also report various industrial applications, such as financial forecasting, process control, and word processing. Most of these applications have been pursued within the last five years only.

4. Conclusion

In the last years, AI technology has been applied to a variety of real-world problems. In this paper we have outlined some examples from knowledge systems and machine learning/neural networks, along with some recent applications to industry, business and science.Due to lack of space we cannot here expound on other, relevant research areas that also have produced impressive real-world results, such as robotics, vision, speech technology, and evolutionary programming. Indeed, successful ideas and applications that grew out of AI are, once established, often not explicitly recognized as AI technology. The true number of AI applications is unclear, as many companies, like financial forecasting companies or the military, do not publish their techniques and results.

In virtually all successful parts of AI, independent scientific work carried out at research institutions preceded the industrial applications. Without the world-wide interest in matters of the brain, cognition and intelligence, many of current and—more importantly—future technology would not have been and will not be possible. Most of these applications have emerged in the last five to ten years only, based at research that has being carried out at universities and academic institutions for more than 30 years. In recent years AI has undergone a change. Partially based on funding policies of government agencies and industry , partially due to the maturing field, and partially since today's hardware has grown powerful enough, practical applications have gained in importance. Many of today's applications, however, would have been impossible without the decades of fundamental research carried out mainly at universities and academic institutions.

After decades of AI research, we are beginning to witness an ever growing diffusion of AI results into technology, along with the derivation of new products and services. Thus, we firmly believe that a well-defined goal-driven effort in future research is well warranted.

REFERENCES

1. Dana Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, 1988.
2. Kazuo Asakawa and Hideyuki Takagi. Neural networks in Japan. *Communications of the ACM*, 37(3):106–112, March 1994.
3. Ivan Bratko. Applications of machine learning: Toward knowledge synthesis. In *Proceedings of the International Conference on Fifth Generation Computing Systems*, Tokyo, 1992. ICOT.
4. B.G. Buchanan and Feigenbaum E.A. DENDRAL and meta-DENDRAL: Their applications dimension. *Artificial Intellicence*, 11:5–24, 1978.
5. W. Burgard, A.B. Cremers, J. Grebe, R. Grevé, S. Lüttringhaus, F. Mücher, and L. Plümer. Knowledge based planning of underground lighting in hardcoal mining. In R. Vichnevetsky and J.J.H. Miller, editors, *Proceedings of the 13th IMACS on Computers and Applied Mathematics*, pages 1751–1752, 1991.
6. P. Clark, B. Cestnik, C. Sammut, and J. Stender. Applications of machine learning: Notes from the panel members. In *Proceedings of the Sixth European Working Session on Machine Learning*. Springer Publisher, 1991.
7. Babak Hassibi, David G. Stork, and Takahiro Watanabe. Optimal brain surgeon: Extensions, streamlining and performance comparisons. In *Advances in Neural Information Processing Systems 6*, San Mateo, CA, 1994. Morgan Kaufmann.
8. F. Hayes-Roth and N. Jacobstein. The state of knowledge-based systems. *Communications of the ACM*, 37(3):27–39, March 1994.
9. D. O. Hebb. *The Organization of Behavior*. Wiley, NewYork, 1949.
10. John Laird, Paul Rosenbloom, and Allen Newell. Chunking in SOAR: The anatomy of a general learning mechanism. *Machine Learning*, 1(1):11–46, 1988.
11. Pat Langley. Areas of application for machine learning. In *Proceedings of the Fifth International Symposium on Knowledge Engineering*, Sevilla, 1992.
12. W. J. Leech. A rule-based process control method with feedback. *Advances in Instrumentation*, 41:69–175, 1986.
13. David J. C. MacKay. *Bayesian Methods for Adaptive Models*. PhD thesis, California Institute of Technology, Pasadena, California, 1992.
14. E.H. Mamdani. Applications of fuzzy algorithms for control of simple dynamic plant. *Proc. IEE*, 121(12):1585–1588, 1974.
15. Olvi L. Mangasarian, W. Nick Street, and William H. Wolberg. Breast cancer diagnosis and prognosis via linear programming. In *AAAI Spring Symposion on AI in Medicine*, Stanford, CA, March 1994.

16. W. S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.
17. John McDermott. Domain knowledge and the desing process. In *ACM/IEEE 13th Design Automation Conference Proceedings*, 1981.
18. M. Minsky and S. Papert. *Perceptrons*. The MIT Press, Cambridge, Massachusetts, 1969.
19. Steven Minton. *Learning Search Control Knowledge: An Explanation-Based Approach*. Kluwer Academic Publishers, 1988.
20. Tom M. Mitchell, Rich Caruana, Dayne Freitag, John McDermott, and David Zabowski. Experience with a learning personal assistant. *Communications of the ACM*, July 1994. (to appear).
21. Dean A. Pomerleau. Rapidly adapting neural networks for autonomous navigation. In R. P. Lippmann, J. E. Moody, and D. S. Touretzky, editors, *Advances in Neural Information Processing Systems 3*, pages 429–435, San Mateo, 1991. Morgan Kaufmann.
22. J. Rennie. Cancer catcher: Neural net catches errors that slip through pap tests. *Scientific American*, 262, May 1990.
23. Martin Röscheisen, Reimar Hofmann, and Volker Tresp. Neural control for rolling mills: Incorporating domain theories to overcome data deficiency. In J. E. Moody, S. J. Hanson, and R. P. Lippmann, editors, *Advances in Neural Information Processing Systems 4*, pages 659–666, San Mateo, CA, 1992. Morgan Kaufmann.
24. David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning internal representations by error propagation. In D. E. Rumelhart and J. L. McClelland, editors, *Parallel Distributed Processing. Vol. I + II*. MIT Press, 1986.
25. David E. Rumelhart and James L. McClelland. *Parallel Distributed Processing. Vol. I + II*. MIT Press, Cambridge, Massachusetts, 1986.
26. Patrick M. Shea and Vincent Lin. Detection of explosives in checked airline baggage using an artificial neural system. *International Journal of Neural Networks*, 1(4):249–253, 1989.
27. E.H. Shortliffe. *Computer based medical Consultations: MYCIN*. American Elsevier, 1976.
28. P. Sugeno. An introductory survey of fuzzy control. *Information Sciences*, 36:59–83, 1985.
29. Gerald J. Tesauro. Practical issues in temporal difference learning. *Machine Learning Journal*, 8, 1992.
30. Sebastian B. Thrun. The role of exploration in learning control. In David A. White and Donald A. Sofge, editors, *Handbook of intelligent control: neural, fuzzy and adaptive approaches*. Van Nostrand Reinhold, Florence, Kentucky 41022, 1992.
31. Leslie G. Valiant. A theory of the learnable. *Communications of the ACM*, 27:1134–1142, 1984.
32. Donald A. Waterman. *A Guide to expert systems*. Addison-Wesley Publishing Company, 1986.
33. Andreas Weigend. On overfitting and effective number of hidden units. In M. Mozer, P. Smolensky, D. Touretzky, J. Elman, and A. Weigend, editors, *Proceedings of the 1993 Connectionists Summer School*, pages 335–342, Hillsdale, NJ, December 1993. Lawrence Erlbaum Associates Publisher.
34. Paul Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavorial Sciences*. PhD thesis, Harvard University, Committee on Applied Mathematics, Cambridge, MA, November 1994.
35. Bernard. Widrow and M. E. Hoff. Adaptive switching circuits. Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record, Part4, 1960.
36. Bernard Widrow, David E. Rumelhart, and Michael A. Lehr. Neural networks: Applications in industry, business and science. *Communications of the ACM*, 37(3):93–105, March 1994.
37. L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.