RC 21907 (98532) 7 December 2000 **Computer Science**

IBM Research Report

AI at IBM Research

Chidanand Apte, Leora Morgenstern, Se June Hong
IBM Research Division
Thomas J. Watson Research Center
Yorktown Heights, NY 10598

AI at IBM Research

Chidanand Apte Leora Morgenstern Se June Hong

T.J. Watson Research Center
IBM Research Division
Yorktown Heights, NY 10598
e-mail: {*apte,leora,sjhong*}*@us.ibm.com*

Abstract

IBM has played an active role in AI research since the field's inception more than 50 years ago. In a trend that reflects the increasing demand for applications that behave intelligently, IBM today carries out most AI research in an interdisciplinary fashion by combining AI techniques with other computing techniques to solve difficult technical problems.

1 Introduction and Background

IBM's contributions in Artificial Intelligence research date back to the 1950s. Arthur Samuels' Checkers program [43] made history by demonstrating that a computer program could play checkers well enough to beat human experts, and could learn from its experience in playing against humans to become a better player. Gelernter's Geometric Theorem Prover [14] used axioms about geometry and diagrammatic information to prove basic theorems about geometry.

Since the early days of AI, IBM has had a rich history of research across a broad spectrum of AI. Research in this area became especially active after IBM formed an AI group in the early 1980s. Notable contributions in the 1970s, 1980s, and early 1990s include systems for querying databases in natural language (TQA [10]); computer algebra systems (Scratchpad [17]), and the subsequent commercial AXIOM [24] system), robotics [31], conceptual graphs [48], knowledge representation (specifically K-REP [32], a frame-based system that was used to develop pharmacological knowledge bases), expert systems [27], epistemic logics [12], and nonmonotonic reasoning [49].

For many years, AI research at IBM was mostly based on the symbolic paradigm. While such research continues today, we have seen in recent years a trend toward the increased use of statistics in AI research, particularly for such applications areas as machine learning and natural language processing. This has led to the growth of new areas such as Statistical Learning Theory and Bayesian networks as active areas of inquiry. We report on recent research in both traditional and newer areas of AI in the following section.

2 AI Activities in IBM Research

We introduce the extent of AI at IBM computer science research in four broad areas; Representation and Reasoning, Statistical AI, Vision, and Game-Playing. We write about each of these areas:

2.1 Representation and Reasoning

IBM's research in knowledge representation spans a wide variety of areas, including logic programming, nonmonotonic reasoning, planning and scheduling, and intelligent agents. These are discussed below.

2.1.1 Logic Programming and Nonmonotonic Reasoning

IBM has a venerable tradition in logic programming and nonmonotonic reasoning. In the 1980s, research at the Watson Labs was at the forefront of the development of constraint logic programming, a variant of logic programming which facilitates arithmetical reasoning as well as general reasoning about computational systems based on constraints [23].

Recent research in logic programming includes the work of Benjamin Grosof, Hoi Chan, and their colleagues, in *courteous logic programs* [18]. Courteous logic programs allow the specification of the scope of potential conflicts via pairwise mutual exclusion. They also allow the specification of partially-ordered priorities between rules; these priorities can be used to resolve conflicts between rules. While courteous logic programs are a form of prioritized logic programs, and in general, a type of prioritized default reasoning, they do not share the high complexity of other expressive forms of prioritized defaults.

Courteous logic programs have been implemented in Common Rules, a Java library that enables the communication of business policies about pricing and promotions between e-business retailers and customer applications. CommonRules, available at [22], includes an XML rule-interchange format for business rules, called Business Rules Markup Language (BRML). It is being piloted by EECOMS, a $29 million three-year industry consortium development on enterprise supply chain integration for manufacturing.

Default reasoning about business rules has also been studied by Leora Morgenstern and Moninder Singh. They have developed *Formula-Augmented Networks (FANs)*, a knowledge structure that enables efficient reasoning about potentially conflicting business rules. FANs extend the traditional paradigm of semantic networks and inheritance with exceptions [20] by allowing the attachment of logical formulas at nodes in the network, and providing methods to determine which formulas "apply" to a particular node. In cases of conflict, these methods determine which rules can be saved and which ought to be discarded. The decision involves information inherent in the topology of the network, along with specialized rules giving preferences among classes of business rules.

The FANs paradigm works well for business domains, since it is often quite straightforward to interpret business cases or situations as nodes in some network, and business rules as logical formulas. FANs have been used at IBM to develop BenInq, an expert system for benefits inquiry in the medical insurance industry, and ProdCat, a system to dynamically configure banking and insurance products [35].

2.1.2 Networks and Other Schemes for Natural Language Understanding

Various projects at IBM have focused on using networks and other structures in innovative ways to improve performance in the difficult task of text and story understanding. At the IBM research labs in Zurich, Nobel Laureate Gerd Binning, along with Juerg Klenk, leads the MeanX project, which seeks to develop a system to understand meaning in natural language. The research team is trying to depart from the standard natural language processing paradigm, which centers around anaphora resolution, word disambiguation, and object identification. Instead, the system centers around the use of world knowledge to achieve comprehension.

MeanX uses a slot grammar and a set of transformation rules to generate an "input network" from textual input. The innovation lies in the transformation of the "input network" to an "output network" that represents the understood textual input. Central to this part of the analysis is the use of *fractal networks*, self-similar hierarchical networks of arbitrary order that can organize input text into structures that are isomorphic to existing world-knowledge fractal networks. The isomorphism facilitates comparison and thus comprehension of the text.

There are several other efforts at IBM Research to use KR structures to understand natural language. Andrew Gordon's Indexing Stories project seeks to organize a large knowledge base of stories by agents' strategies and motivations. Gordon has developed a representational language for representing strategies which is closely modeled on standard AI planning techniques. In addition to indexing stories, this representational language has applications in the interactive entertainment domain, where explicit representation of and reasoning about an adversary's strategies can enrich the interactive experience [15]. Erik Mueller's work on story understanding uses natural language templates and general principles about agents' cognitive behavior to understand an agent's goals and actions. Mueller's system combines state-of-the art natural language processing with commonsense knowledge about the way agents interact. Mueller's work extends his earlier work [36], which fuses AI techniques with cognitive psychology.

2.1.3 Intelligent Agents: PriceBots, Planning, and Scheduling

IBM's research in intelligent agents focuses both on the development of core technology to facilitate the quick implementation of intelligent agents, and on exploring foundational principles for such applications as auctions and scheduling.

The ABLE (Agent Building and Learning Environment) project [5], led by Joe Bigus, focuses on building hybrid intelligent agents that can both reason and learn. The ABLE framework consists of a set of core JavaBeans and a set of function-specific JavaBeans. Application-level agents (realized as JavaBeans) can then be constructed out of these components using either the ABLE editor or any standard bean-building environment. Project plans call for building a platform compatible with FIPA (Foundations for Intelligent Physical Agents [13]) standards in Java.

Intelligent agents technology forms a central part of the Information Economies project. This project, headed by Jeff Kephart, studies the role of intelligent agents in setting prices in a variety of situations such as auctions and negotiations. Such situations can vary widely in terms of the amount of knowledge that can be assumed of the participants and the amount of computation and reasoning that the participants can be expected to perform. The group has studied different strategies based on different sets of assumptions, and have formally analyzed the behavior of the algorithms based on these strategies. The theoretical analysis is complemented by the implementation of PriceBots and ShopBots, intelligent agents which can be programmed to implement a particular strategy for a retailer or consumer. Kephart and his colleagues have also studied the role of reinforcement learning, information filtering, and information bundling in an e-commerce environment [16].

Intelligent agents technology has also been used for supply-chain management and planning and scheduling problems. Sesh Murthy, Richard Goodwin, Pinar Keskonoak, and their colleagues have examined the difficult problem of scheduling multiple machines when there are multiple objectives and constraints. The particular problem examined is scheduling paper manufacturing jobs where such factors as sequence-dependent setups, job-machine restrictions, batch size preferences, and downstream scheduling consequences have to be considered simultaneously. This is a difficult problem since achieving optimal solutions even in cases where one is only considering minimizing total tardiness on one machine can be NP-hard. This group tackled the problem by developing the *A-Team* architecture, an agents architecture in which agents — solution constructors, improvers, and destroyers — cooperate by exchanging results. Various algorithms were implemented using this basic architecture, and experimental results have shown that impressive results can be achieved in a computationally efficient manner. This group's work has received the prestigious Daniel H. Wagner Prize for Excellence in Operations Research Practice for their contributions [29].

2.2 Statistical AI

We use the term *Statistical AI* to denote several areas of research where the methodologies of AI and Statistics are brought together to achieve higher levels of intelligent performance. The domains where this integration is most noticeably effective are Speech and Handwriting Recognition, Natural Language Understanding, Knowledge Discovery and Data Mining (KDD), and related applications of Machine Learning.

2.2.1 Speech and Handwriting Recognition

Speech recognition technologies have been pioneered at IBM since the early ’70s with one of the largest groups in the world dedicated to solving problems in speech and language using unique statistical modeling approaches. A long term research effort initiated by Jelinek and continued by Nahamoo, Bahl, Mercer, Picheny, Roukos et al, [3, 38, 37] has led to IBM’s highly successful retail dictation product, ViaVoice. This product is available in 12 different languages. Other recognition projects include transcription of radio broadcast news and people’s voicemail. Related activities include speech biometrics (speaker verification and identification), speech synthesis, and audio-visual speech and speaker recognition [44].

Emerging as a new thrust of research activity is Conversational Systems, which combines our speech recognition technology with natural language processing. This includes multi-modal dialog management, natural language generation, and statistical spoken language understanding. Ultimately, this work will lead to systems in which users speak in their natural tongue to control the application in their own personal style. Our emphasis is on scalable dialog systems capable of handling many simultaneous channels, applications, and input modalities (voice, keyboard, gesture, and mouse) [11]. We are prototyping systems for many applications ranging from stock and mutual fund trading systems, to phone banking, air travel reservations, and web-based shopping and in many languages including English, French, German, Mandarin, and Spanish.

Many prefer handwriting to typing, and it is clear that pen and paper are better suited for many tasks because handwriting is frequently less intrusive, less noisy, more flexible, more convenient and simpler than current computers alternatives. Although computers have become ubiquitous, it is clear that pen and paper are still preferred for many tasks because pen and paper note taking is frequently less intrusive, less noisy, more flexible, more convenient and simpler than current computers alternatives. The Pen Technologies Group at IBM’s T.J. Watson Research Center believes that there is a tremendous market for pen computing systems that are capable of satisfying users’ high expectations. Towards the goal of creating such systems, the Pen Technologies Group focuses its research on all aspects of pen as a user interface. Recent and on-going algorithmic research include methods for performing handwritten document retrieval and improving handwriting recognition accuracy [39, 50], and statistical language modeling [40].

2.2.2 Natural Language Processing

Natural Language Processing (NLP) at IBM is a dynamic research area spanning a wide range of topics, frequently overlapping with projects in the Human Computer Interaction and Data Mining communities. Over the years, we have pioneered algorithms and systems that have significantly increased the community’s understanding of the field and created new research directions. These accomplishments include an early machine translation system used for many years by the Air Force (AN/FSQ-7), our TQA [10] natural language database query system, the source-channel statistical machine translation paradigm [7], grammar-free statistical parsing, mathematical investigations of compositionality in natural language, and new algorithms for text categorization based on symbolic rule induction and statistical theories of pattern recognition. Currently, we are working on theoretical issues of computational linguistics as well as practical algorithms for text analysis, machine translation and knowledge management.

Algorithms are being developed by Prager and colleagues [41] that can extract useful information from enormous collections of documents like the web. Many of these investigations are being driven by the question answering paradigm: envision that a user has a specific question in mind and that there are many valid answers with differing levels of granularity; paragraphs, individual sentences, key phrases, and summaries.

Research in lexical navigation [9] permits a user to find related documents and see their connections to the original document. This is based on a set of techniques to extract vocabulary items for named entities (people, places, organizations, domain terms, dates, etc.) along with the named relations (CEO, teacher, etc.) that link them, using lexical networks for linking the concepts.

Although the amount of information present on the web and other computer databases is exploding rapidly, an individual's ability to read and absorb it remains essentially fixed. Because of this, we have developed machine learning algorithms which mine recurring patterns and linguistic associations from this wealth of data. For example, we have developed a new approach to linear classification [55] that trains very quickly and is one of the most accurate. Recent applications include text categorization for email routing and auto-response systems, and interactive search for the web based on a novel concept of dialog categorization. We have developed a current state-of-the-art algorithm for automatic new topic discovery in newswire feeds and a help center's productivity aid that analyzes call databases and automatically identifies topics associated with an increased numbers of calls. Further, we are developing algorithms for structuring extracted information to support automatic database updates from unstructured text. Our intent is that future NLP systems will be adapted to new domains without mandatory manual intervention, thereby enabling out-of-the-box applications.

Machine translation (MT) technology is becoming crucial due to the wide usage of the Internet for exchanging information and communicating between people in the world. Our MT effort has three principal focuses. First, we conduct research on fully automatic machine translation algorithms [33]. Second, we explore methods of assisting the manual translation process by source-side linguistic annotation [53]. Third, the EasyEnglishAnalyzer tool [4] has been developed for grammar checking and controlled language checking for use as a pre-editing step to make texts more machine translatable [4]. We are also developed extremely rapid corpus-based statistical machine translation algorithms specifically tuned for cross-lingual information retrieval from large multilingual document collections. Several of these efforts take advantage of Slot Grammar [34] based bilingual lexicons and transformational systems to produce more accurate machine translation. Broad coverage Slot Grammar parsers in several European and Asian languages are being currently developed. The English Slot Grammar parser is also the basis for grammar checking and controlled language checking.

2.2.3 Machine Learning and Related Areas

The ongoing rapid growth of on-line data due to the Internet and the widespread use of databases have created an immense need for Knowledge Discovery and Data Mining (KDD) methodologies. The challenge of extracting knowledge from data draws principally upon research in statistics, data management, pattern recognition, and machine learning. IBM Research has been at the forefront of this exciting new area from the very beginning. Key advances in robust and scalable data mining techniques, methods for fast pattern detection from very large databases, text and web mining, and innovative business intelligence applications have come from our worldwide research laboratories.

A key focus area in KDD research at IBM continues to be on high performance scalable data mining techniques for large-scale databases and data repositories. IBM's early lead in this area was established by the Quest (Agrawal, Srikant, Shafer, and colleagues [1, 45]) project, when the association rule and sequential patterns technology for efficiently detecting patterns in large-scale databases and fast scalable decision tree modeling was invented. These and other technologies for scalable and parallel data mining developed as part of this project provided the original basis and impetus for IBM's flagship data mining product, Intelligent Miner.

Another principle focus area of research in KDD is that of supervised learning systems (or predictive data mining). One long-term focus area of researchers like Apte, Hong, Pednault, and Weiss has been on rule based predictive modeling and its integration into data mining frameworks [19, 54, 2]. This work has resulted in new data mining middleware for rule based probabilistic estimation (the ProbE framework of Pednault and colleagues [2]), which combines machine learning with principles from statistical learning theory and data management. This technology has been embedded in innovative business intelligence applications for areas such as insurance risk management and retail targeted marketing [2].

Another focus area is applications of machine learning to Performance Management. It is necessary to manage many kinds of computer services, such as web transactions and electronic mail, in order to guarantee good service and availability. Research at IBM seeks to increase the level of automation for performance management. Work conducted in the early 80s included the YES/MVS expert system (Hong, Kastner, and colleagues [26]) for automated management of MVS operations consoles. Current work includes the study of predictive detection (Hellerstein et al. [25, 52]), which gives advanced knowledge of performance degradations, and event mining [30], a form of data mining that recognizes situations which require specific action to ensure good performance.

Intelligent tutoring systems construct models of a student's understanding and interact with a student based upon this model. Because learning often takes place in a group setting, research scientists at IBM [47]) are expanding this paradigm by considering the interaction among a team of students. They are developing an intelligent tutoring system that monitors and manages teams of students as they collaborate to solve distributed, multi-step problems. Techniques include defining a typology of collaborative problem-solving roles, and storing evidence about individual and group problem-solving performance in a Bayesian network.

A closely related area to machine learning is applications of pattern recognition techniques in Computational Biology solutions. Discovery of repeated patterns in amino acid sequences and gene function identification uses advanced pattern matching algorithms such as TERESIAS [42] developed at IBM's bioinformatics group which actively pursues the mining of information from such sequences.

2.3 Vision

Intelligent vision requires the ability to understand the image that a camera or other sensory device produces. This requires providing the computer with a representation that it can manipulate, along with the background knowledge, context, and computational and inferential methods that it needs to understand the image. The Exploratory Vision group at IBM Research, headed by Ruud Bolle, studies how vision systems can be used for applications ranging from biometrics to video browsing to automated object recognition at supermarkets.

For these varied applications, researchers rely on several innovative technologies. IBM research scientists have developed a method to transform an ordinary video camera into a range-based segmenter [6]. This allows such functions as hand pose decoding, fingertip tracking, gesture recognition, and even imaging a user's iris.

Researchers have also developed methods to recognize objects even when they are partly occluded. Other technologies developed include semi-automated color segmentation techniques based on simulated annealing, methods for visualizing motion in video, methods for face and gesture recognition, and innovative sensors and motion trajectory extrapolation for tracking a person's motion.

These techniques are foundational to several high-profile projects within IBM Research.

The Biometrics project explores the ways in which security systems can be enhanced by fingerprint recognition and face recognition techniques.

The Video Analysis project explores the use of video browsing[46]. Video browsing is functionally akin to searching for strings in bodies of text, but is much harder to perform with high accuracy and within reasonable time constraints. Face recognition and event recognition techniques are central to this project.

The Human and Computer Interaction aims to allow a camera, aimed at a person, to be employed as a user-interface device. The user can communicate by aiming his face, touching a target, or using specific hand gestures such as pointing or grasping. This work is intended partly for people such as the severely handicapped who have trouble with traditional user-interface devices. Face and gesture recognition technologies are central to this project as well.

The VeggieVision project explores the use of automated object recognition at the supermarket. The particular application studied is the recognition of vegetables at the checkout register. The problem is particularly difficult because vegetables must be recognized through the traditional supermarket plastic bag, which is milky and partly occludes the object(s) within. VeggieVision uses a combination of background subtraction and chromakeying to separate the vegetable from the superfluous information. It then concatenates all relevant information into a feature vector, uses a nearest-neighbor matcher to find a comparable item in its memory base, and registers that number. In case of several close matches, the clerk is asked to choose the correct match. Over time, the system learns from its mistakes. VeggieVision is currently being field tested by scanner manufacturers.

2.4 Game-Playing Programs and Applications

IBM's history in game playing began when Arthur Samuels developed an expert checkers-playing program that learned from experience [43]. More than forty years later, IBM Research continues to break new ground in developing game-playing programs.

2.4.1 Chess

Perhaps the most famous of IBM's game-playing programs in Deep Blue [8], which made headlines when it beat Gary Kasparov, thus becoming the first chess-playing program to defeat a reigning world champion. Developing a champion-class chess-playing program had long been a challenge problem for AI. Game-playing programs have traditionally used a combination of evaluation functions to determine how good a particular position is, and effective search techniques to search through a space of possible states. The difficulty in developing a champion-level chess-playing program lies in the facts that first, the state space is so large (since the *branching* factor or ratio, the number of states that can be reached from a given state of the game, is 30-40) and second, the sophisticated evaluation of chess positions can be difficult to develop and computationally expensive to perform.

Deep Blue, developed by C. J. Tan, Murray Campbell, Feng-hsiung Hsu, Joe Hoane, Jerry Brody, and U.S. chess champion Joel Benjamin, relies on a combination of brute force and sophisticated knowledge of chess. Fundamental to its success is a 32-node IBM RS/6000 SP high-performance computer which uses Power Two superchip processors (P2SC). Each node yields 256 processors which work in tandem. The net result is a system capable of calculating 200 million moves per second. Deep Blue's understanding of chess strategy is implemented by its evaluation function, which considers material (the value of a particular chess piece), king safety (a measure of the defensive strategy), board position, and tempo (a measure related to the aggressiveness of a player's strategy, relative to the way the game unfolds).

2.4.2 Backgammon

While backgammon is seemingly a simpler game than chess, developing a champion-level backgammon-playing program has also proved difficult. Indeed, the random element introduced into the game by the rolling of dice (which makes backgammon an apparently less intellectual game than chess) makes the branching ration of backgammon much higher than than of chess: several hundred as opposed to 30-40). Thus, a solution cannot (yet) rely on brute force computing power as its primary technique. Gerald Tesauro has developed TD-Gammon [51], a champion-level backgammon-playing program that relies on a form of reinforcement learning known as *temporal difference learning* or TD-learning. TD-Gammon explores how a TD-learning machine trains a multi-layer neural network to learn complex non-linear functions. In this case, these functions are the evaluation functions that TD-Gammon uses to evaluate game positions and plan its strategy.

Interestingly, TD-Gammon does not rely on sophisticated evaluation functions learned from experts. Because experts often change their minds about the value of a particular board position, and because the probabilistic nature of the game means that experts may sometimes only rarely see some interesting game configurations, Tesauro believes that it is best for the backgammon-playing program to learn its evaluation function on its own.

2.4.3 Applications of Game-Playing Programs

Game-playing programs are valued at IBM Research not only because they solve challenging AI problems, but because they lay the foundations for crucial applications and development. The technology behind Deep Blue is central to the ongoing work at IBM's Deep Computing Institute [21], which seeks to integrate the power of intensive computation with advances in algorithms, analytic methods, and other areas for applications in such areas as computational biology, bioinformatics and pattern discovery, data mining, and optimization. Likewise, the neural network technology developed for TD-Gammon has been used to enhance IBM's anti-virus software [28]. On the horizon is BlueGene, a petaflops machine being developed in IBM Research, which we hope to leverage to extend our horizon in computational aspects of AI techniques.

3 Concluding Remarks

In contrast to its earliest years, when Artificial Intelligence was often considered a luxury, a novelty, or a science-fiction fantasy, AI is now recognized to be a central part of computer science. Much of the transformation in attitude is due to the changing nature of the computer in people's lives. As computers and computing devices become ubiquitous in the everyday world, applications need to become easier to use for non-technical users. User-friendly applications often need to understand and anticipate the needs of the users, and thus in some sense must behave intelligently. AI has therefore become an integral part of the information revolution. At the same time, it has become clear that it is the combination of AI with other disciplines that will work most successfully to solve challenging technological problems.

These considerations are reflected both in today's technological landscape, and in academic settings. AI techniques can be found in many IT projects, and a large proportion of the subdisciplines in a university's computer science department are often AI-related.

These trends are evident in IBM as well. Because IBM research today is largely motivated by user-driven and market-driven concerns, projects are generally organized by problems in technology rather than by traditional academic disciplines. However, we are also working hard to ensure a sense of community among scientists with similar research interests. To that end, IBM Research has in the past few years organized Professional Interest Communities (PICs) which closely parallel traditional academic disciplines. PICs augment but do not replace our traditional project organization.

AI is a strong component of both the project and the PIC organization at IBM Research. AI and AI-related technologies play a major role in many projects in computer science research at IBM. In addition, at least eight of the twenty-some current PICs are closely related to AI. Besides the Artificial Intelligence PIC, we have PICs in the areas of Knowledge Discovery and Data Mining, e-Commerce, Human Computer Interaction, Natural Language Processing, Performance Modeling and Analysis, Web, and Computational Biology. Information on these PICs is available on the internet at http://www.research.ibm.com/compsci. Additional websites (all reachable from the previous top-level site) include http://www.research.ibm.com/compsci/ai for AI, http://www.research.ibm.com/compsci/kdd for KDD, http://www.rsearch.ibm.com/compsci/nlp for NLP, and http://www.research.ibm.com/compsci/hci for Speech and Handwriting Recognition.

Acknowledgment

We thank the many AI researchers at IBM whose work we have introduced here and who provided us with useful information for this paper.

References

[1] R. Agrawal, T. Imielinski, and A. Swami. Mining Association Rules between Sets of Items in Large Databases. In *Proceedings, ACM SIGMOD Conference on Management of Data*, pages 207–216, 1993.

[2] C. Apte, E. Grossman, E. Pednault, B. Rosen, F. Tipu, and B. White. Probabilistic Estimation Based Data Mining for Discovering Insurance Risks. *IEEE Intelligent Systems*, 14(6), November/December 1999.

[3] L.R. Bahl, P.V. de Souza, P.S. Gopalakrishnan, D. Nahamoo, and M.A. Picheny. Robust Methods for Using Context-Dependent Features and Models in a Continuous Speech Recognizer. In *Proc. of the Int. Conf. on Acoustics, Speech, and Signal Processing*, volume 1, pages 533–536, 1994.

[4] A. Bernth. EasyEnglish: A Tool for Improving Document Quality. In *Proceedings of Fifth Conference on Applied Natural Language Processing, Association for Computational Linguistics*, pages 159–165, 1997.

[5] J. Bigus. Agent Building and Learning Environment Project. http://www.research.ibm.com/able, 2000.

[6] R. Bolle, J.H. Connell, N. Haas, R. Mohan, and G. Taubin. Object Imaging System. U.S. Patent No. 5,631,976, 1997.

[7] P.F. Brown, J. Cocke, A. Stephen, S.A. DellaPietra, V.J. DellaPietra, F. Jelinek, J. Lafferty, R. Mercer, and P.S. Roosin. A Statistical Approach to Machine Translation. *Computational Linguistics*, 16:79–85, 1990.

[8] M. Campbell. Knowledge Discovery in Deep Blue. *Communications of the ACM*, 42(11):65–67, November 1999.

[9] J. Cooper and R. Byrd. Lexical Navigation - Visually Prompted Query Expansion and Refinement. In *Proceedings of DIGLIB 97*, 1997.

[10] F. Damerau. The Transformational Query Answering System Operational Statistics. *American Journal of Computational Linguistics*, 7:30–42, 1981.

[11] K. Davies. The IBM Conversational Telephony System For Financial Applications. In *Proceedings of EuroSpeech-99*, volume 1, pages 275–278, 1999.

[12] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning About Knowledge*. MIT Press, Cambridge, 1995.

[13] FIPA. Website of foundation of intelligent physical agents. http://www.fipa.org, 2000.

[14] H. Gelernter. Realization of a Geometry Theorem Proving Machine. In *Proceedings of the International Conference on Information Processing*, pages 273–282, 1959.

[15] A.S. Gordon. Playing Chess with Machiavelli: Improving interactive entertainment. In *Proceedings of the 2001 Spring Symposium on AI and Interactive Entertainment*, 2001.

[16] A. Greenwald and J. Kephart. Shopbots and Pricebots. In Thomas Dean, editor, *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence*, pages 506–511. Morgan Kaufmann, 1999.

[17] J.H. Griesmer and R.D. Jenks. The SCRATCHPAD System. In *Proceedings of ACM '75*, pages 266–267, October 1975.

[18] B. Grosof, Y. Labrou, and H.Y. Chan. A Declarative Approach to Business Rules in Constraints: Courteous Logic Programs in XML. In Michael P. Wellman, editor, *Proceedings of the First ACM Conference on Electronic Commerce, EC-99*. ACM Press, 1999.

[19] S.J. Hong. R-MINI: An Iterative Approach for Generating Minimal Rules from Examples. *IEEE Transactions on Knowledge and Data Engineering*, 1997.

[20] J.F. Horty, R. Thomason, and D. Touretzky. A Skeptical Theory of Inheritance in Nonmonotonic Semantic Networks. *Artificial Intelligence*, 42:311–349, 1990.

[21] IBM. Deep Computing Home Page. http://www.research.ibm.com/dci, 2000.

[22] IBM. IBM alphaWorks. http://www.alphaworks.ibm.com, 2000.

[23] J. Jaffar and J.L. Lassez. Constraint Logic Programming. In *Proceedings of POPL-87, Conference on Principles of Programming Languages*. 1987.

[24] R.D. Jenks and R.S. Sutor. *AXIOM: The Scientific Computation System*. Springer-Verlag, New York and Numerical Algorithms Group Limited, Oxford, England, 1992. ISBN #0-387-97855-0, ISBN #3-540-97855-0, 742 pages.

[25] T.S. Jayram J.L. Hellerstein and I. Rish. Recognizing End-User Transactions in Performance Management. In *Proceedings of AAAI2000*, 2000.

[26] J.K. Kastner. YES/MVS A Contunuous Real Time Expert System. In *Proceedings of AAAI-1984*, 1984.

[27] J.K. Kastner, C. Apte, J.H. Griesmer, S.J. Hong, M. Karnaugh, E. Mays, and Y. Tozawa. A Knowledge-Based Consultant for Financial Marketing. *AI Magazine*, 7(5):71–79, 1986.

[28] J. Kephart, G. Sorkin, W. Arnold, D. Chess, G. Tesauro, and S. White. Biologically Inspired Defenses Against Computer Viruses. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)*, pages 985–996. Morgan Kaufmann, 1995.

[29] P. Keskinocak, S. Murthy, R. Akkiraju, R. Goodwin, J. Rachlin, F. Wu, and J. Yeh. Cooperative Multiobjective Decision Support for the Paper Industry. *Interfaces*, 29(5):5–30, 1999.

[30] S. Ma and J.L. Hellerstein. Mining Partially Periodic Event Patterns With Unknown Periods. In *International Conference on Data Engineering*, 2000.

[31] S. Mahadevan and J. Connell. Automatic Programming of Behavior-Based Robots Using Reinforcement Learning. *Artificial Intelligence*, 55(2):311–365, 1992.

[32] E. Mays, R. Weida, R. Dionne, M. Laker, B. White, C. Liang, and F. J. Oles. Scalable and Expressive Medical Terminologies. 1996.

[33] A. Bernth M.C. McCord. The LMT Transformational System. pages 344–355, 1998.

[34] M.C. McCord. Slot Grammars. 6:31–43, 1980.

[35] L. Morgenstern. Inheritance comes of age: Applying nonmonotonic techniques to problems in industry. *Artificial Intelligence*, 103(1–2):237–271, 1998.

[36] E.T. Mueller. *Daydreaming in Humans and Machines: A Computer Model of the Stream of Thought*. Ablex/Intellect, 1990.

[37] J. Navratil, J. Kleindienst, and S.H. Mae. An instantiable speech biometrics module with natural language interface: Implementation in the telephony environmen. In *Proceedings of ICASSP 2000, Istanbul, Turkey*, June 2000.

[38] C. Neti, G. Iyengar, G. Potamianos, A. Senior, and B. Maison. Perceptual interfaces for information interaction: Joint processing of audio and visual information for human-computer interaction. In *Proceedings of the International Conference on Spoken Language Processing*, volume III, pages 11–14, 2000.

[39] M.P. Perrone and S.D. Connell. K Means Clustering for Hidden Markov Models. In *Proceedings of Internationall Workshop on Frontiers in Handwriting Recognition*, 2000.

[40] J.F. Pitrelli and E.H. Ratzlaff. Quantifying the Contribution of Language Modeling to Writer Independent On Line Handwriting Recognition. In *Proceedings of Internationall Workshop on Frontiers in Handwriting Recognition*, 2000.

[41] J. Prager, E. Brown, A. Coden, and D. Radev. Question-Answering by Predictive Annotation. In *Proceedings of of SIGIR-2000, Athens, Greece*, 2000.

[42] I. Rigoutsos and A. Floratos. Combinatorial Pattern Discovery In Biological Sequences: The TEIRESIAS Algorithm. 14(1), 1998.

[43] A. Samuel. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research & Development*, (3):211–229, 1959.

[44] A. Senior, C. Neti, and B. Masion. On the Use of Visual Information for Improving Audio-Based Speaker Recognition. In *Proceedings of Audio-Visual Speech processing conference (AVSP99), Santa Cruz, CA*, August 1999.

[45] J. Shafer, R. Agrawal, and M. Mehta. SPRINT: A Scalable Parallel Classifier for Data Mining. In *Proceedings of the 22nd International Conference on Very Large Databases*, 1996.

[46] J. Shim, C. Dorai, and R. Bolle. Automatic Text Extraction from Video for Content-Based Annotation and Retrieval. In *Proceedings of the 14th International Conference on Pattern Recognition*, 1998.

[47] M.K. Singley, M. Singh, P. Fairweather, R. Farrell, S. Swerling, and J. Algebra. Supporting Teamwork and Managing Roles in a Collaborative Learning Environment. In *Proceedings of Computer-Supported Cooperative Work 2000 conference*, 2000.

[48] J.F. Sowa. *Principles of Semantic Networks*. Morgan Kaufmann, 1991.

[49] L. Stein and L. Morgenstern. Motivated Action Theory: A Formal Theory of Causal Reasoning. *Artificial Intelligence*, 71(1):1–42, 1994.

[50] J. Subrahmonia. Similarity Measures for Writer Clustering. In *Proceedings of Internationall Workshop on Frontiers in Handwriting Recognition*, 2000.

[51] G. Tesauro. TD-Gammon, A Self-Teaching Backgammon Program, Achieves Master-Level Play. *Neural Computation*, 6:215–219, 1994.

[52] R. Vilalta, C. Apte, and S.M. Weiss. Operational Data Analysis: Improved Predictions Using Multi-Computer Pattern Detection. In *Proceedings of the 11th IFIP/IEEE International Workshop on Distributed Systems: Operations & Management (DSOM 2000)*, 2000.

[53] H. Watanabe, K. Nagao, M.C. McCord, and A. Bernth. Improving Natural Language Processing by Linguistic Document Annotation. In *Proceedings of Coling 2000 Workshop on Semantic Annotation and Intelligent Content*, pages 20–27.

[54] S.M. Weiss and N. Indurkhya. Lightweight Rule Induction. In *Proceedings of the International Conference on Machine Learning (ICML 2000)*, 2000.

[55] T. Zhang and F. Oles. Text Categorization Based on Regularized Linear Classification Methods. 2000. to appear in Information Retrieval.