

The Role of Intelligent Systems in the National Information Infrastructure

Edited by
Daniel S. Weld

Associate Editors:
Joe Marks, and Daniel G. Bobrow

■ This report stems from a workshop that was organized by the American Association for Artificial Intelligence (AAAI) and cosponsored by the Information Technology and Organizations Program of the National Science Foundation. The purpose of the workshop was twofold: first, to increase awareness among the artificial intelligence (AI) community of opportunities presented by the National Information Infrastructure (NII) activities, in particular, the Information Infrastructure and Technology Applications (IITA) component of the High Performance Computing and Communications Program; and second, to identify key contributions of research in AI to the NII and IITA.

The workshop included a presentation by NSF of IITA program goals and a brief discussion of a report aimed at identifying important AI research thrusts that could support the development of twenty-first century computing systems. That report, as well as the full set of initial suggestions for it from AAAI fellows and officers, was circulated to attendees prior to the workshop. Workshop attendees identified specific contributions that AI research could make in the next decade to the technology base needed for NII/IITA and the major research challenges that had to be met. This report records the results of these discussions. It is organized to follow the IITA program description produced by the HPCCIT IITA Task Group.

The time from workshop presentation to written report was long, arduous, and fraught with debate and difficult decisions. We thank the editors for their efforts in producing this report. Special thanks to Dan Weld for his dedication and perseverance; his skill in unifying the varied contributions was critical to this report.

—Barbara Grosz, President, AAAI

1. Introduction

The National Information Infrastructure (NII) will have a profound effect on the education, lifestyle, and well-being of Americans from every corner of society. The infrastructure will transport critical information and software to every home, open educational and training opportunities to remote communities, and accelerate commerce by reducing the time to develop new products and increasing the efficiency of markets. Because electronic delivery is orders of magnitude faster than traditional transport, the NII will create new markets in information services and will spur development of strategic applications in areas such as health care, environmental monitoring, and advanced manufacturing.

The NII is expected to grow to include a million networked information repositories that support fast access to medical images, interactive product simulations, digital libraries, and multimedia educational materials. Current trends in semiconductor density, processor speed, and network bandwidth suggest that the infrastructure will be thousands of times larger than existing systems such as the Internet; the array of services supported by the NII will be unimaginably vast.

But, who will be able to use the NII and take advantage of the opportunities it offers? Most people have no formal training in computers. They have little interest in the computer itself; rather, they want to find some-

thing or someone or accomplish some task. No matter how fast the computers of the future become, the NII will not achieve its full potential unless the infrastructure is flexible and easy to use. Instead of forcing a user to remember how and where to access information, NII computers need to understand a user's task, guide him or her to the correct place, and show the user what he or she wants. Instead of requiring that users "surf" the net to find new sites, the NII should automatically track users' interests and inform them of relevant possibilities. Instead of being a source for data, the NII should be a source for services and solutions. Today's computer systems are rigid and complex; they require users to learn arcane languages rather than adapting to the way people naturally communicate and work. To prevent critical limitations in the NII, we must understand how people reason about the world and how they interact with each other; and we must engineer our machines to do the same.

Artificial intelligence (AI) uses the theoretical and experimental tools of computer science to study the phenomena of intelligent behavior and to construct intelligent systems. The field is diverse and multifaceted—it addresses one of the most profound scientific problems, and also develops practical technology. AI research has also produced an extensive body of principles, representations, algorithms, and spin-off technologies. Successful applications range from the DART system, which was used in deployment planning for Desert Shield, to broadly adopted symbolic math packages, such as Mathematica®, to thousands of fielded expert systems³. Incorporating AI technology into the next generation of computers forming the NII can help ensure that the nation's information infrastructure is both flexible and easy to use.

In the next subsection, we provide an overview of the technical challenges confronting the NII. Then, we outline specific research areas with potentially large payback. Sections 2 and 3 elaborate each of these points and link the challenges to the research areas.

Executive Summary

The National Information Infrastructure (NII) will have profound effects on the lives of every citizen. It promises to deliver to people in their homes and offices a vast array of information in many forms, changing the ways in which business is conducted, offering new educational opportunities, bringing geographically dispersed library resources and entertainment materials to everyone's doorstep. It will connect people to people, and help them with their jobs and tasks.

For the NII to be useful, however, people will need easy and efficient access to its resources. Today's computers are complex and difficult to use, even for experts. The NII will be orders of magnitude more complex than current systems; it could easily become a labyrinth of databases and services that is inconvenient for experts and inaccessible to many Americans.

The field of artificial intelligence (AI) can play a pivotal role in meeting major challenges of the NII. AI uses the theoretical and experimental tools of computer science to study the phenomena of intelligent behavior. The field not only addresses a profound scientific problem, but also develops practical technology for constructing intelligent systems. AI research has produced an extensive body of principles, representations, and algorithms. Successful AI applications range from custom-built expert systems to mass-produced software and consumer electronics. AI techniques can play a central role in the development of a useful and usable National Information Infrastructure (NII) because they offer the best alternative for addressing three key challenges.

First, AI technology can help make computers easier to use. It will support the development of computer interfaces that collaborate with users to meet their information needs. These interfaces will handle multiple modalities including natural language, gestures, graphics, and animation and will be able to employ whichever modality best suits a particular user request. The interfaces will operate as intelligent agents, allowing users to state what they want accomplished

and automatically determining the actions required to satisfy these needs and when to perform them. Over time, these intelligent agent systems will build a model of the user's needs and will adjust automatically to an individual's skills and pattern of usage.

Second, AI representations and techniques can support the development of a flexible infrastructure. To be useful, the NII must have intelligent indexing and provide convenient access to all forms of information. Doing so presents a significant challenge because the NII will contain information on a multitude of diverse subjects, and data represented in a wide variety of forms, including various natural languages, digital and video images, audio, geometric computer-aided design (CAD) models, mathematical equations, and database relations. Various areas of AI research and technology can help. Speech- and image-processing algorithms will allow systems to extract and identify multimedia content and index it with symbolic descriptions, thus enabling fast, flexible retrieval of answers to queries. Knowledge representation and reasoning methods will enable data translation services to convert information from one format to another, subject to semantic constraints. Work on agent architectures will provide the basis for constructing specialized software agents to act as subject-specific brokers, tracking the creation of new databases, noting updates to existing repositories, and answering queries in their targeted area.

Third, AI techniques can assist in the development of more powerful software tools and environments to support all stages of a project's life cycle: specification, design, adaptation, construction, evaluation, and maintenance. These software development tools and environments will be used in constructing advanced user interfaces and the complex systems needed for National Challenge applications as well as the software needed for the NII itself. Currently, construction of large software systems is costly; subsequent evolution and reuse is problematic. AI representations and algorithms can aid the con-

struction of rapid prototyping systems and enable the management of libraries containing reusable software modules and large knowledge bases. Advanced planning systems, combined with group-enabling software, will support increased efficiency of multiperson projects. The ability to populate synthetic environments with simulated people will enable virtual product testing and improve education and training.

The results of previous research and applications development place the field of AI in position to make enormous contributions to the NII and the National Challenge applications. Although the state-of-the-art offers a substantial body of methods, representations, and algorithms, the full realization of this promise requires a concerted attack on several fundamental scientific problems.

This report recommends several basic research initiatives in AI, each of which has high potential for large payback to the NII endeavor. Speech and image processing will contribute to improved user interfaces and enable automatic classification of multimedia content. Knowledge representation structures, plausible reasoning algorithms, and large-scale ontologies will enable NII systems to reason about user objectives and abilities and infer the databases and services of most interest. Machine-learning and planning methods will provide the basis for systems that relieve the user from the need to memorize details of database protocols or personally track changes to network services; they can also be used to construct systems that automatically adapt to human preferences. Research on software agent architecture will enable more sophisticated interfaces, software development aids, and simulation systems. Development of computational models of collaboration will enable multiple software agents to coordinate and thus furnish enhanced network services; they can also provide the basis for building human-computer interface systems that collaborate with people in using NII resources to solve problems and perform tasks.

1.1 Technical Challenges

Numerous obstacles block the development of both the NII and National Challenge applications in education, health care, advanced manufacturing, and electronic commerce.⁴ In this report, we focus on three of the most difficult challenges: ease of use, flexible infrastructure, and powerful development tools. We discuss these challenges briefly here; in Section 2 we elaborate and explain the potential role of intelligent software systems in meeting these challenges.

1.1.1 Ease of Use

Current computer systems are complex and difficult to use even for experts. The NII will be orders of magnitude more complex than the Internet, and could easily become a labyrinth of databases and services. For the NII to be accessible to all citizens, dramatic improvements must be made in the design of user interfaces (This point is elaborated in Subsection 2.1).

Today's interfaces require that users memorize cryptic commands, menu selection sequences, and button clicks; people are forced to adapt to the machine. In contrast, NII interfaces will need to be intelligent, adjusting automatically to a person's skills and pattern of usage. An intelligent interface to NII resources could help people find and do what they want, when they want, in a manner that is natural to them, and without their having to know or specify irrelevant details of NII structure.

A natural metaphor for such an interface is a software agent, an *intelligent agent* (i.e., an entity capable of autonomous goal-oriented behavior in some environment, often in the service of larger-scale goals external to itself) that acts as a personal assistant to the user. Users will want to communicate with their agents in familiar and flexible ways—by speaking English, drawing diagrams, or providing concrete examples. Agents should be goal oriented, allowing users to state what they want accomplished, then automatically determining how and when to achieve the goal. These agents should understand an expressive range of commands so that users can form questions or re-

quests without having to learn—or be limited by—an artificial query language. They should be *cooperative*, collaborating with the user to refine incorrect or incomplete requests. Furthermore, personal software agents should have the ability to be customized, automatically adapting to different users by following direct requests from users and learning from experience with them.

1.1.2 Flexible Infrastructure

Just as the nation's highway system would be an unnavigable maze without such services as maps, gas stations, and signposts, the NII will be unusable without a flexible system of support services. Because the majority of transactions will be interactions between two autonomous programs, an NII equivalent of maps and signposts must be designed to provide guidance to software agents as well as people. Context sensitivity is important: by accounting for a user's (or software agent's) objectives, NII services can provide guidance more like a chauffeur or a tour guide than a map.

Three factors conspire to confound the task of developing flexible NII support services: scale, scope, and heterogeneity. Together, the NII's information repositories will hold information on a truly vast scale. The scope too will be vast: the stored information will range across all subjects. In addition, data will be represented in an incredible variety of forms, including various human languages, digital and video images, audio, geometric computer-aided design (CAD) models, mathematical equations, and database relations.

A foundational suite of high-level information infrastructure services (Subsection 2.2) could provide significant assistance to NII applications, enabling them to handle the variability in these three dimensions and ensure full access to these data in a manner that supports interoperability. We envision at least three types of infrastructure services that could provide critical support for common problems: data and knowledge management services, integration and translation services, and knowledge discovery services.

Data and knowledge management services address two common NII needs: (1) finding information that is relevant to your task or goals; (2) finding the right audience for a piece of information you have produced. These services allow information consumers to quickly locate useful facts and software resources in a huge morass of heterogeneous, distributed data.

Data that are similar in content can vary greatly in form and in the operations that can be performed on them. Integration and translation services might convert information from one format to another subject to semantic constraints. For example, a financial translation service would not just perform the unit conversion from Japanese yen into U.S. dollars, but could convert from raw cost to total cost, including import duties, taxes, and fees.

Because the NII's information repositories will be huge and because many will evolve rapidly, it will be impossible for people, unassisted, to check the repositories for consistency. Instead, knowledge discovery services could track the creation of new databases and updates to existing repositories. These services could cross-index related topics to discover new correlations and produce summaries.

1.1.3 Powerful Development Tools

Today's tools and programming languages make the construction of software systems tedious and error prone. Current technology provides less support than the development of the NII and ambitious National Challenge applications require. Better support would be useful at all points of a project's life cycle: specification, design, adaptation, construction, evaluation, and maintenance. Many software development problems could be ameliorated by devising a set of powerful tools and environments (Subsection 2.3).

We envision at least three distinct kinds of tools to which AI techniques could contribute: (1) rapid prototyping systems that combine services for specifying and refining designs with modular libraries of previously developed software and

world knowledge; (2) intelligent project management aids that include software to promote collaboration and distributed decision making as well as next-generation project management software capable of checking resource utilization and assisting group leaders in replanning when unexpected conditions occur; (3) distributed simulation and synthetic environments to be used by applications for education, training, and computational prototyping of products.

1.2 Recommendations

A substantial body of AI research has addressed both the underlying nature of intelligence and the development of engineering algorithms necessary to reproduce rudimentary machine intelligence. This research has placed the field of AI in position to make enormous contributions to NII interfaces, flexible infrastructure, and development tools as well as to National Challenge applications. However, a concerted attack on several fundamental scientific problems is required to fully realize this promise. Here, we briefly present several key subareas within AI that we believe to be especially relevant to the development of a flexible and adaptive NII; in Section 3 we describe the state of the art of each of these AI subfields and suggest promising directions for research.

Research in knowledge representation (Subsection 3.1) seeks to discover expressive and efficient methods for representing information about all aspects of the world. Knowledge representation is important to the NII because almost every intelligent computational activity depends on it to some degree. Knowledge representation systems offer the benefits of object-oriented databases and the structuring capabilities of hypertext-based libraries; they also provide increased expressiveness and more powerful algorithms for information retrieval and update.

Machine learning methods (Subsection 3.2) extend statistical techniques in order to enable systems to identify a wide range of general trends from specific training data. These methods can be used to con-

struct interface systems that adapt to the needs of individual users, programs that discover important regularities in the content of distributed databases, and systems that automatically acquire models of the capability of new network services.

The field of planning (Subsection 3.3) develops algorithms that automatically construct and execute sequences of primitive commands in order to achieve high-level goals. By reasoning about formal models of the capabilities and content of network services and databases, AI planning systems can focus information-gathering activities in profitable directions. Because planning systems take a declarative goal specification as input, they can also help raise the level of user interfaces, allowing users to specify what they want done, then computing actions needed to achieve the goal and determining when these actions should be executed.

Work in plausible reasoning (Subsection 3.4) has leveraged statistical principles to devise principled encodings for many forms of uncertain information. Algorithms have been developed to support diagnostic reasoning, causal inference, and evaluation of the tradeoffs between plan cost and goal satisfaction. Plausible reasoning techniques are especially appropriate for National Challenge application areas such as health care, but are applicable to the information infrastructure as well. For example, intelligent help systems can use behavior traces to assemble probabilistic profiles of user goals, and personal assistants might assess tradeoffs between the user's conflicting objectives.

The study of agent architecture (Subsection 3.5) seeks to integrate specialized AI subfields to create *intelligent agents*, robust entities that are capable of autonomous, real-time behavior in an environment over an extended period of interaction. Agent architectures could provide the integration needed to support a variety of critical roles in the NII, including personal assistants; intelligent project coaches; and large-scale, distributed, group-trainers.

Research into multiagent coordination and collaboration (Subsection

3.6) has developed techniques for representing the capabilities of other agents and has specified the knowledge needed by agents to collaborate. Negotiation algorithms have been developed that allow two intelligent agents to determine areas of shared interest and compute agreements that increase the utility of all participants. This area is crucial to the NII because the sheer scope of the infrastructure will demand that much activity be performed by software agents, without detailed supervision by people. Techniques developed in this area will also play central roles in developing more collaborative and flexible systems for human-computer communication.

The goal of ontological development (Section 3.7) is to create explicit, formal, multipurpose catalogs of knowledge that can be used by intelligent systems. In contrast with knowledge representation research that focuses on the form of representation and methods for reasoning using those forms, research in ontological development focuses on content. An ontology for finance, for example, would provide computer-usable definitions of such concepts as money, banks, and compound interest. Creation of shared systems of vocabulary is crucial to the NII because ontologies provide the conceptualizations and basic knowledge required for communication and collaboration among different agents and between a person and his or her personal intelligent agent.

The fields of speech and language processing (Section 3.8) seek to create systems that communicate with people in natural languages such as written and spoken English. Applications to the NII are vast. Speech systems could revolutionize user interfaces, especially for small, mobile computers. Textual analysis could lead to superior indexing systems and improved information retrieval.

Research in image understanding and synthesis (Subsection 3.9) is leading to algorithms for analyzing photographs, diagrams, and video as well as techniques for the visual display of quantitative and structured information. NII applications for image un-

derstanding and synthesis will range from the extraction of semantic content for use in browsing and searching image data to intelligent compression schemes for storage and transmission to enhanced medical imaging to the generation of realistic (or schematic) artificial scenes from models extracted from world images.

2. The Role of Intelligent Systems

Intelligent interfaces, advanced infrastructure services, and powerful tools are necessary prerequisites to the construction of National Challenge applications in areas such as health care, electronic commerce or twenty-first century manufacturing. In this section, we describe three roles for intelligent software systems for the foundational substrate of the NII, and suggest the types of AI research necessary to transform the NII vision into reality.

In Subsection 2.1, we explain how intelligent user interfaces could act as assistants to both novice and expert users, helping them navigate the NII's labyrinth of databases and efficiently interact with advanced services. By responding to high-level requests in spoken language and other natural modalities, by communicating information both verbally and graphically, by automatically determining how and when to accomplish the goals of individual users, and by adapting to the skills and desires of those users, personal assistant agents will allow humans to benefit from information resources and facilities that might otherwise overwhelm them with their size, complexity, and rate of change.

In Subsection 2.2 we describe a set of general infrastructure services that could act as a foundation for construction of intelligent interfaces and development of the National Challenge applications. Examples include intelligent indices that help track resources, integration and translation services that convert between heterogeneous representations, softbots and brokers that act on behalf of information consumers and producers to create an efficient information economy.

In subsection 2.3 we define several software development tools and environments that could speed the construction of the advanced user interfaces and network-resident applications and services described previously. By providing intelligent support for specification and refinement tasks, rapid system prototyping will be possible; by exploiting work on collaborative planning and agent architecture, current work on computer-supported cooperative work and distributed simulation and training environments will be made more effective. The effective marriage of modern software-engineering methods with state-of-the-art AI technology will provide the means for constructing and strengthening the virtual organizations needed to develop and maintain new software and hardware resources for the NII.

2.1 Intelligent Interfaces

The resources and facilities provided by the NII will only be useful insofar as users have simple and effective ways of finding and using them. With the increasing number and kinds of services and resources available electronically (such as news wire stories; medical images; government information; electronic commerce and banking services; multimedia libraries and tutoring systems; music, film, and interactive entertainment; scientific articles; and online debates) potential users will be overwhelmed and frustrated unless access is both simple and effective. The conventional tools of computer-human interaction will not suffice for this next generation of applications; experienced users are already overwhelmed on today's fledgling Internet. The gap between current tools and the NII's human-computer communication demands leads to a crucial challenge: providing intelligent interfaces to resources so that people can use the NII without difficulty. From the perspective of the "NII as information superhighway," an intelligent agent is the competent chauffeur who knows every road and quietly performs routine errands. This software agent acts as a personal assistant. To be called "intelligent," it must satisfy several interrelated criteria:

Integrated: Users should not be forced to remember the details of particular databases or the wide and growing variety of services and utilities to use them effectively. Instead, the system should support an understandable, consistent interface that tunes itself to the task at hand.

Expressive: Users should be able to form arbitrary questions and requests easily, without being limited by restrictive menus or forced to learn artificial query languages. Intelligent interfaces should accept requests in whichever modality (e.g. speech, text, gestures) the user chooses.

Goal oriented: Users should be able to state what they want accomplished. The intelligent interface should determine how and when to achieve the goal, then perform the actions without supervision.

Cooperative: Instead of the passive-aggressive error messages that are currently given in response to incorrect or incomplete specifications, intelligent agents should collaborate with the user to build an acceptable request.

Customized: Personal assistant agents should adapt to different users, both by receiving direct requests from the user and by learning from experience.

These criteria—and their consequences—are explored further in Subsections 2.1.1 through 2.1.3. In some cases, however, the best interface will be one that gives the impression of directly manipulable, three-dimensional space; the contribution of AI to these virtual-reality interfaces are described in Subsection 2.1.4.

2.1.1 Integration and Expressivity

Two decades ago, window-based graphical interfaces and the direct-manipulation metaphor revolutionized human-computer interaction. However, few fundamental changes have occurred since then, and computers remain intimidating to the vast majority of the population. If the NII is to be both broadly accessible and flexible, people will need to interact with it in a natural manner, much like they do with one another. For example, users will want to access NII resources using a combination of speech and text (typed or handwrit-

ten) in their own natural language, and with hand and facial gestures. Furthermore, an interface should be able to present information in the manner most conducive to interpretation, be it text, graphics, animation, audio, or some coordinated combination of several modalities. Whereas today's application interfaces offer, at most, a help command or menu option, NII interfaces will increase acceptance by offering customized, intelligent help and training, especially for the nonexpert user. Development of such a flexible interface paradigm raises several challenges in the areas of machine perception and automatic explanation.

2.1.1.1 Machine Perception. Because people converse using speech, written language, gesture and facial expression, the ability to communicate seems effortless. If we want to ensure that user interactions with the NII are as natural, computers will require more advanced perceptual capabilities. As a result of research in the AI community, such capabilities are becoming technically feasible: given a controlled environment, existing computer vision algorithms (Subsection 3.9) can recognize eye and lip movements as well as hand gestures. Speech systems (Subsection 3.8) are currently capable of robust speaker-independent recognition for small vocabularies, and practical speaker-dependent recognition for vocabularies of ten thousand words or more; real-time natural language processing systems (Subsection 3.8) have been used in numerous database-query applications. However, technical problems still remain: many current technologies are brittle and thus break too easily for them to be considered fully mature and ready to use.

2.1.1.2 Automatic Explanation. Computers acquire, process, and generate data far more readily than they can present or explain it. If the NII were only to provide more ways for data to be produced and transferred, it would have limited success. To complement existing conventional abilities to store and move raw data, we need intelligent agents that are both linguistically and graphically articulate. If a query returns huge

amounts of data, the intelligent agent should be able to compute a salient summary and present it using whichever modality best suits interpretation; it should support and be able to choose from among a wide range of options, including chart graphics, natural-language text, volumetric visualizations, animation, music, or speech. Furthermore, current interface capabilities that provide formatted data must be supplemented by automatic explanation systems that consider the background, abilities, and interest of the requester.

The technical challenges here lie in computing appropriate summaries, synthesizing output in a given modality, and determining which modality (or combination) is appropriate for communicating a particular message or response. Current technology can generate grammatical text from certain knowledge-representation formats and can synthesize acceptable speech from unrestricted text (Subsection 3.8). Pioneering work in the automatic design of effective graphics holds much promise, but the field is still in its infancy (Subsection 3.9). Other important questions—how to select and combine appropriate interface modalities automatically, and how to model and exploit discourse structure in multimodal interfaces—have scarcely been considered. Research in several areas is needed to enable agents to tailor explanations to individual users (e.g. Subsections 3.2, 3.3, 3.5, 3.6).

2.1.2 Goal Orientation and Cooperation

Effective use of today's technology requires memorizing the peculiarities of many resources, databases, and network services. For example, the Internet already supports simple video on demand, but relatively few users know how to use it. The problem of comprehensible access will only worsen with the widespread deployment of digital libraries and commercial transactions; it could become an insurmountable barrier for new users. The problem can be addressed in part with the application of interface conventions and standards, but the fundamental difficulty

is that interaction with most applications is far too particular and detailed. What is needed is a technology that will raise the level and quality of discourse between people and machines.

Supporting a truly high level of discourse involves several challenges. Users should be able to phrase requests in terms of *what* they want accomplished and leave the problem of determining *how* to achieve that goal to the interface. For this to be possible, agents must be able to understand a wide range of goals, access thousands of NII databases and utilities, negotiate for desired resources owned by different entities with different pricing structures, and combine results obtained from diverse sources. AI planning techniques (Subsection 3.3) provide a solid basis for meeting this challenge; they enable a system to use a logical encoding of a user's goal and a library of *action schemata* that describe available information sources, databases, utilities, protocols, and software commands to build, interpret, and execute a plan that will accomplish the desired objective. Unlike standard programs and scripts that are committed to a rigid control flow determined a priori by a programmer, a planner automatically and dynamically synthesizes and executes plans to accomplish a user's goals. Recent work in collaborative planning even addresses the problem of how a user and computer might collaborate to formulate a shared plan when neither human nor computer alone knows how to achieve a desired goal (Subsection 3.6). In addition to raising the level of discourse, the planning approach avoids the problematic task of writing programs that anticipate all possible changes in system environment, network status, and error conditions.

2.1.3 Customization and Adaptivity

Many users like to customize the look and feel of their computer interface. Witness the proliferation of screen-saver applications, window backgrounds, and custom keyboards and mice. Why stop there? Users should be able to specify preferences about all aspects of system behavior, leaving it to the personal assistant agent to

handle conflicts (for example, the conflict between a stated desire to use inexpensive services and an urgent demand). In addition to providing enhanced customization abilities, an intelligent agent-oriented interface must be adaptive. It must adjust automatically to the needs and idiosyncrasies of individual users, and it must change as the user's experiences or requirements change.

Intelligent agents must also adapt to their environment. As network utilities become heavily loaded or raise prices, the agent should shift to alternative services. When new NII databases and facilities are introduced or upgraded, it should explore them (or consult a broker) to determine applicability.

Machine learning techniques (Subsection 3.2) are crucial for building adaptive interfaces. Learning programs could unobtrusively watch over a user's shoulder during his or her normal interaction with the computer and later could generalize from its observations to customize the software. Prototypes of such interfaces have already been implemented. For example, an intelligent scheduler learns preferred meeting times and locations, and a correspondence assistant learns from a user's behavior how best to prioritize email messages. Future applications include personalized news streams as well as shopping assistants that learn a user's tastes and price ranges during the course of home shopping, access online consumer reports, and suggest new items for purchase.

2.1.4 Virtual Reality, Telepresence, and Interface Immersion

In many cases, the best interface gives the impression of directly manipulable, three-dimensional, physical reality. For example, an advanced CAD system might provide an automobile designer with the sensation of walking around, climbing inside, or driving a new vehicle around the test track—before a prototype is ever built. Virtual environments could have application in education, training, and entertainment as well. For example, virtual Japanese shopkeepers in a virtual Tokyo could provide students with an opportunity for lan-

guage immersion without the expense of a trip to Japan. Telepresence might allow people to manipulate hazardous environments and, thus, safely perform tasks such as undersea exploration or nuclear reactor maintenance.

At present, virtual environments provide only a limited approximation of reality, and advances in both interface hardware and software are required for wide-scale use. Subsection 2.3.3 elaborates on the major problems in the development of virtual environments and the potential solutions to which AI can contribute. Populating a virtual environment with seemingly intelligent agents will require substantial advances in all areas of AI, especially the real-time issues of agent architecture (Subsection 3.5). Synthesizing realistic facial expressions that synchronize with spoken language and provide feedback about an agent's understanding during dialogue will demand progress in both language and image processing (Subsections 3.8 and 3.9). Advances in both knowledge representation and ontological development (Subsection 3.1 and 3.7) will enable constructing synthetic environments that require specification of large amounts of knowledge about all aspects of the world.

2.2 Information Infrastructure Services

Advanced infrastructure services are a necessary prerequisite to the development of intelligent NII interfaces (Subsection 2.1) and the construction of National Challenge applications such as health care and electronic commerce. Each of these applications involve network-resident services. Just as the nation's road system would be a confusing maze without such services as maps, gas stations, and signposts, the NII will be unusable without an advanced infrastructure. Because the NII will be dramatically larger and more complex than the nation's road system, the demand for sophisticated NII infrastructure services will be vast.

As a simple example, a monolithic map will not suffice for NII navigation, because users will want cus-

tomized directions that are sensitive to factors such as individual objectives and local network congestion. Instead of being forced to rely on general-purpose signposts, users will prefer customized maps and signs that emphasize information relevant to the user's objectives. Agent-oriented navigation tools could also be popular; customized tour guides could take into account an individual's special interests and actively search for the best routes. The challenge for infrastructure developers is to provide efficient, effective search methods that can navigate the network and transact queries only where appropriate, handling a vast variety of interfaces to resources, and interpreting and collating the results.

This section discusses three types of infrastructure services: (1) *Data and knowledge management services* that allow information consumers to quickly locate relevant facts and software resources from a huge morass of heterogeneous, distributed data; (2) *Integration and translation services* that convert information from one format to another subject to semantic constraints; (3) *Knowledge discovery services* that scan rapidly evolving databases in order to produce summaries, discover new correlations, and check consistency.

Before discussing the details of these services, we note that the vast majority of interactions between entities on the network will not be between people and people or between people and programs but between programs and programs. People will, of course, also operate in the network. Many of the functions discussed here (including search, information brokering, network guidance, resource market research and marketing) can and will on occasion be performed by people—as they currently are in the physical economy. However, because of the NII's potential size, complexity, and rate of change, intelligent software systems will initiate a large fraction of network activity. If these programs are to be useful, they must be both intelligent and knowledgeable. For example, people can use a freeway signpost that reads "I-95 New England" to get

to Boston because they know that I-95 is the name of a freeway, the sign means I-95 goes to New England from here, and Boston is in New England. Network-resident intelligent agents will need similar kinds of general knowledge to infer that a seismic-activity database might hold the answer to a query about earthquakes.

2.2.1 Data and Knowledge Management Services

To realize the NII's potential, two closely related problems must be solved. Information consumers need effective ways to locate relevant information and software resources in a huge, distributed sea of heterogeneous data. Conversely, publishers must disseminate new information and services to interested people and software agents. Two challenges—heterogeneity and scalability—make location and dissemination services difficult to provide.

2.2.1.1 Heterogeneous Data. As we discussed in the Introduction, the information distributed on the NII will be stored in a wide variety of forms, from video images, audio and byte-coded and scanned text in various languages to database relations and mathematical equations. Indexing this information will be difficult because there are so many ways to categorize each item. For example, a photograph of Bill Clinton standing in front of the White House with Al Gore is indeed a picture of Bill Clinton. However, it can also be categorized as a picture of the White House, as well as a picture of Al Gore, a picture of a president, and a picture of the residence of a head of state. Similarly, a speech by Bill Clinton could be indexed by any portion of its content, any aspect of the style in which it was delivered, or any aspect of the circumstances of its delivery.

Because it would be grossly inefficient to index photos and audio clips under all possible terms, NII databases will need to use another method to provide flexible access. The polynomial-time inference and classification schemes of knowledge representation (Subsection 3.1) provide the desired functions, but multiple taxonomies must be supported, and classification schemes must allow

evolution over time. The sheer quantity of available data will require that many of the indices be created autonomously, which, in turn, will require information retrieval and natural language parsing techniques (Subsection 3.8) as well as algorithms from computer vision (Subsection 3.9). Because many queries will be underspecified and return too many matches, the information infrastructure must support quality determination by evaluating completeness, consistency, and relevance. Plausible and probabilistic reasoning algorithms (Subsection 3.4) have already demonstrated their utility for representing medical information, and their application to educational and help systems is growing.

2.2.1.2 Scalability. The NII must support both a vast amount of data and a huge number of users with different interests and needs. Although a monolithic information index could combine sources such as telephone books, airline schedules, and encyclopedias, such a centralized scheme would likely suffer from *rigidity* (being unable to respond to rapid changes in the world) and represent a likely point of failure, especially in high-activity situations. A similar problem results from naive attempts to scale current technology for information dissemination—indiscriminate advertising causes both cognitive and network congestion. In addition to broadcasting, the NII must support “narrow casting.” The AI challenge is to identify the select set of people and agents who are likely to be interested in an announcement of a new service (or, symmetrically, who could provide a desired service).

Several factors will contribute to a solution. First, instead of relying on passive indices, NII data access could use brokers that actively scan for new relevant resources. Each broker would specialize on a different subject area and would monitor closely related brokers in addition to primary sources. The onus for advertising rests on the information provider; interested brokers do the rest. An information seeker contacts an appropriate local broker and its request is

passed along until it is fulfilled. A centralized index can be seen as the degenerate case of this distributed scheme, but the existence of multiple, competing brokers could provide faster response time, improved specificity, and better adaptation to change in primary sources.

Second, the indices could utilize semantic information rather than mere syntactic criteria. Brokers could reason about index terms, their relation, and their relevance. Simple hypertext versions of extant reference books would serve poorly, because they require a human for navigation. Codified semantic information, however, can be processed by automated agents (as well as a manual browser). Knowledge representation techniques (Subsection 3.1) could enable inheritance along multiple taxonomic dimensions during index formation and allow a broker to determine which other brokers might be relevant during query processing. For this approach to be feasible, however, substantial effort must be invested to leverage existing classification schemes (such as the Dewey decimal system and Chemical Abstracts) and develop comprehensive new ontologies and encodings of commonsense knowledge (Subsection 3.7). Only then will search engines be able to deduce connections between relevant information sources.

2.2.2 Integration and Translation Services

We already commented on the problems posed by heterogeneous data in the NII (Subsection 2.2.1.1). However, even data that are similar in content can vary greatly in form and in the operations that can be performed on them. For example, even within a relatively constrained domain such as health care, the languages of doctors, nurses, patients, and insurance agents differ dramatically. Although much data-format conversion and some level of application-system interoperability are achievable through the development of standards (e.g. the RTF data format, the COBRA and OLE interapplication wrappers), standards enforce the lowest common denominator. By the very nature of the consensus that creates them, stan-

dards will always lag behind the needs of NII users.

Instead of data translation, the NII infrastructure should support semantic translation. Instead of simply scaling between currencies (for example, Japanese yen to U.S. dollars) semantic translation would use world knowledge to convert between derived quantities (for example, from raw cost to total cost including applicable import duties, taxes, and fees). Because determining which duties must be paid requires reasoning about the type of merchandise, the scope of services provided, and the relevance of tax codes, AI techniques can provide significant assistance in meeting the challenges of integration and translation. Knowledge representation systems (especially those using modal and higher-order logics for reasoning about representations, Subsection 3.1) form a substrate into which relational and other database systems can be embedded and over which standardized ontologies (Subsection 3.7) can be defined. When the information to be converted is relatively standardized and well structured, representation transformation techniques from automated software development are appropriate (Subsection 3.1). When the information is less structured, machine-translation technology (Subsection 3.8) can be used. In many cases, applications will use negotiation techniques (Subsection 3.6) to select a common language and ontology in order to facilitate coordination and translation.

2.2.3 Knowledge Discovery Services

The rapid growth of data and information has already created both a need and an opportunity for extracting knowledge from databases and ensuring their consistency. Because the development of the NII will release a flood of new databases, these needs and opportunities will surge. The sheer quantity of available information will preclude manual tracking. Automated tools, however, could scan databases; check consistency; produce summaries; support logical inference and abduction; facilitate browsing, question answering, explanation and justification; and discover new connections between data that

were previously unconnected.

The problem of consistency checking illustrates why knowledge discovery is difficult in general. The process is inherently intractable—algorithms take time that is exponential in the size of the database. Furthermore, consistency is unattainable in a large multisourced system that is being updated continually. Heuristic techniques can help by focusing attention on the most likely sources of inconsistency and suggesting plausible alternatives for dealing with inconsistencies when they are detected.

To date, knowledge discovery applications have been developed for astronomy, biology, finance, insurance, marketing, medicine, and many other fields. Next-generation systems could benefit from advances in knowledge representation (Subsection 3.1), techniques for the estimation of the quality and reliability of information (Subsection 3.4), the ability to exploit previous knowledge (Subsection 3.7), unsupervised machine-learning algorithms (Subsection 3.2), machine perception (Subsections 3.8 and 3.9), and statistics.

2.3 System Development and Support Environments

System development and support environments provide the tools and environments to build advanced user interfaces and network-resident systems like those needed for National Challenge applications. The objective is a set of tools for making specification, design, adaptation, construction, and evaluation straightforward in spite of the complexity and diversity of the underlying architectures, languages, protocols, applications, and systems. In the following subsections, we discuss three such facilities: rapid prototyping systems, intelligent project-management aids, and synthetic environment testbeds.

2.3.1 Rapid System Prototyping

It is commonly acknowledged that software carefully designed to satisfy someone's needs (for example, the designer's) often fails to satisfy everyone else's. In many cases, determining what other people want means letting them try it out, but the sheer cost of constructing and re-

working complex software systems calls for using prototypes that give the feel of the nominal product but can be constructed quickly at low cost. That is, designers formulate tentative, possibly partial, specifications of a desired system and then repeatedly test and revise the formulation until they obtain a satisfactory specification for postprototype versions. Facilities for rapid prototyping of systems are an essential infrastructure for the NII. AI can contribute to two aspects of prototyping systems: facilities for specifying and refining systems, and facilities for sharing and reusing designs and software.

2.3.1.1 Specification and Refinement Support Services. The process of rapid prototyping demands tools for specifying, testing, and revising prototypes. Specification calls first for formal languages (such as logical specification languages) in which to formulate descriptions. Once a tentative specification is in hand, the behavior it entails should be vividly rendered to the designer. Subsequent testing and design validation processes demand efficient algorithms. However, specification and refinement environments must support more activities than simple construction of an executable design; they must also facilitate revision and modification of descriptions, because any specification of a massively complex system will evolve with experience. Furthermore, support environments could leverage specification encodings to semiautomatically generate documentation, tutorial, and online help systems for the software under construction.

AI techniques can contribute to each of these challenges. Specification complexity can be alleviated by encoding designs in terms of the high-level concepts defined in a repository of world knowledge (Subsection 4.7). Because specification descriptions will still be huge, specification languages will use knowledge representation techniques (Subsection 3.7) to support modularity, hierarchical structure, and multiple inheritance.

Testing and validation of the specification can be done in several

ways. Theorem proving and logical inference algorithms (Subsection 3.1) can detect ambiguities, incompleteness, and inconsistency in the specification. Decision-theoretic techniques (Subsection 3.4) could select economic tradeoffs among possible alternatives given encodings of the available components. Another approach to validation uses knowledge-based automatic programming techniques (Subsection 3.3) and modules already available in software libraries (Subsection 2.3.1.2) to mechanically construct a program meeting the specifications; by executing this prototype, a user could see whether the specification behaves as expected and whether the expected behavior is desirable. Although it is easy to generate random test cases, robust validation requires using knowledge of user aims and the environment (Subsection 3.7) to generate qualitatively different tests that exercise every dimension of the system under construction. The same knowledge can be used to generate help systems.

Finally, system support environments can support evolutionary design by acting as an intelligent project coach (Subsection 3.5) that records (Subsection 3.1) and explains (Subsection 3.8) choices behind design decisions. When the system design is a collaborative effort, the project coach could track the responsibilities of team members (Subsection 3.6).

2.3.1.2 Software and Knowledge Library Support Services. Reuse of well-designed tools can provide one of the greatest economies available in any activity but especially in the labor-intensive practice of software engineering. Accordingly, the specification and refinement support environments described in Subsection 2.3.1.1 require access to libraries of software modules, module specifications, and formalized knowledge about these modules, users, applications, and the world.

Libraries house materials, but they also offer an internal structure that facilitates finding and using entries. In the case of software libraries, software modules in the collection must be indexed by their specification and

accompanied by a recording of their design rationale (Subsection 3.7). Additional encodings can facilitate modification of existing components for new purposes, automatic translation and optimization of modules, and safe replacement of parts of existing systems with independently developed improvements of the parts (Subsection 3.3). Providing intelligent access to these library functions will require all of the information infrastructure services, such as indexing and translation, described in Subsection 2.2.

2.3.2 Intelligent Project Management Aids

Constructing, maintaining, and extending the National Challenge applications and other complex systems poses many hard problems apart from the issues of prototyping, specification, and reuse just discussed. Even with these technologies at hand, there must be some way of managing all the people, systems, and agencies involved. Although many project management systems are currently available, the enormous scope of the NII project means moving beyond the current state of the art (for example, PERT charts or MacProject). Managing large-scale projects requires facilities for coordinating independent activities with groupware and managing the project plans themselves.

2.3.2.1 Collaboration and Group Software. To provide supportive environments for collaboration and group-cooperative work, software systems must provide participants in the collaboration with facilities for information sharing, virtual collocation, and task coordination. For example, coordination mechanisms could facilitate group discussions (possibly distributed in time, space, and participant background). Because organizations (and virtual communities) will likely be composed of a large and diverse collection of individuals, tools could inform users of recommended policies, procedures, and processes as well as facilitate the evolution of these guidelines and agreements.

Building cooperation-supporting environments requires developing

tools that model processes and plans, coordinate projects, and manage workflow constraints (Subsection 3.3). Such tools could prove most adaptable if organized around declarative models (Subsection 3.1) of the social organizations and entities involved in the activity, the relationships among them, and the constraints imposed on them by the nature of the activity. Cooperative efforts require mechanisms for managing communications to make participants more productive, and as in the information infrastructure generally (Subsection 2.2), this approach calls for developing software agents to filter the information posted to groups, broker information that moves from one team member to another (Subsection 3.6), and scour various databases in response to a user query.

In addition, collaboration-supporting environments must supply means, such as notification facilities or requirement utilities, to connect people and tools available on the network into a collaborative environment. For example, to support continuing conferences among participants with different interface capabilities, an environment must be able to translate "utterances" between different modalities (Subsection 3.9) and generate summaries to quickly bring offline participants up to date (Subsection 3.8).

2.3.2.2 Problem Solving and System Design Environments. Although collaboration and group software provides the basis for effective communication of teams or larger groups working together, large projects call for additional help in modeling the problem or task being addressed by the group. Existing management aids provide some help in this direction but do not offer much assistance in representing knowledge about plans and designs or provide mechanisms for reasoning about plans and designs in flexible ways. As the National Challenge application of crisis action planning illustrates, substantial knowledge about the likely consequences of actions and the utilities and intentions of the multiple actors involved is required to exploit inter-

actions and synergies between complex subprojects while keeping options open to maintain flexibility. Operations research techniques suffice for problems with simple utility characteristics, but for large collaborative projects, techniques need to be augmented with fast algorithms for managing plans that explicitly account for uncertainty, incomplete information, interactions, and tradeoffs (Subsections 3.1, 3.4, and 3.3).

2.3.3 Distributed Simulation and Synthetic Environments

To evaluate a prototype software system (Subsection 2.3.1), a designer must be able to simulate at least part of the environment in which the system is to operate. However, simulations of specialized environments are useful far beyond the issue of software design; other important applications include management of complex systems (such as air traffic control, telephone switching networks, power networks, distributed sensors, resource allocation in factories, remotely piloted vehicles), training (such as education, flight simulators, virtual surgery), planning and optimization, remote diagnosis, sensing, and cooperative interaction. Such synthetic worlds could integrate real, as well as virtual objects; combine visual, as well as computational, descriptions; and support human users as well as intelligent software agents.

Exploiting the potential for synthetic environments means facing some difficult problems. The first problem is constructing such environments, but because synthetic environments are just complex systems themselves, the techniques discussed in Subsection 2.3.1 can also be used to apply intelligence to ease and speed their construction as well. The second problem is populating the environment with simulated people or creatures. Robotics path planning and dynamic control algorithms (Subsection 3.9) can be used to generate realistic three-dimensional movements. AI planning algorithms (Subsection 3.3) and agent architectures (Subsection 3.5) could generate realistic behaviors. Environments that must contain life-like characters can successfully em-

ploy characters that act human even when these characters are considerably simpler than characters acting intelligently (Subsection 3.5). The third problem is that synthetic environments often reflect parts—in some cases, very large parts—of the world; so, the specification of the system must include large amounts of knowledge about the world. (Subsections 3.1 and 3.7).

3. Research Thrust Areas

Research on the underlying nature of intelligence and the development of practical algorithms necessary to reproduce rudimentary machine intelligence leaves the field of AI strategically situated to contribute to the design and construction of the NII. However, the full realization of this promise requires a concerted attack on a variety of fundamental scientific problems. This report recommends support of AI research in eight key areas, each of which has substantial promise for high payback to the NII effort.

In this section, we describe these key subareas:

- Knowledge representation
- Learning and adaptation
- Reasoning about plans, programs, and actions
- Plausible reasoning
- Agent architecture
- Multiagent coordination and collaboration
- Ontological development
- Speech and language processing
- Image understanding and synthesis.

This list is not a comprehensive enumeration of every interesting AI research topic. Rather, it elaborates areas that clearly and obviously emerged as important to the discussion within Section 2. Each is relevant to the development of a flexible and adaptive NII.

3.1 Knowledge Representation

Work on knowledge representation seeks to discover expressive, convenient, efficient, and appropriate methods for representing informa-

tion about all aspects of the world. *Expressive* here means being capable of capturing both general and specific information in broad and narrow domains; it also implies the ability to express weak or incomplete statements (for example, whoever is president of the company will chair the board of trustees) as easily as strong and concrete statements (for example Charles Diamond will chair the board of trustees). *Convenient* means permitting acquisition and reporting of the information in terms close to those used by either lay persons or experts. *Efficient* means supporting rapid extraction of common and important conclusions from the information. *Appropriate* means translating between component representations when the representations that maximize expressiveness, convenience, or efficiency differ. Unfortunately, there are usually tradeoffs between these properties; no known knowledge representation method scores well along all dimensions.

3.1.1 Relevance to the NII

Knowledge representation problems are important because almost every intelligent computational activity depends on solving them to some degree. Most information currently stored on the Internet uses one of two degenerate knowledge representation methods: databases or natural language text. We use the word *degenerate* because these representations are extreme points on the expressiveness-efficiency spectrum. When information can be encoded in a relational database, one can quickly answer any query expressible in a language such as SQL, but only restricted and concrete bodies of knowledge can be encoded in a database. Natural language text, however, is expressive enough to encode much of human knowledge, but no one has yet efficiently mechanized inference over unrestricted natural language text. No foolproof algorithms exist for answering questions or extracting conclusions from natural language documents. Because relational database and natural language representations are insufficient, new knowledge representation methods are required to achieve NII objectives,

such as accurate location of relevant information, narrow casting, semantic translation, and reusable software libraries.

3.1.2 State of the Art

Knowledge representation formalisms have been applied successfully to a variety of commercial and government applications, such as configuration, scheduling, customer service support, financial management, and software information systems. In each case, the knowledge representation methods offered considerably more flexibility than relational database systems. In general, knowledge representation systems offer syntax and structure much closer to natural languages but provide the semantic precision and inferential capability of logical languages. Many years of theoretical and experimental work have refined the core of these systems into *description logics*, which structure knowledge into modular, multiply connected taxonomic hierarchies of concepts, abstractions, and approximations. These logics offer the benefits of object-oriented databases and the structuring capabilities of hypertext-based libraries but go far beyond them in their expressiveness and in the algorithms available for retrieving information about entries and revising the hierarchies as information is added or updated. Guaranteed polynomial-time retrieval and inference algorithms have been developed for useful classes of description languages; still richer languages admit algorithms that are usually fast in practice.

These techniques for structuring, using, and transforming representations form a solid basis for knowledge representation systems, but they do not address many problems of how to express specific types of knowledge. Thus, numerous researchers have searched for suitable representations for commonsense information, such as quantities, time, physics, uncertainty, and knowledge representations themselves. These efforts have resulted in several well-understood libraries of techniques, for example, temporal reasoning. Because these issues prove important in virtually all parts of AI, progress on them offers

great advantage to the whole field.

3.1.3 Research Opportunities

Several knowledge representation research directions have the potential for exceptional payback for NII infrastructure and applications. The integration of description languages with object-oriented and relational databases could help provide value-added services on top of conventional data management platforms, examples of such services include more semantically oriented queries and knowledge discovery. The improvement of specialized languages for temporal, probabilistic, and non-monotonic languages could provide support for natural language processing and information retrieval. The development of standard languages for encoding knowledge of scientific fields and World-Wide Web hypertext libraries could facilitate support of semantically rich queries. The elaboration of metalanguages for describing knowledge representation systems (for example, their accuracy, relevance, efficiency, and completeness) could expedite automatic translation and interoperability.

3.2 Learning and Adaptation

Machine learning addresses two interrelated problems: the development of software that improves automatically through experience; and the extraction of rules from a large volume of specific data. Machine learning is of growing importance because of the rapidly increasing quantities of diverse data on the NII and the expanding need for software that can automatically adapt to new or changing users and runtime environments. The central technical problem in machine learning is developing methods to automatically form general hypotheses from specific training examples.

3.2.1 Relevance to the NII

Machine-learning methods offer new capabilities for the NII that are unavailable using current software technology. Machine-learning algorithms identify general trends from specific training data, offering the promise of programs that examine gigabytes of network-accessible data to extract trends that would otherwise go unno-

ticed by people. Machine learning also offers approaches to automatically modeling the NII itself by learning probabilistic regularities in server loads, security breaches, correlations among user accesses to data repositories, and the identity of services that typically are appropriate for recurring information needs.

The usefulness of *data mining* (the extraction of general regularities from online data) may be illustrated by the problem of learning which medical treatments are most effective in particular situations or which land-zoning policies produce the best outcomes. Current learning methods are able to find regularities provided large data sets of single-media data (as opposed to mixtures of images, logical descriptions, text, and sound). New learning methods that address multimedia data and generalize more accurately could have a significant impact on our ability to use the ever-growing amount of data that will be available on the NII.

As a second application for machine learning, consider the problem each user faces in locating information in the flood of data that will be available on the NII. Machine-learning techniques could lead to electronic news readers that learn the interests of each user by observing what they read, then use this knowledge to automatically search thousands of news sources to recommend the ten most interesting articles. Similar applications include building intelligent agents that provide current awareness services, alerting users to new web pages of special interest, or providing "What's New" services for digital libraries. Although information retrieval provides a baseline capability (such as keyword search on large stores of text), more accurate learning methods are needed.

3.2.2 State of the Art

Practical methods for learning from large volumes of single-media data have been demonstrated in a number of areas. For example, methods such as decision tree learning, neural network learning, genetic algorithms, and Bayesian methods have been applied to data-mining problems such as assigning credit ratings based on

bank records, recognizing human faces, and predicting medical treatment outcomes based on medical symptoms. New approaches have recently been developed, such as inductive logic programming, which enables learning more expressive hypotheses than earlier learning methods. Significant progress has occurred recently in developing a theory of machine learning. For example, there is now a quantitative understanding of how the error in learned hypotheses depends on the amount of training data provided and the complexity of the hypotheses considered by the learner. The field is moving forward rapidly, pushed by recent technical advances and the growing need for this technology.

3.2.3 Research Opportunities

Current learning methods already provide positive value, but basic research in new learning algorithms is likely to have significant payoff. Research is needed to develop methods that generalize more accurately from limited data, further develop the underlying theory of machine learning, and understand how to employ user-provided conjectures and background knowledge when analyzing available training data.

In addition, the opportunities for machine-learning applications in the NII suggest specific research directions. Methods for learning over multimedia data will be increasingly important. New methods will be needed for combining information from multiple databases. Given that much information on the NII will be in the form of text, one ripe area for basic research involves combining natural language-processing techniques with machine-learning methods. This area, which has been overlooked by researchers to a surprising degree in the past, seems especially important for the NII. A recognized but unfulfilled promise of machine learning is to aid the continual maintenance of large knowledge bases for knowledge-based assistant programs. Many systems are currently deployed throughout every sector of society; machine learning can help reduce the amount of labor needed

for both development and maintenance.

An additional NII-related topic is social learning methods. Here, small amounts of information from multiple users are combined to provide individually customized advice to each. The recommendation of news articles to individuals provides an example. After observing a small number of articles that user A reads and likes, a social learning system might suggest additional articles by correlating A's interests with other users, then recommending articles liked by the most similar other users.

3.3 Reasoning about Plans, Programs, and Action

The field of planning develops algorithms that automatically construct and execute sequences of primitive commands in order to achieve high-level goals. Research focuses on designing languages for modeling dynamic systems and devising algorithms that synthesize possible courses of action. Issues revolve around tradeoffs involving the expressiveness of modeling languages, the specification of performance measures, and the complexity of the underlying search problems.

3.3.1 Relevance to the NII

Using networked computing and information services effectively requires an understanding of their capabilities and the ability to chain services together to achieve complex objectives. For example, the Internet Netfind service can determine a person's email address but only if provided with distinguishing information about the person, such as his or her city or institutional affiliation. AI planning systems can automatically reason about formal models of Netfind and other utilities to focus information-gathering activities in profitable directions (for example, first determining the person's city, then calling Netfind). Because NII users will routinely generate information gathering tasks, AI planners can efficiently assist users in navigating networks and managing the costs of access and retrieval. In contrast, existing search tools (such as Archie, Veronica, and Anarchie) are limited

by inflexible strategies and the lack of a predictive model of the dynamic environment in which they operate. This dynamic aspect is one of the most important and challenging features of the NII. The associated decision problems are extremely complex given the heterogeneous computing environment, scores of separate and largely incompatible databases, diverse methods of access, and complicated protocols for communication.

Planning algorithms could provide hidden but essential functions for the NII. They could enable development of software robots that use planning technology to consult repositories of programs, protocols, and indices and construct plans to satisfy users' requests for information or services. As a result, users would not have to concern themselves with new programs, archives, and protocols as they become available. Instead, they could specify what they want, and leave to planning algorithms the determination of how to achieve the goals. By exploiting planning technology, intelligent software agents could serve as user advocates and make the best use of available resources.

3.3.2 State of the Art

Successful planning systems have been developed for several tasks, including factory automation, military transportation scheduling, and medical treatment planning. Researchers interested in NII applications are beginning to develop software agents that take information-gathering goals supplied by users and then plan and execute actions to achieve these goals. The generation of plans of action for using NII services is a special case of automatic programming in which the programs (plans) involve loops and conditional branches, with primitive statements couched in terms of basic commands to local and remote networking and database servers. Although early planning systems could only generate straight-line programs, recent work has extended the plan language to include conditionals (for example, if the NC-SA site is available, then get the file there, otherwise get it from CERN), and prototype planners are being developed to automatically synthesize

loops (for example, repeatedly attempt to access this server on one minute intervals until successful). Current planners can handle the expressive goal languages that will be demanded by NII applications—least-commitment planning algorithms can satisfy goals that are composed using disjunction, negation, and nested quantification. To date, most planning research assumes complete information; this restriction is a definite obstacle to the application of planning technology to NII domains where incomplete information is the norm. The research community has recognized this challenge; several recent planning systems include principled techniques for coping with uncertainty and sensing actions.

3.3.3 Research Opportunities

The NII offers a perfect target for modern planning theory. First, the critical information required for decision making is readily available in symbolic form from electronic sources, thereby eliminating or reducing many of the difficult interpretation problems encountered in planning in areas like robotics. Second, the actions involve computer programs with well-understood semantics and input-output behavior that is easily observed and interpreted by other computer programs. Finally, NII-related problems encourage the use of predictive models, involve manageable levels of uncertainty, and are characterized by clear performance criteria.

Planning systems for the NII will have to cope with uncertainty regarding the availability of services. They will have to make plans with radically incomplete and possibly out-of-date information. Planning systems will need to cope with the tradeoffs between the benefit of computing the best possible plan and the need to act quickly—before all avenues have been explored. Similarly, NII planning systems must balance the benefit accrued from high-quality information sources with the cost of invoking premium databases. Although modern planning algorithms can handle expressive action representation languages that are capable of representing the rich variety of In-

ternet utilities and NII services, combinatorial problems might prevent such algorithms from scaling to handle large problems. To combat this difficulty, research is needed on search-control languages and domain compilation techniques.

Modern planning theory is beginning to understand the issues involved in reasoning with incomplete information, and coping with costs and time pressure. The first working systems to realize existing theories are now being tested on real applications. The NII will provide an ideal framework in which to evaluate and refine these theories. The success of these prototype systems will be measured by the increased efficiency of information workers who can spend less time searching for and through information.

3.4 Plausible Reasoning

The AI field of plausible reasoning has tackled the problem of representing, understanding, and controlling the behavior of agents or other systems in the context of incomplete or incorrect information. This research has led to a number of techniques, algorithms, and implemented systems for describing, diagnosing, and manipulating both natural and man-made artifacts. By basing these techniques on the sound foundation of probability theory, AI researchers have been able to assemble large knowledge bases in a principled way; effectively perform both predictive and diagnostic reasoning about the system; and develop control policies or plans that could enable the system to act safely, effectively, and efficiently.

At least four themes underlie this work: (1) making the structure of the target system explicit; (2) representing uncertainty about the system coherently and explicitly; (3) updating beliefs about the system as new information about it is received; and (4) reasoning about tradeoffs or relative likelihoods in the prediction, diagnosis, or decision task. A hallmark of this approach has been the development of models that combine both structural and numeric components. The *structural* (or symbolic)

component indicates dependencies or independencies among system components; the *numeric component* quantifies the extent of the dependency, the strength of belief in a relationship, or the relative likelihood of various results.

3.4.1 Relevance to the NII

Uncertainty, incomplete or incorrect information, and the need to generate high-quality behavior or make high-quality predictions or diagnoses will be central to NII applications. An intelligent agent cannot possibly have complete and timely information about Internet, diagnosing problems with human or complex non-human systems is inexact and characterized by noisy and conflicting information, and manufacturing and logistical control problems are fraught with uncertainty. In all cases, there is a need for developing cost-effective solutions, which requires reasoning about tradeoffs between the cost and likelihood of success, the relative quality of alternative courses of action, and the value of obtaining more information versus the cost of doing so.

3.4.2 State of the Art

Most of the historical success of probabilistic or decision-theoretic models has been expert applications. In these applications, a model (or at least the structural component of the model) for a domain is elicited from a human expert and is then used to solve a wide range of problems from the domain. These systems have achieved a high level of effectiveness, size, speed, and accuracy for tasks such as diagnosing diseases and suggesting treatments in human medical care as well as troubleshooting problems with complex devices such as aircraft. This success is in large part owed to the fact that the framework provides a natural way to capture crucial knowledge about the system: structural regularities, degree of uncertainty, value of information, preferences or utilities, and relative likelihood.

Research has also addressed the planning or decision-making problem, generating courses of action that solve a problem or carry out a task effectively. This work has been applied both to aiding human decision mak-

ers in constructing and solving domain models and to importing representations and algorithms from classical AI, control, and stochastic optimization to build and solve these models automatically. Work in decision making has also been extended to the concept of *meta-rationality*; in planning an effective course of action the agent must take into account not only the cost of taking action but also the cost of delaying action.

Recent work has also been directed toward learning structural models and numeric parameters from data from existing data or knowledge bases, statistical databases, or the agent's own experience in diagnosis or problem solving.

3.4.3 Research Opportunities

Four basic challenges face current research in plausible inference; two concern development of new methods and two concern modelling new kinds of information.

First, methods must be devised for automatic construction of network structures from fragmentary input, especially input involving a combination of formal specifications, natural language databases, raw statistical data, or existing databases or knowledge sources. In addition to the current emphasis on building a monolithic model that is then applied to many situations, some applications require that a model be built on demand to solve a new and novel problem. Second, abstraction and approximation methods are needed for handling very large network structures. Promising avenues include sampling techniques, methods for pruning scenarios of low likelihood or low relevance, and exploitation of structural regularities in large databases. The third challenge is developing models of user preferences. To build high-quality solutions to problems, an agent must have a good definition of quality that takes into account a user's preferences. Representations must be developed that simultaneously are rich, are easily elicited, and can be used to solve the problem effectively. Finally, the capabilities of current systems must be expanded to encompass the capability of communicating with, and reasoning about, mental at-

titudes, including the development of semantics and algorithms for *introspective reasoning* (reasoning about the agent's own mental states) and *social reasoning* (reasoning about the mental states of other agents).

3.5 Agent Architecture

Agents, as we defined previously, are entities capable of autonomous goal-oriented behavior in some environment, often in the service of larger-scale goals external to themselves. The *architecture* of an agent is the computational structure that, along with the more dynamic knowledge represented within it, generates the agent's behavior in the environment. The architecture must contain structures that enable representing knowledge (Subsection 3.1), representing and achieving goals (Subsection 3.3), interacting with the environment, and coping with unexpected occurrences. Moreover, for many domains, these capabilities must be exhibited in real time. Depending on the nature of the environment, other agents (either human or virtual) in the environment, and the kinds of task the agent should perform in the environment, other capabilities may also need to be supported in the agent's architecture; for example, coordination and collaboration (Subsection 3.6), language use (Subsection 3.8), learning (Subsection 3.2), and humanlike behavior and affect.

3.5.1 Relevance to the NII

Agent architectures provide the necessary infrastructure for agents that fill critical roles in both intelligent user interfaces (Subsection 2.1) and software development tools and environments (Subsection 2.3). For example, an *intelligent project coach* (Subsection 2.3.1.1) is an agent that helps analysts and designers achieve their larger-scale goals in a project environment by recording and explaining choices behind design decisions. The architecture for such an agent needs to provide a basis for representing design knowledge; interacting with the design environment; coping with unexpected occurrences; collaborating with designers; using language (for explanations); and learning about designs, designing, and designers. Simi-

larly, an agent that assists in large-scale group training on the NII by populating a virtual-reality environment that includes other agents (Subsection 2.1.4), such as collaborators, competitors, assistants, leaders and instructors, needs to be built on an architecture that provides most of the capabilities listed previously (and quite possibly more).

The key problem in agent architecture is finding a compatible mix of the necessary capabilities and integrating them together to support appropriate, most likely real-time, behavior. Compatibility is key here. True intelligent behavior requires such a close integration of these capabilities that arbitrary combinations of them are as likely to degrade performance as enhance it. For example, a main result from research on the combination of learning and planning is that the acquisition of rules intended to improve the speed of the planner usually has the inverse effect unless the combination is done just right.

3.5.2 State of the Art

The field of agent architecture is just starting to reach maturity. A rich body of work is now available on the individual capabilities, along with real-time versions of several of them. Building on this base, dozens of proposals have recently been generated about how to combine small subsets of the total set of capabilities (although mostly not in real time). A small number of systems have demonstrated the ability to exhibit a handful or more of the key capabilities in real time. For example, one such system has yielded automated intelligent pilots that have successfully participated in an operational military-training exercise.

3.5.3 Research Opportunities

An effective and efficient integration of all the key capabilities is still a long-term project. However, many high-value applications only require subsets of these capabilities; for example, an intelligent project coach might have only weak real-time requirements and no requirement to behave in a humanlike manner. Applications might also be able to get by with approximate versions of oth-

er capabilities and still perform useful functions. What is therefore needed now—and what the field is clearly ready to provide—are significant pushes to develop both real-time versions of a wider set of the key capabilities and larger and more varied combinations of capabilities (even if not in real time). The investigation of incremental approaches to integration—in which initially small subsets of prototype capabilities are combined and applied, and then both the number of capabilities and the quality of the resulting behavior are gradually improved and applied in a wider range of domains—is one promising means of addressing this latter need.

3.6 Multiagent Coordination and Collaboration

The field of multiagent coordination has studied the problem of endowing agents with the ability to communicate with each other to reach mutually beneficial agreements. Specialized techniques have been developed to enable an agent to represent and reason about the capabilities of other agents. Research on collaboration has led to representations of the information agents must establish and specifications of the information agents must communicate in order to collaborate. Research has also led to algorithms that enable two agents to communicate their differing objectives, determine areas of shared interest, and converge on Pareto-optimal agreements that increase the utility of all participants.

3.6.1 Relevance to the NII

In Section 2, we described intelligent agents that act as personal assistants and software brokers that support information retrieval and other advanced services. These possibilities suggest that the majority of network interactions will eventually be between programs. To make these interactions flexible, intelligent coordination and collaboration between agents will be essential. Furthermore, collaborative capabilities could significantly enhance human-computer interfaces. A sampling of tasks such agents could perform illustrates the promise and challenge of multiagent coordination and collaboration.

Payment and delivery of services: As electronic commerce blossoms on the NII, organizations and individuals will require assistance in finding the most attractive product or service among potentially hundreds of such products and services advertised on the NII. Personalized and trusted bargaining agents could act on behalf of their users by first finding information and then negotiating with selling agents over price, conditions of payment, and delivery schedules.

Yellow-pages and consumer reports agents: Creating, finding, and providing new information services, and helping people find the information they seek present major problems for the NII. Capabilities for constructing information brokers that constantly monitor a single, relatively narrow area of interest can help address these problems. Each broker would construct an extensive representation of the content of relevant information sources, the capabilities of service providers, and the scope of related brokers.

3.6.2 State of the Art

The field of multiagent systems has integrated ideas from economics and linguistics with those of computer science. Game theory provides a solid mathematical foundation for studying collaboration and negotiation algorithms, but game theory by itself is insufficient because it does not provide algorithms for computing optimal strategies or determining equilibrium courses of behavior. AI research uses game-theoretic concepts as a guide in the design and analysis of practical agent collaboration and negotiation algorithms. Achievements include the following: identification of *protocols* (global constraints on messages between the agents) that lead to quick agreements and reduce the incentive for one agent to try and deceive another; design of interagent communication paradigms and languages (for example, formalizations of speech-act theory, KQML, and AOP), for making interagent requests and coming to consensus about the mental states of other agents; and successful implementation of multiagent systems in a variety of application domains (for

example, transportation planning, distributed resource allocation, telephone network management, sensor interpretation, manufacturing, and factory automation).

3.6.3 Research Opportunities

The NII presents a variety of challenges for AI research in collaboration and coordination algorithms. The vast number of communicating intelligent agents will challenge the scalability of current theories and methods. Agents must be able to efficiently select the most knowledgeable set of partners with whom to coordinate in service of a task or information-gathering objective. As a result, agents need means for advertising their existence, interests and services; narrow casting methods will be a crucial component in insuring that agents do not get swamped with irrelevant junk messages.

Because different intelligent agents can have different mandates, flexible incentive structures will be necessary to assure cooperation. Negotiation will be crucial for resolving conflicts in goals, information and results, and negotiation algorithms must take into consideration tradeoffs between the time spent searching for appropriate agents and information sources, the time to access a given service, and the information quality and timeliness of information delivery. Techniques developed for negotiation must be extended to deal with situations in which people as well as computer agents participate in the collaborative or coordinated activity.

Research on theories of collaboration and coordination among multiple agents provides insight into these issues and tradeoffs as they occur for a relatively small number of nearly homogeneous agents. The NII provides a test-bed environment for scaling up and refining these theories to deal with very large and heterogeneous communities, in which the set of agents changes dynamically; new services appear; and the underlying languages, protocols, and ontologies evolve over time.

3.7 Ontological Development

The goal of research in ontologies is to create explicit, formal catalogs of

knowledge that can be used by intelligent systems. An *ontology* is a theory of a particular domain or sphere of knowledge, describing the kinds of entity involved in it and the relationships that can hold among different entities. An ontology for finance, for example, would provide working definitions of concepts like money, banks, and stocks. This knowledge is expressed in computer-usable formalisms; for example, an agent for personal finances would draw on its finance ontology, as well as knowledge of your particular circumstances, to look for appropriate investments. Ontologies are broad, in that they cover a wide range of phenomena and situations. They are multi-purpose in that the same ontology can be used in different programs to accomplish a variety of tasks.

Building ontologies is difficult for three reasons. First, articulating knowledge in sufficient detail that it can be expressed in computationally effective formalisms is hard. Second, the scope of shared background knowledge underlying interactions of two agents can be enormous. For example, two doctors collaborating to reach a diagnosis might combine commonsense conclusions based on a patient's lifestyle with their specialized knowledge. Third, there are unsolved problems in using large bodies of knowledge effectively, including selecting relevant subsets of knowledge, handling incomplete information, and resolving inconsistencies.

3.7.1 Relevance to the NII

The creation of repositories of ontologies that can be used by intelligent software robots is crucial to the NII because ontologies provide the shared conceptualizations that are required for communication and collaboration. For example, an engineer's software agent needs to understand the design rationale and function of a subsystem to detect the possible impact of other design decisions on the subsystem. Data-mining intelligent agents need to understand the contents of databases to integrate information from disparate sources. Robust representations of commonsense knowledge will be essential for agents that communicate with peo-

ple; the inability to draw on the background knowledge we share with other people is one reason computers today are so difficult to use.

3.7.2 State of the Art

Despite its fundamental importance, the accumulation of ontologies has only just begun. Techniques for organizing ontologies, combining smaller ontologies to form larger systems, and using this knowledge effectively are all in their infancy. There are few collections of ontologies in existence; almost all are still under development, and currently none of them are widely used.

Efforts are under way to create ontologies for a variety of central commonsense phenomena, including time, space, motion, process, and quantity. Research in qualitative reasoning has led to the creation of techniques for organizing large bodies of knowledge for engineering domains and automatic model-formulation algorithms that can select what subset of this knowledge is relevant for certain tasks. Although these efforts are promising, they are only in the preliminary stages of development. The natural language community has invested in a different form of ontological development. WordNet is a simple but comprehensive taxonomy of about 70,000 interrelated concepts that is being used in machine translation systems, health care applications, and World Wide Web interfaces.

Another important development has been the creation of easy-to-use tools for creating, evaluating, accessing, using, and maintaining reusable ontologies by both individuals and groups. The motivation is that ontology construction is difficult and time consuming and is a major barrier to the building of large-scale intelligent systems and software agents. Because many conceptualizations are intended to be useful for a wide variety of tasks, an important means of removing this barrier is to encode ontologies in a reusable form so that large portions of an ontology for a given application can be assembled from smaller ontologies, that are drawn from repositories. This work is also only in the preliminary stages of development.

3.7.3 Research Opportunities

Several research directions offer exceptional payback for NII infrastructure and applications: developing reusable ontologies for commonsense concepts, such as physical concepts (for example, time, space, material properties), NII concepts (such as, computers, networks, documents, bandwidth), social concepts (such as privacy and harm), and mental concepts (such as forgetting and attention); defining semiformal representation languages that support descriptions both informally in natural language and formally in a computer-interpretable knowledge representation language; implementing the next generation of ontology construction tools (these tools should include capabilities for browsing and visualizing ontologies, detecting inconsistencies, and semiautonomously synthesizing ontologies based on the use of terms in natural language documents; and devising strategies that agents can use to detect communications problems stemming from inconsistent ontologies and developing translation algorithms so that intelligent agents can agree on a common communication substrate.

3.8 Speech and Language Processing

The ultimate goal of natural language-processing (NLP⁵) research is to create systems able to communicate with people in natural languages. Such communication requires an ability to understand the meaning and purpose of communicative actions, such as spoken utterances, written texts, and the gestures that accompany them and an ability to produce such communicative actions appropriately. These abilities, in their most general form, are far beyond our current scientific understanding and computing technology.

Ambiguity is one reason general natural language processing capabilities are difficult to achieve. Human languages all use a small set of resources (such as words, structures, intonations, and gestures) to convey an exceedingly wide, rich, and varied set of meanings. Any one word, structure, or gesture will often be used in

many different ways. Although people rarely notice such lexical, structural, semantic and intonational ambiguities, their identification and resolution challenge current speech- and language-processing systems. Another source of difficulty is the difference between what people say (or write) and what they actually mean. People rely on their audience to understand much that is not explicitly said or written, deriving this information from context and common knowledge. Furthermore, people often begin to speak or write before their ideas are well thought out, using the formulation of real utterances as a step in understanding their own partially formed ideas. Both practices result in partial and imperfect evidence for what people are really trying to communicate.

3.8.1 Relevance to the National Information Infrastructure

Because natural language is often the preferred way to communicate with people, it will undoubtedly become a popular means for communicating with computers. Natural language is also currently the most prevalent medium for knowledge representation; most of what humankind knows is stored as written text. Thus, the potential relevance of natural language processing to the NII is immense. Natural language understanding and generation could be central to the next generation of intelligent interfaces; in the short term at least, information management will largely mean text management; and natural language processing of some kind will be needed to facilitate the cooperative work environments required for efficient design and development of the NII.

3.8.2 State of the Art

Natural language processing research has resulted in several significant achievements, including the following: techniques for parsing, semantic interpretation, and discourse modeling sufficient to process realistic database queries posed in natural language; the reuse of language-understanding techniques for generation of reports customized to context, task, and user; statistical models of speech acoustics, word pro-

nunciation, and word sequencing that are sufficiently accurate to support usable speech understanding of restricted vocabulary utterances; speech-generation systems that can generate spoken language with intonation contours that begin to conform to and reflect intended meaning and underlying context; machine-translation systems that can improve the efficiency of human translators by providing useful first drafts; and content-based retrieval systems that can glean useful information from unstructured text documents.

A relatively recent development is the use of statistical models. Typically generated automatically, statistical models can predict with good accuracy simple grammatical features of utterances such as a word's part of speech, as well as semantic properties such as a word's most likely sense in a given context; they thus reduce problems caused by ambiguities in the grammatical and semantic properties of words. Improved discourse modeling is another result of recent research; idealized models of purposive communicative action have been developed based on empirical studies, planning, and collaboration, and the models have been incorporated in experimental systems.

3.8.3 Research Opportunities

Much of the success of current natural language processing technology has come from a long, tedious process of incremental improvement in existing approaches. More work of this kind is needed to extract the best possible performance from known techniques. In addition, there are significant research opportunities in exploring new and combined approaches. For example, although statistical and machine-learning techniques in natural language processing offer broad (but shallow) coverage and robustness with respect to noise and errors, grammatical and logical techniques offer deeper analyses of meaning, purpose, and discourse structure. These two types of techniques could complement one another. The symbolic techniques might serve to specify a space of interpretation possibilities and the statistical techniques might serve to

evaluate efficiently the evidence for alternative interpretations. The results of such integration will be of value to all natural language processing applications, from information extraction and machine translation to collaborative interfaces. Another way to exploit known technology maximally might be to determine how natural language processing technology can be combined most effectively with other AI and non-AI technologies to forge effective multimodal user interfaces.

3.9 Image Understanding and Synthesis

The speed with which people extract information from images makes vision the preferred perceptual modality for most people in the majority of tasks, thus implying that easy-to-use computers should be capable of both understanding and synthesizing images. One of the goals of computer-vision research is image understanding and classification. Depending on the application, the imagery to be understood might include a scanned document page, a mug shot, an aerial photograph, or a video of a home or office scene. In each instance, what it means to understand the image (or image sequence) and the ways this understanding is accomplished, can be very different.

In contrast, *image synthesis* is the task of generating artificial imagery; it is the goal of computer-graphics research. Again, the types of image vary greatly (for example, charts and maps, interior and exterior views of buildings, biomedical and scientific visualizations, and cartoon animations), as do the methods for generating them. In the near future, some computer applications might even use a tightly coupled combination of image understanding and synthesis. For example, a three-dimensional fax might scan and understand two-dimensional images of an object at one location; resolve structural ambiguities; compress and transmit a symbolic encoding of the object; and then synthesize a virtual three-dimensional replica that can be manipulated, molded, or even physically manufactured somewhere else. Such

an application would, in effect, be sending a physical object through a computer network.

A general, robust, three-dimensional fax capability is beyond current technology. It is not easy to say in general what characterizes the hard or easy problems in computer graphics and computer vision. Although some vision and graphics problems have yielded to persistent research (such as optical character recognition and photorealistic image synthesis), others continue to challenge us (for example, general real-world scene interpretation and motion control for animated characters). Nonetheless, these two areas are currently among the most dynamic in the field of AI, and the state of the art is constantly changing.

3.9.1 Relevance to the National Information Infrastructure

Image understanding and synthesis are relevant to the development of intelligent user interfaces, digital libraries, and 3D fax capabilities. Automatic design of informational graphics (for example, charts, maps, and scientific visualizations) is a necessary complement to natural language generation; both modalities are needed to facilitate computer applications that can explain themselves. Virtual environments seem sterile and unbelievable unless the human- and computer-controlled agents that inhabit them move in plausible ways. Automatic motion synthesis and motion tracking from video are two technologies that address this problem. Interpretation of manual gestures and facial expressions is an important aspect of human communication that might be incorporated effectively into computer-human interfaces if the necessary computer-vision problems can be solved. On a more mundane level, robust handwriting recognition would create many opportunities for developing new, more natural user interfaces.

Infrastructure services and development tools will also benefit from advances in computer graphics and computer vision. Although the majority of human knowledge remains stored in paper documents, document analysis and recognition will

be needed to convert scanned text and illustrations into symbolic form, thereby facilitating data and knowledge management services. More speculatively, general image understanding would revolutionize knowledge discovery and acquisition; even limited success (for example, the ability to reliably find specific people in photographs or videos) would support extremely useful services. Finally, the difficulty of designing and modeling large-scale virtual environments can be mitigated by applying intelligent modeling techniques that can automatically build scene models from replicas of natural and manmade objects, which might even be acquired initially through computer-vision techniques.

3.9.2 State of the Art

In general, current computer-vision techniques are capable of impressive feats under controlled conditions, but these techniques often prove to be brittle and nonrobust under real-world conditions. The state-of-the-art in four typical tasks illustrates this point.

Facial recognition: A variety of different algorithmic approaches can recognize standard mug shots front-facing, head-only photographs under controlled lighting with high accuracy; however, identifying a face in an image taken in a more realistic setting cannot be done reliably.

Object recognition and reconstruction: Under ideal lighting and viewing conditions, simple known objects (for example, a coffee mug, a rubber duck) can be recognized, and simple unknown objects can be reconstructed, but these techniques often fail completely under less favorable conditions.

Hand tracking and gesture recognition: Under ideal conditions, the movement and configuration of a human hand can be tracked with high accuracy; however, no current system can interpret sign language in a practical setting. Automatic recognition of facial expressions is becoming a reality.

Document analysis and recognition: A one-column document that is cleanly typed in an appropriate font can be interpreted with high fidelity; however, a poor-quality multicolumn docu-

ment with irregular layout and mixed fonts can baffle the best of the current systems.

Three-dimensional computer graphics problems can be neatly divided into two categories: modeling and rendering. Modeling is the problem of acquiring, representing, and manipulating a symbolic description of the objects in a static or moving scene. Rendering is the problem of converting a scene model into the appropriate two-dimensional image. In general, image rendering is well understood; current techniques are capable of accounting for great subtlety in lighting and shading phenomena to generate amazingly realistic imagery. Modeling, however, is not as advanced. Making an articulated figure move in a visually plausible way or designing a solid model of a manmade artifact are difficult. Typically, the best tools available for these tasks are direct-manipulation user interfaces, which are tedious and hard to use. Only recently have researchers begun to apply, to some degree, AI tools and ideas with a view toward automating the hard modeling tasks. Informational graphics, which are mostly two-dimensional but can be three-dimensional, form an essentially distinct category of graphics, but one that is of great importance for information analysis and presentation. Research in this area has also focused mostly on improved user interfaces for the manual creation of images such as charts and maps; relatively little research has been conducted to date on the automatic design of such graphics.

3.9.3 Research Opportunities

Much of the low-level image-understanding technology is relatively mature. For example, there are fairly reliable techniques for recovering visual motion, stereo reconstruction, and determining structure from motion. Similarly, in graphics, low-level processes such as rendering are almost-solved problems. These successes account for the positive capabilities listed earlier. Increasingly, the research challenge and opportunity is in higher-level processing, representation, and reasoning about visual and geometric information. Only at

higher levels is there enough information to counter the noise and ambiguities that are the source of brittleness in current techniques. For example, emerging AI areas such as artificial life and evolutionary computation are beginning to show promise for such representation and reasoning problems. The key challenge in this area will be to build on the prior results of AI, vision, robotics, and graphics, and incorporate new concepts and technologies.

4. Conclusions

The National Information Infrastructure (NII) promises to deliver to people in their homes and businesses a vast array of information in many forms. It could significantly improve many aspects of citizens' lives. However, for the NII to realize its potential, it must provide efficient and easy access without requiring specialized training. The field of AI is positioned to make substantial contributions to the NII. AI techniques can play a central role in the development of a useful and usable NII by (1) enabling the construction of human-computer interface systems that are goal-oriented, cooperative, and customizable; allow users to communicate in natural ways in a variety of modalities; and provide a consistent interface to the full range of NII services; (2) providing services for data and knowledge management integration and translation, and knowledge discovery in support of a more flexible infrastructure; and, (3) assisting in the development of more powerful software tools and environments by adding a range of advanced capabilities to rapid prototyping systems; enabling the development of intelligent project management aids; and supporting the construction of simulation systems that include more sophisticated simulated agents, including characters that act in human-like ways.

This report recommends support of AI research in eight key areas, each of which has substantial promise for high payback to the NII effort: knowledge representation; learning and adaptation; reasoning about plans, programs, and actions; plausi-

ble reasoning; agent architecture; multiagent coordination and collaboration; development of ontologies; speech and text processing; and, image understanding and synthesis.

Acknowledgments

This report was created with support from the Information Technology and Organizations Program of the IRIS Division of the National Science Foundation. The opinions presented in the report do not represent the views of the National Science Foundation.

Paul Beame, Pat Hayes, Richard Korf, Ed Lazowska, Aaron Sloman, and Lynn Stein gave many thoughtful suggestions for ways to improve this report. Many thanks also to Alicen Smith, Sunny Ludvik, and Carol Hamilton who helped with proofreading and copyediting.

Report Contributors

Ruzena Bajcsy (University of Pennsylvania) Ronald J. Brachman (AT&T Bell Labs), Bruce Buchanan (University of Pittsburgh), Randall Davis (Massachusetts Institute of Technology), Thomas L. Dean (Brown University), Johan de Kleer (Xerox Palo Alto Research Center), Jon Doyle (Massachusetts Institute of Technology), Oren Etzioni (University of Washington), Richard Fikes (Stanford University), Kenneth Forbus (Northwestern University), Barbara Grosz (Harvard University), Steve Hanks (University of Washington), Julia Hirshberg (AT&T Bell Labs), Ed Hovy (University of Southern California, Information Sciences Institute), Daniel Huttenlocher (Cornell University), Robert E. Kahn (CNRI), Jean-Claude Latombe (Stanford University), Yvan LeClerc (SRI International), Thomas Mitchell (Carnegie Mellon University), Ramesh Patil (University of Southern California, Information Sciences Institute), Judea Pearl (University of California, Los Angeles), Fernando C. N. Pereira (AT&T Bell Labs), Paul S. Rosenbloom (University of Southern California, Information Sciences Institute), Stuart J. Russell (University of California, Berkeley), Katia

Sycara (Carnegie Mellon University), Bonnie L. Webber (University of Pennsylvania), Beverly Woolf (University of Massachusetts)

NSF Observers

Su-Shing Chen and Y. T. Chien

References

- 1 Grosz, B., and Davis, R., eds. 1994. A Report to ARPA on Twenty-First Century Intelligent Systems. *AI Magazine* 15(3): 10-20. <http://www.aaai.org>
- 2 Clement, M., Katz, R., and Chien, Y., eds. 1994. Information Infrastructure Technology and Applications. <http://www.hpcc.gov/reports/reports-nco/iita.rpt.html>
- 3 The Innovative Applications of Artificial Intelligence Conference *Proceedings* series (Menlo Park, California: AAAI Press) contains numerous examples of successfully deployed expert systems.
- 4 Committee on Physical, Mathematical, and Engineering Sciences (CPMES) and the Federal Coordinating Council for Science, Engineering, and Technology (FCC-SET). *High Performance Computing and Communications: Toward a National Information Infrastructure*. Washington, D.C.: Office of Science and Technology Policy, 1994; and IITA Task Group. *Information Infrastructure Technology and Applications*. Washington, D.C.: Office of Science and Technology Policy, February 1994.
- 5 We use the term "NLP" here both for speech and text processing, although speech and text processing have distinct histories, research communities, and terminology.