<h1>The Artificial Intelligence Laboratory </h1>

<h2>From the MIT President's Report July 1993-June 1994.</h2>

The primary goal of the Artificial Intelligence Laboratory is to understand how computers can be made to exhibit intelligence. Two corollary goals are to build intelligent systems and to understand certain aspects of biological intelligence. Current research in the Laboratory includes work on robotics, vision, natural language, enhanced reality, learning, reasoning and problem solving, model-based expert systems, engineering design, supercomputing, and basic theory. Major new applications themes include information transportation, access, and analysis, enhanced reality technology, and applications of artificial intelligence in medicine. Professor Patrick H. Winston works on the problem of learning from precedents. rofessor Marvin Minsky develops general theories of intelligence and knowledge representation. Professor Robert C. Berwick studies fundamental issues in natural language, including syntactic and semantic acquisition. Professor David A. McAllester works on knowledge representation and automated reasoning. Professor Lynn Andrea Stein works on integrated architectures for intelligence. Professor W. Eric L. Grimson, Professor Berthold K. P. Horn, and Professor Tomaso A. Poggio do research in computer vision. Professor Rodney A. Brooks, Professor Tomas Lozano-Perez, Professor Marc H. Raibert, Professor Warren P. Seering, and Dr. J. Kenneth Salisbury work on various aspects of Robotics. Professor Randall Davis and Dr. Howard E. Shrobe work on expert systems that use both functional and physical models. Professor Carl E. Hewitt studies distributed problem-solving and parallel computation. Dr. Thomas Knight and Professor William J. Dally work on new computer architectures. Professor Gerald J. Sussman and Professor Harold Abelson lead work aimed, in part, at creating sophisticated problem-solving partners for scientists and engineers studying complex dynamic systems.<p>

The Laboratory's 188 members include 18 faculty members, 24 academic staff, 27 research and support staff, and 119 graduate students active in research activities funded by the Advanced Research Projects Agency, System Development Foundation, Office of Naval Research, Air Force Office of Sponsored Research, National Science Foundation, Apple Computer, Bear Stearns Company, Digital Equipment Corporation, Fujitsu, Hitachi, Hughes Research Foundation, International Business Machines, Jet Propulsion Laboratory, Korean Atomic Energy Research Institute, Loral Systems Company, M & M Mars, Inc., Matsushita Electric, Mazda Motor Corporation, MCC Corporation, Mitre Corporation, Mitsubishi, NASA, Panasonic Technologies, Sandia National Laboratory, Siemens, Sperry Rand, Sumitomo Metal Industries, Systems Development Foundation, University Space Research Association (CESDIS).<p>

<h2>VISION</h2>

<h3>Object Isolation and Identification</h3>

Professor Grimson's group has focused primarily on methods for recognizing objects in cluttered, noisy, unstructured environments. Such systems have been incorporated into a hand-eye system, into a navigation system for autonomous vehicles, and into an inspection and process-control system for industrial parts. Recent efforts have focused on formal methods for evaluating alternative recognition methods, on grouping methods for preprocessing the input data into salient sets of features, on the role of visual attention in recognition, and on the development of efficient methods for indexing into large libraries of objects. These efforts have been integrated into a system that uses a movable eye-head to find objects hidden in a room by focusing attention on interesting points in the room, and then using grouping and recognition methods to identify such objects.<p>

One key new project is directed at using such methods to enhance the performance of brain surgeons by merging ordinary images of a patient with synthetic images produced using Magnetic Resonance Imaging. We call this enhanced reality. These visualization methods have been used in neurosurgery, and as part of a clinical study of Multiple Sclerosis.<p>

<h2>Motion Vision, Low-Level Integration, and Photogrammetry</h2>

Professor Horn and his students work on problems in motion vision. Currently, the extension of existing methods in the time direction is being explored. While one can get good motion information from just two image frames, distances to objects are determined only rather coarsely. Methods from computer graphics are used to predict the shape and position of an object at the next image frame time, based on the estimated shape and position and the estimated motion at the present time. Dramatic improvements in the accuracy of the reconstructed object shape are attained in this fashion, although after about ten frames the errors introduced by the prediction phase begin to balance out the improvements obtained from continuing the solution in time.<p>

Because the problem of dealing with the arbitrary motion (both translation and rotation) of arbitrary surface shapes presently appears intractable, they are exploring a number of special cases, such as the use of fixation. The equations relating image brightness changes and motion simplify when the camera system is servoed to maintain the image of a particular point in the environment stationary in the image plane. Methods have been developed for fixation that do not depend on tracking identifiable features in the environment. This is of importance in unstructured environments, where there may be texture on the surface being viewed, but no obvious features.<p>

Because recovery of information about the world from a single cue such as motion parallax, binocular stereo disparity, or shading in images tends to not be very robust, there is now a great deal of interest in integrating information from multiple cues. The intimate integration of early vision modules will be required for most practical applications of vision systems. Professor Horn's approach to the problem focuses on intimate integration at the lowest level of vision modules. In the simplest case, this means interlacing iterations of different schemes for recovering shape, or more formally, constructing a compound functional that contains penalty terms for mismatching information available from both cues being considered. Preliminary results in integrating motion vision and shape from shading, and in integrating binocular stereo and shape from shading show great promise. The new approach is particularly attractive because it suggests a systematic methodology for this integration, enabling new visual cues to be included easily.<p>

Professor Horn and his students are also looking into the object recognition through the computation of stable "invariants." While the idea of invariants is an old one in the field, it has in the past been applied to two-dimensional patterns, not three-dimensional objects. There has been some work recently elsewhere on invariants for recognition, but none that actually took into account the constraint provided by perspective projection. Ignoring this constraint leads to unstable methods of little interest in the real world of noisy data. New "invariants" are now being developed that exploit an accurate knowledge of the imaging model. Those require fewer features and lead to stable recognition schemes. Finally, work on a new special-purpose early-vision analog VLSI chip will be completed soon. This chip determines the "focus of expansion"--that point in the image towards which the camera appears to be moving. It does this without the need to detect and analyze image features. The result can be used to compute time to impact and possibly to recover shape information. The chip is expected to operate at 1000 frames per second. While it has only a 32 x 32 array of sensors and processing elements, it is expected to be able to recover the focus of expansion with sub-pixel resolution. Work is also starting on the next step, a chip that can deal with arbitrary combinations of translational and rotational motions, provided that the scene being viewed is approximately planar. This chip will be an order of magnitude more complex than the previous one and require considerable innovation before the circuitry can be fitted into the available space.<p>

<h2>Machine Learning and Networks</h2>

Professor Poggio's research focuses on the problem of learning a multivariate function from sparse data. Professor Poggio's group has developed a theoretical framework, based on regularization theory, that has roots in the classical theory of function approximation. The research effort at the Artificial Intelligence Lab and at the connected Center for Biological and Computational Learning continues to grow in several directions: theory and mathematical issues, algorithms for learning, use of prior information to augment the example set, exploiting virtual examples, applications, and neuroscience.<p>

In particular, Dr. Federico Girosi, Michael Jones and Professor Poggio have now succeeded in showing that regularization networks encompass a much broader range of approximation schemes, including many of the popular general additive models, some tensor product splines, and some neural networks. The result is significant because it provides a unified theoretical framework for a broad spectrum of neural network architectures and statistical techniques. Dr. Girosi and Professor Anzellotti have shown that several neural network architectures and function approximation techniques are closely related also from the point of view of the curse of dimensionality. Partha Niyogi and Girosi considered these classes of functions and showed how generalization error depends on the number of examples and on the number of parameters.<p>

Several applications have been demonstrated in visual object recognition, computer graphics, very low-bandwidth video conferencing, video e-mail, feature detection, target detection, visual inspection, visual database search, face detection, time series analysis, and pricing models in finance. For instance, David Beymer and Professor Poggio have developed and tested a view-based pose-invariant face recognition system using several views per person and running on the Connection Machine parallel computer on a data base of about 1000 face images. The recognition rate is about 98 percent on a data base of 68 people. They have also obtained encouraging preliminary results with a pose-invariant face recognition system which relies on only one example per person and the generation of virtual examples through transformations learned from face prototypes.<p>

Kay Kah Sung and Professor Poggio are developing a face detection system, capable of localizing faces in cluttered pictures. This is a specific example of a technology that is critical for object recognition, trainable visual inspection systems (for manufacturing and medical applications) and indexing n large data bases of images.<p>

Professor Poggio, Beymer, Sung, Stephen Lines and Amnon Shashua have developed and tested learning networks for the analysis and synthesis of images that could be used for trainable man-machine interfaces and very efficient computer graphics, respectively. Together the two systems may provide very-low bandwidth video-email and very-low bandwidth video-conferencing (in the order of a few bytes per frame).<p>

On the biological side, on-going physiological experiments with monkeys are providing exciting results that suggest that the mammalian visual system may use view-based recognition strategies similar to what the model of Poggio and coworkers predicts.<p>

The plan for the future includes further development of the mathematical theory and the development of additional applications. Applications to be explored include object recognition, image data base search and object detection, computer graphics, computer interfaces, and financial models. Biologically, work will focus on exploring how the brain may work and in particular how the visual system may recognize objects.<p>

<h2>NATURAL LANGUAGE UNDERSTANDING AND ACQUISITION</h2>

Professor Berwick and his colleagues have finished constructing the next generation of natural language parsers and translators, based on modular, universal linguistic theories. During the past year, a completely modular parser for Japanese was extended to cover a large subset of this language, and the system has now been extended to Dutch, German, Spanish, and Italian. By taking advantage of a parallel "dataflow" system organization, they have demonstrated nearly linear speedup in parsing, a first in this area.<p>

The results of this work were published in an MIT Press book, Cartesian Computation, and will appear in summary form this year in a Scientific American article. additionally, in collaboration with researchers from the Department of Linguistics and Philosophy, the universal parser has been extended to handle examples from Arabic and Hindi.<p>

Professor Berwick also led a special research group on developing multilingual tools for building dictionaries for languages as different as Hindi, Bengali, Greek, and Japanese. He has completed a system for automating this otherwise difficult task.<p>

Finally, Professor Berwick and his students have developed a new model for language acquisition based on the theory of dynamical systems. For the first time, this model can measure the actual number of examples needed to learn a particular language. The system can also model how languages change over time.<p>

<h2>ROBOTICS</h2>

<h2>Mobile Robots</h2>

Professor Brooks and his students have been working on embodied theories of intelligence. Previous work has concentrated on small mobile robots, and that work has been developed further. But, over this past year, Professor Brooks, in cooperation with Professor Stein, has embarked on a new project-to build a humanoid robot with human-like capabilities.<p>

During the past year two students completed work on ways to program groups of multiple mobile robots in an unstructured environment. Dr. Lynne Parker tackled the heterogeneous case and developed an architecture where the robots could cooperate without explicit negotiation. Instead, the robots all broadcast there actions continuously, and other robots, upon hearing this, may change their actions. Dr. Parker showed the system was robust in the face of great communication degradation, and would converge under a wide variety of conditions. She also added a component so that robots could learn the characteristics of other robots and improve their cooperative performance. Dr. Maja Mataric looked at groups of homogeneous robots, and was particularly interested in the case where no explicit communication was possible. The robots sense the presence of other robots like themselves and change their behavior appropriately, just as social insects interact. Dr. Mataric added a shaped learning capability to the robots and showed that it converged faster than other well known algorithms.<p>

Andrew Gavin completed work on new vision algorithms for a Mars rover, and tested his algorithms in a simulated Martian environment. Michael Binnard developed new techniques for fabricating and controlling a pneumatic six- legged robot. He showed that such a robot could plausibly operate on the Martian surface.<p>

Along with Professor Stein and her students, and students of other faculty in the laboratory, Professor Brooks and his students started work on Cog, a robot with humanoid form. The purpose of this project is to learn what constraints on the organization of human intelligence are imposed by having a body. The first year of the project was dedicated to building hardware and software so as to get a useful substrate upon which higher level work can be undertaken.<p>

<h2>Planning For Collision-Free, Compliant, and Grasping Motions</h2>

Professor Lozano-Perez and his associates work on computational problems in the areas of robot manipulation, computer vision, computer-aided design and computational biology. During the past year, they have made progress on a variety of fronts. They have demonstrated a system that uses off-line motion planning and simulation to enable a robot to use information from on-line visual sensing to choose appropriate task-achieving commands in the face of substantial sensing and modeling error. They have demonstrated a planning system, implemented on the Connection Machine, for selecting grasps for the Salisbury three-fingered robot hand. And together with Professor Grimson, they have begun field-testing a prototype system for the automatic registration of three-dimensional medical images, such as from tomography and magnetic resonance imaging, with three-dimensional range data from a patient in the operating room.<p>

Work is currently underway on developing autonomous robot agents for distributed simulation systems such as SIMNET, on the use of visual sensing for advanced human-computer interfaces, on combining off-line robot motion planning with on-line force feedback, and on applying geometric algorithms to the three-dimensional packing of amino acid residues in proteins.<p>

<h2>Legged Locomotion</h2>

Professor Raibert and members of the Leg Laboratory study legged locomotion in robots, animals, and computer animation. During the past year one of Raibert's students discovered that the somersaulting motion of a human-like object can be stabilized passively with the right kind of springy arms. Two other students have simulated the physics, appearance, and walking behavior of a giant cockroach and a robot-like cockroach.<p>

<h2>Robot and Human Arms and Hands</h2>

Dr. Salisbury's group has focused this year on three areas: sensor guided grasping, study of human and robot hands, and the development of haptic interfaces.<p>

Dr. Salisbury's study of human and robot hands, in collaboration with Dr. Srinivasan at the Research Laboratory for Electronics, has focused on the development of touch perception algorithms to enable robots to deduce contact conditions, including contact, texture, constraint, and motion. The algorithms use simple force-sensing fingertips and force time-history information. Dr. Salisbury's group is also developing precision force and motion manipulators for dual use in robot palpation experiments and human psychophysical experiments.<p>

Significant progress also has been made in hardware and program development for our new PHAN ToM haptic interface. This device exerts precisely controlled force vectors on user's fingertips or a stylus and is used to present touch (or "haptic") information to the user. This permits users to mechanically interact with virtual objects, permitting perception of properties including touch, shape, texture and motion. A larger "tool-handle" version of the device has been build and is undergoing testing. These devices are being used in our laboratory and in RLE as components of the Virtual Environment Training Technology program.<p>

<h2>REASONING TOOLS FOR ENGINEERS AND SCIENTISTS</h2>

<h2>Mixed Symbolic and Numerical Computation</h2>

The research of the MIT Project for Mathematics and Computation (Project MaC), under the direction of Professors Abelson and Sussman, is working to demonstrate breakthrough applications that exploit new computer representations and reasoning mechanisms that they have developed. These mechanisms enable intelligent systems to autonomously design, monitor, and understand complex physical systems through appropriate mixtures of numerical computing, symbolic computing, and knowledge-based methods. They call this mixed approach intelligent simulation.<p>

Systems incorporating intelligent simulation can automatically prepare numerical experiments from high-level domain descriptions. They automatically select and configure appropriate numerical methods. They actively monitor numerical and physical experiments. They automatically analyze the results of such experiments, using domain knowledge to interpret the numerical results, and they report these results to their human users in high-level qualitative terms. In favorable cases intelligent simulation pro grams can automatically configure special-purpose hardware for efficient execution of computationally demanding numerical experiments.<p>

The group has demonstrated the basic capabilities of intelligent simulation systems. They have implemented computer programs that interpret numerical simulations of nonlinear systems, automatically producing summary descriptions similar to those in the published literature.<p>

They have shown that programs can harness techniques from computer vision to "look at" simulation data and isolate regions of interesting behavior, and they have demonstrated this automatic analysis by discovering new results of current interest in theoretical hydrodynamics.<p>

During the past year, Dr. Andrew Berlin developed and demonstrated novel techniques for using active control to wiggle a structural element in such a way that buckling is prevented. He has performed analysis, simulation, and physical experiments to demonstrate his techniques. These techniques promise to lead to intelligent physical structures with unprecedented strength-to-weight ratios.<p>

The buckling of compressively-loaded members is one of the most important factors limiting the overall strength and stability of a structure. Dr. Berlin has constructed a small-scale model railroad-style truss bridge that contains compressive members that actively resist buckling through the use of piezo electric actuators. He has also constructed a prototype actively controlled column in which the control forces are applied by tendons, as well as a composite steel column that incorporates piezo-ceramic actuators. Active control of buckling allows this composite column to support 5.6 times more load than would otherwise be possible.<p>

<h2>Model Based Reasoning Systems</h2>

Professor Davis, Dr. Shrobe, and their associates are building knowledge-based systems that use models of structure, function, and causality to perform a wide range of problem solving and reasoning tasks. The systems they have built can reason about how a device works and how it fails in a manner similar to an experienced engineer. This is an important advance in the art of knowledge-based systems construction, because it provides the system with a more fundamental understanding of the device than is possible using traditional approaches.<p>

Where previous work dealt with digital electronic systems, more recent work is focused on understanding how things work in domains that include simple mechanical devices like four-bar linkages and mechanistic explanations of biological phenomena. Examples of understanding include the ability to produce descriptions of device behavior from a description of their structure, the ability to predict behavior under unusual circumstances, and the ability to redesign to fit those new circumstances.<p>

This work is based in part on the belief that the next major innovation in computer-aided design will be the construction of tools that understand (and can be told) how devices work and that can use this knowledge to support intelligent design modification. In order to be successful, such systems must already know a great deal of the basics of its design domain. A mechanical engineer takes it as given that a human colleague will understand terms such as "Scotch Yoke," "Four Bar linkage," and "Trip Mechanism," and would never employ a human assistant who did not know the meaning of these and hundreds of other basic terms, their common usages, and constraints on their application. Yet engineers endure such ignorance in their design aids daily.<p>

Making automated design tools powerful and easy to use will require constructing a very large knowledge base of engineering "know how." Such a system will contain roughly an order of magnitude more knowledge than existing commercial knowledge based systems and will lead to a new level of flexibility and power.<p>

Professor Davis has also been leading the Intelligent Information Infrastructure project, which is concerned with the next generation of ideas and software to support the National Information Infrastructure. Their basic assumption is that the National Information Infrastructure should have intelligence embedded into it, allowing it to understand the information it is carrying and enabling it to provide the foundation for new ways to gather, organize, and transmit knowledge, as well as new ways to operate organizations to take advantage of new knowledge structures.<p>

The members of the project have built a variety of systems, including the publication/distribution system used by the White House Office of Media affairs, in use routinely since January 20, 1993 to distribute

OMA publications nationally and internationally, and an on-line surveying system used to determine the size and character of the audience receiving the documents. They have also developed and used the START system to provide a natural-language based information resource.<p>

They have also worked with Senator Kennedy's office to establish email and World-Wide-Web connectivity between the Senator and his constituents. They drafted the early versions of the Senator's homepage and continue to assist in updating.<p>

<h2>Engineering Problem Solving and Design</h2>

During the past year, Professor Seering's group has made advances in understanding the dynamic response characteristics of robot arms. The focus of the group's work has been on system identification and on exploring control strategies for improving robot performance. Efforts are ongoing on finding ways of minimizing undesirable vibrations in dynamic systems. In collaboration with NASA the group has been working to improve the response of flexible space structures. An evaluation of the group's ideas will be run on the Space Shuttle on a flight tentatively scheduled for January of 1995. <p>

Professor Seering's students have also been looking for ways to use computers to help mechanical designers capture design decisions so that information about those decisions can be retrieved efficiently and at the desired level of detail by a designer. This work is being conducted in collaboration with a team of researchers at the General Motors Technical Center and a group of GM design engineers.<p>

<h2>HARDWARE AND SOFTWARE ARCHITECTURES</h2>

<h2>Symbolic Parallel Architectures</h2>

The Symbolic Parallel Architecture group, under the direction of Dr. Knight, has been developing a uniform, large scale, parallel symbolic supercomputer called Transit. Unlike most parallel machines, this architecture has been explicitly designed to support a wide range of parallel programming models with excellent performance. The key realization is the critical importance of low latency in the processor to-processor communications path. This low latency communications is used as a substrate for coherent caches and processor-to-processor message passing. The implementation of Transit is being done in three phases: construction of the routing network, coherent cache implementation, and finally processor design. The routing network is currently under detailed design and simulation. Its construction involves novel three-dimensional packaging and cooling technology, novel VLSI techniques for chip-to-chip communications, and a very simple, high speed routing component. The initial prototype is expected to yield a remote memory access latency of about 300ns and a per-port peak bandwidth of 800 megabaud. The aggregate switch bandwidth approaches a terabaud.<p>

Initial design of a massively parallel SIMD computer for early vision applications has also begun. The architecture maximizes the number of bits transformed each cycle by allocating a very simple processing element to each bit of a data item. Simulation of low-level arithmetic primitives has been completed, and preliminary VLSI layout issues are now being addressed. A system composed of 144 chips is expected to deliver between approximately 100 billion 16-bit arithmetic operations per second.<p>

<h2>Concurrent VLSI Architecture</h2>

The Concurrent VLSI Architecture Group under the direction of Professor Dally has been developing techniques for applying VLSI technology to solve information processing problems. The group has been developing the J-Machine, a fine-grain concurrent computer that offers supercomputer performance and tests a number of new concepts in interconnection networks, addressing mechanisms, processor architecture, and concurrent software systems. During the past year, the group completed the design of a single node of the J-Machine, known as the Message Driven Processor. The MDP chip was fabricated by our industrial partner, and first samples of the chip worked correctly, at greater than predicted performance. They have built an initial multi-node system, and expect to complete a 1024 node prototype J-Machine by year's end. They are continuing to develop system software, languages, applications, and high-speed peripherals for the machine. They have written a distributed operating system and compilers for the Concurrent Smalltalk and Concurrent Aggregates programming languages. As one of the machine's initial applications, they are developing a high performance, reliable transaction processing system.<p>

<h2>Message-Passing Semantics</h2>

The Message-Passing Semantics group, under the guidance of Professor Hewitt, has been developing the foundations for Open Systems that perform robustly in changing environments. An Open System is one that is always subject to unanticipated communications from outside and whose operations are subject to indeterminate results. Robustness means the ability to keep commitments in the face of conflict and indeterminacy, which are ubiquitous in Open Systems. Robust computer systems are needed to meet the challenge of Open Systems so as to gain from the advantages of openness while meeting the requirements that are imposed by openness. Open Systems undergo continual change: some change coming from within, through communication among internal parties, some from without through interaction with the environment.<p>

The primitives of ultraconcurrent systems are called ACTORs. These can be organized into systems of ORGs (Organizations of Restricted Generality). The Actor model provides a scientific and technological basis for Open Systems because it supports dynamic reconfigurability, compositionality, and extensibility. The ORG model provides a scientific and technological basis for organizational systems because it supports teamwork, management, liaison, and organizational representation. The group's research focuses on theoretical, architectural and on linguistic aspects of organizational systems composed of humans and telecomputer systems.<p>

<h2>BASIC THEORY</h2>

<h2>Society of Mind</h2>

Professor Minsky has continued to develop the theory of human thinking and learning called the "Society of Mind." This theory explores how phenomena of mind emerge from the interaction of many disparate agencies, each mindless by itself. For example, one aspect of the theory explains the combination