RIACS

Research Institute for Advanced Computer Science
NASA Ames Research Center

# RIACS FY2002 Annual Report

**Barry M. Leiner**

**RIACS Technical Report AR-02**

**November 2002**

**THIS PAGE IS INTENTIONALLY LEFT BLANK**

# Research Institute for Advanced Computer Science

## ANNUAL REPORT

**October 2001 through September 2002**

**Cooperative Agreement: NCC 2-1006**

Submitted to:

Contracting Office Technical Monitor:
Anthony R. Gross
Associate Director for Programs
Information Sciences & Technology Directorate
NASA Ames Research Center
Mail Stop 200-6

Submitted by:
**Research Institute for Advanced Computer Science (RIACS)**

Operated by:
**Universities Space Research Association (USRA)**

RIACS Principal Investigator:
Dr. Barry M. Leiner
Director

Dr. Barry M. Leiner

**THIS PAGE IS INTENTIONALLY LEFT BLANK**

# TABLE OF CONTENTS

## LIST OF FIGURES AND TABLES

# I. Introduction and Overview

The Research Institute for Advanced Computer Science (RIACS) carries out basic research and technology development in computer science, in support of the National Aeronautics and Space Administration's missions. Operated by the Universities Space Research Association (a non-profit universitiy consortium), RIACS is located at the NASA Ames Research Center, Moffett Field, California. It currently operates under a multiple year grant/cooperative agreement that began on October 1, 1997 and is up for renewal in September 2003.

Ames has been designated NASA's Center of Excellence in Information Technology. In this capacity, Ames is charged with the responsibility to build an Information Technology (IT) Research Program that is preeminent within NASA. RIACS serves as a bridge between NASA Ames and the academic community, and RIACS scientists and visitors work in close collaboration with NASA scientists. RIACS has the additional goal of broadening the base of researchers in these areas of importance to the nation's space and aeronautics enterprises.

RIACS research focuses on the three cornerstones of IT research necessary to meet the future challenges of NASA missions:

- Automated Reasoning for Autonomous Systems
  Techniques are being developed enabling spacecraft that will be self-guiding and self-correcting to the extent that they will require little or no human intervention. Such craft will be equipped to independently solve problems as they arise, and fulfill their missions with minimum direction from Earth.

- Human-Centered Computing
  Many NASA missions require synergy between humans and computers, with sophisticated computational aids amplifying human cognitive and perceptual abilities.

- High Performance Computing and Networking
  Advances in the performance of computing and networking continue to have major impact on a variety of NASA endeavors, ranging from modeling and simulation to analysis of large scientific datasets to collaborative engineering, planning and execution.

In addition, RIACS collaborates with NASA scientists to apply IT research to a variety of NASA application domains including aerospace technology, earth science, life sciences, and astrobiology. RIACS also engages in other activities, such as workshops, seminars, visiting scientist programs and student summer programs, designed to encourage and facilitate collaboration between the university and NASA IT research communities.

## *I.A. Summary of FY2002 Activity*

During the year October 1, 2001 through September 30, 2002, RIACS engaged in a number of research projects collaboratively with NASA scientists. Over 180 publications and presentations resulted from this work. Section II describes each of those projects along with a list of publications. In addition, RIACS continued to run a successful bi-weekly seminar series, augmented with special seminars, as well as support and participate in a number of workshops. During FY2002, we hosted 56 seminars and supported four workshops. RIACS also maintains an online library of the technical reports generated by staff scientists, and provides hard copies to interested parties upon request. Twelve technical reports were produced this year. The seminars, workshops and technical reports are shown in Section III. Finally, RIACS continued the Ames/RIACS Summer Student Research Program (SSRP), a competitive program in which selected students spend the summer at Ames working with Ames researchers. A description of the students' specific efforts may be found in Section II.E.1.

In addition to participating in initiatives by others at Ames (described in the various projects below), we have initiated several activities. We have continued to work with Ames personnel in establishing an Ames thrust in spoken dialogue natural language interfaces. This activity is in the process of transitioning from a RIACS activity to an Ames (Code IC/CAS) activity in which RIACS participates. See Section II.B.2 for more detail.

In support of Ames interest in leveraging its strengths in biology, astrobiology, nanotechnology, and IT, RIACS worked with USRA and NASA Ames to organize a community forum/workshop to discuss such leverage against the needs of space-based research. This workshop, titled *Biology/Information Science/Nanotechnology Fusion, and NASA Missions*, was held in October 2002 and is described in Section I.B.

Information management technologies are important to NASA and more generally, the public sector. Late last year, we initiated a panel study into the information management technology requirements for applications of particular government interest, such as digital libraries, mission operations, and science data management. The study was jointly supported by NASA, NSF, and DARPA, and the final report will be available shortly. See Section II.B.5 for more detail.

Staff size resumed its growth this year, with the number of staff scientists growing from 35 last year to 41 at the end of this year. In addition, five scientists - Daniel Clancy, James Hieronymus, Butler Hine, David Maluf, and Kanna Rajan – are now on loan to NASA under the IPA program. 52 visiting scientists and 44 visiting students spent time at RIACS during the year. Section IV provides more detail on staffing.

We succeeded in recruiting a new Deputy Director, Dr. Serdar Uckun. Dr. Uckun has an M.D. from Ege University, Izmir, Turkey, an M.S. in Biomedical Engineering from Bogazici University, Istanbul, Turkey, and a Ph.D. in Biomedical Engineering from Vanderbilt University. Between 1992 and 1994, he received postdoctoral training in Computer Science at the Knowledge Systems Laboratory at Stanford University. Prior to joining RIACS in March 2002, Dr. Uckun was the Director of Advanced Technology at Blue Pumpkin Software, and before that, Assistant Director and Manager of the Intelligent Systems Department at Rockwell Science Center, Palo Alto, CA. His research interests include decision making under uncertainty, situational awareness, and scheduling.

Three Associate Directors coordinate areas with related projects. Kathleen Connell is responsible for projects investigating the synergies of biology and astrobiology activities with IT. Richard Washington is responsible for projects in autonomous systems, including topics such as automated reasoning, planning and scheduling. Johann Schumann is responsible for projects in software engineering, including the areas of software synthesis, verification and validation, and distributed frameworks. Their overviews of these areas follow.

## *I.B. Bio/IT Fusion Area Overview*

Kathleen Connell (Associate Director)

The promise of convergence science and technology areas, combined with continued findings in cell and molecular biology, proved to be a fertile investment for RIACS in 2002. The 'Bio/IT' Group at RIACS is one of only a handful of multi-disciplinary groups within the civilian space sector. Nevertheless, this sort of convergence effort promises to be a pathfinder for future science and technology breakthroughs, which will, in turn, generate the next technical foundations for the future of space exploration in support of our national objectives.

In addition to the accomplishments of RIACS Bio/IT Staff, as described in the task descriptions II.D.1 through II.D.5 below, a milestone accomplishment for RIACS was the successful execution of the 'Biology, Information Technology and Nanotechnology And NASA Missions' workshop, co-organized by the RIACS Associate Director K.M. Connell, USRA Chief Engineer Lewis Peach, NASA Ames Director G. Scott Hubbard, and Dr. T.R. Govindan, of Code I, NASA Ames, with support from the RIACS Director Barry Leiner. This workshop brought together over 100 representatives of every NASA enterprise, NASA personnel from advanced technology in biology, information technology and nanotechnology, and over 50 representatives of universities from around the country. The findings of this first of its kind workshop will be published later this year, and are likely to be a pathfinder in this new and important field for the Agency and for the science and technology community as a whole. See http://binfusion.arc.nasa.gov for further information.

**RIACS Bio/IT Staff (as of 30 Sep 02)**

Kathleen Connell, Associate Director
Lisa Faithorn
Karl Schweighofer
Xander Twombly
Senior Team Advisor: Lewis Peach, Chief Engineer, USRA

## *I.C. Autonomous Systems Area Overview*

Richard Washington (Associate Director)

NASA is extending our view of our own planet and the universe with missions that push the limits of today's technologies. The new generation of missions must accomplish difficult tasks with only limited ground support. As NASA spacecraft venture further to places that severely limit communication, such as the seas of Europa or the surface of Titan, the spacecraft must be able to perform largely on their own for long periods of time. In the past, the need for autonomous behavior has led to complex, one-off engineered approaches, leading in turn to missions with ever-increasing complexity and cost. Current fiscal realities make this no longer a tenable path. Instead, the new generation of explorers will rely on autonomous systems built from the range of reusable autonomy technologies under development at NASA and RIACS.

On-board autonomy is complemented by ground-based automation of mission operations. As systems and mission operations become more complex, a manual approach to labor-intensive daily mission planning activities will prove impractical given tight timelines and budget constraints. Assistance from automated planning and scheduling systems augments the human decision-making process, improving efficiency and ensuring safety. In contrast to on-board autonomy technology, the ground assistance must work in tight collaboration with human operators; this provides both the potential for synergy as well as a challenge to design autonomy technology adapted for human interaction.

The RIACS Autonomous Systems group plays critical roles in the development of autonomy technology, with contributions ranging from advances in fundamental technology research to infusion of autonomy technology into missions. RIACS scientists are primarily focused on the topics of Automated Planning, Robust Execution, Fault Diagnosis, and Computer Vision. In addition, RIACS personnel have played central roles in the infusion of NASA Ames autonomy technology into the upcoming Mars Exploration Rover (MER) and Mars Science Laboratory (MSL) Mars rover missions.

*Automated Planning*. Automated planning is necessary both for on-board decision making and for ground support. The Europa planning system is a constraint-based planning system that has served as a base for a large number of application projects as well as a platform for testing fundamental research concepts. It is currently being integrated into a mixed-initiative science planning tool for the Mars Exploration Rover (MER) mission. Other applications include observation scheduling for the SOFIA airborne observatory and planning for constellations of Earth-observing satellites. Europa also forms the core planning engine for the NASA IDEA research project, which is investigating the unification of planning and execution. In addition, core research on temporal planning have been explored and published within this effort.

*Robust Execution*. Robust execution provides the link between human or automated planning and real-time systems. As such, execution systems must enforce the constraints of planners and respect the real-time constraints of running on computationally constrained platform. The Contingent Rover Language (CRL) and the CRL Executive have been designed to provide a flexible and powerful language for plan execution and on-board autonomous operations. The CRL Executive has served as the high-level control for the NASA Ames K9 rover over the past three years, and it is now extended to support concurrency and a rich set of temporal constraints, allowing it to work with a contingent version of the Europa planner. The new version is serving as the command language for a set of Intelligent Systems demonstrations being developed at NASA Ames. The ideas from this system are influencing early design work for the upcoming Mars Science Laboratory

(MSL) rover mission. In addition, the executive system is designed as a platform for research on issues of execution under uncertainty, in particular efficient plan adaptation algorithms.

*Fault Diagnosis.* A spacecraft operating autonomously, potentially out of contact of Earth, must be able to diagnose potential problems to allow an intelligent choice of whether to continue operating or to react to the fault. This capability allows a spacecraft to avoid false positives and false negatives from simple sensor thresholds, which reduce operating efficiency and increase risk. RIACS efforts in diagnosis have concentrated in model-based diagnosis and hybrid diagnosis.

The integrated vehicle health maintenance (IVHM) project, a collaboration between multiple NASA centers and Northrop Grumman, is demonstrating model-based diagnosis for a reusable launch vehicle prototype. The RIACS contribution to this project is primarily in the area of domain knowledge acquisition and modeling, drawing on expertise crossing the boundaries of computer science and aerospace engineering.

Hybrid diagnosis is particularly well suited to planetary rovers, where the continuous nature of the system and its constant interactions with the environment require a diagnostic system that can reason about uncertain, continuous quantities. The goal is to provide on-board capabilities for situation assessment that are practical for upcoming missions such as the MSL rover mission, while providing a significant advance over current state identification techniques.

*Computer Vision.* A vital capability for a planetary rover is to use its sensors to navigate in its environment. The primary external sensors for planetary rover navigation are its cameras. RIACS is involved in a collaboration between NASA Ames and JPL to produce algorithms for "visual odometry" to augment internal position estimation. These algorithms contribute to the task of "visual servoing," or navigating to a target using visual information.

Autonomous systems are critical for the missions that will enable NASA's vision of better understanding our universe. The contributions of the RIACS Autonomous Systems area are central to this effort, both in overcoming the fundamental challenges of autonomous, reliable operations and in applying autonomy technologies to solve important NASA problems.

**RIACS Autonomous Systems Staff (as of 30 Sep 02)**

Richard Washington, Associate Director
Anupa Bajwa
Esfandiar Bandari
Jimi Crawford
Richard Dearden
Ari Jónsson

## I.D. *Software Engineering Area Overview*

Johann Schumann (Associate Director)

Several recent mission failures have been attributed to software-related glitches: Mars Climate Orbiter failed because of a mismatch of unit systems (Metric vs. English); Mars Polar Lander probably crashed because an erroneous signal from a landing sensor was not processed appropriately and therefore directly caused the rocket engine to be cut while still 50 ft. above the landing site. The international space station (ISS) was uncontrollable for a period of time because, after a hardware failure (disk crash), this exception was not propagated and handled appropriately, thus disabling the backup processing units.

In general, ambitious mission profiles and the necessity to obtain best-possible science return with reduced budgets generate a huge pressure on the development of software within NASA. In addition to reliability, many application areas are highly safety-critical (e.g., manned space flight, air traffic control, air/rotocraft research). Any piece of software in these applications must undergo rigorous testing and certification procedures- nowadays a major cost driver despite their inability to catch all bugs in the software.

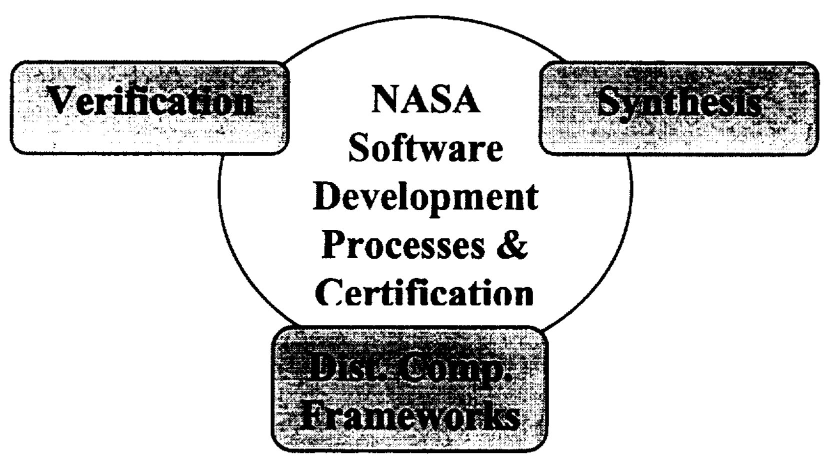


**Figure 1: Software Engineering Activities**

The *Software Engineering* area of RIACS was created to address NASA's needs with this regard. Its global structure is shown in Figure 1. The sub-area of NASA Software Development Processes and Certification works on important issues addressing *how* NASA can develop and deploy software in an orderly fashion. This research theme relies on and is supported by three equally important pillars which help to accomplish these goals:

*Verification* is a methodology to formally demonstrate that a program (or software system) fulfills certain properties (e.g., memory safety), or that certain situations can never occur (e.g., deadlock between multiple processes). More specifically, this area is comprised of the following projects:

- Java Pathfinder is a model checker which, given a Java program, can automatically check if a certain property holds or is violated in certain (often hard to find) cases. This tool has now been released and enjoys a growing number of regular users. Java Pathfinder is being extended to enhance its usability by the software engineer (e.g., by a Java Runtime Monitor, and automatic abstraction).
- With higher and more complex levels of autonomy required for all areas of modern missions, more and more autonomy software is being developed. Because such pieces of software (e.g., planning systems, failure detection, isolation and recovery) are central and important for mission success, *verification of autonomy software* is an important topic. Our sub-area addresses this topic in two projects: system-level verification (i.e., developing methods for handling larger, modular systems), and verification of Livingstone failure models in a closed loop.
- Feedback control systems are an important component in most complex machinery (e.g., aircraft). Classical control systems, however, fail when the underlying characteristics of the plant change. Here, neural network based controllers have been developed to recognize such a change (e.g., a broken wing-tip or a stuck rudder) and to automatically *adapt* the control system accordingly. For such systems (e.g., F15-Active), verification and certification is a central point when such systems are to be used in manned aviation.
- Because traditional V&V processes cannot be applied directly to adaptive systems, we are developing a Software Development and V&V Process Guide for systems with neural network based controllers. We also perform research on advanced methods to ensure that the neural network always stays within its limits.

*Program Synthesis* addresses the need for high-quality and safety-critical software from the other side: instead of having the code being written by humans, and then trying to eliminate the errors using V&V techniques, program synthesis *automatically* generates running code from a compact, high-level specification, thus not introducing most of the errors in the first place. A program synthesis system *automatically* transforms the specification into executable code (e.g., C or C++) using a large body of background knowledge, whereby the correctness of the code is always ensured. Thus, a synthesis system can be seen as a very sophisticated, knowledge-based, correct compiler. In this area, we are developing world-class program synthesis tools in the following application domains:

- *AutoBayes* is a tool for the automatic generation of complex data analysis algorithms (currently up to 2000 lines of C++ code) from compact and high level specifications. Typical data analysis tasks, which can be handled by the system include clustering and classification (e.g., vegetation clustering), filtering (e.g., using Kalman filters), detection of change-points (e.g., when a sensor breaks), and statistical modeling of software reliability. The schema-based synthesis is based upon Bayesian networks.
- The *AutoFilter* tool has been designed to automatically synthesize high-quality code for state-estimation problems (e.g., navigation of aircraft or spacecraft) using Kalman filters. In FY'02, AutoFilter has been used to successively generate major parts of the state estimation code as was used in the space probe DS1. In collaboration with JPL, the

project team used the tool to generate code which could be directly compiled into the existing DS1 code and run on a board-level simulation system.

As mentioned earlier, *certification* of safety-critical software is important (if not required), but extremely burdensome. Our research effort in program synthesis aims at the development of tools which automatically support the certification process. In addition to the generation of highly documented code, we have developed extensions of the Autobayes and AutoFilter systems to automatically verify important *properties* of the generated code. Typical properties can be specific to the programming language (e.g., array bounds, divide-by-zero), or domain specific (e.g., checking on the correct units, or proving the optimality of the Kalman filter equations). Our systems provide these certificates individually for each generated program, thus substantially reducing the remaining certification effort.

- *Synthesis of UML designs*. In this project, we have developed a prototypical tool to generate highly structured and readable UML statechart designs from requirements in the form of scenarios. This tool is embedded in a commercial UML environment (Rhapsody and Rhapsody/RT) and had been used within a successful case study to "re"-implement the weather control logic in the CTAS air-traffic control system.

*Object Oriented Frameworks*. The development of distributed software applications poses considerable challenges for the software designer. The main activity of this task is the development of a framework to simplify the creation of distributed applications, using techniques from aspect-oriented programming. This work is being applied within DARWIN and the NASA-wide Intelligent Synthesis Environment.

The RIACS area of software engineering is performing important and NASA relevant research work. For the future, it is anticipated that RIACS' software engineering area and its importance within NASA will grow. Size, functional complexity, and criticality of software developed within NASA and for NASA will increase tremendously in order to meet ambitious mission goals with tight mission budgets. In order to prevent mission failures that can lead to loss of scientific data or, in the extreme case, loss of human life, the issue of producing high quality, highly reliable code will become more and more important. The RIACS software engineering area is playing a pioneering role in addressing and solving these problems. With world-class basic research and NASA-relevant case studies, the group is developing techniques and tools, which, together with the group members' experience, will make a huge impact on how software is developed, verified and certified at NASA.

**RIACS Software Engineering Staff (as of 30 Sep 02)**

Johann Schumann, Associate Director
Robert Filman
Bernd Fischer
Dimitra Giannakopoulou
Charles Pecheur
Grigore Rosu (until July 2002)
Willem Visser

# II. RIACS Projects During FY2002

NASA Ames has identified three cornerstones of IT research necessary to meet the future challenges of NASA missions:

- Automated Reasoning for Autonomous Systems
- Human-Centered Computing
- High Performance Computing and Networking

RIACS research focuses on all three of these areas, as well as collaborating with NASA scientists to apply IT research to a variety of NASA application domains. Research projects and the overall RIACS program are regularly reviewed by an eminent Science Council. See Section IV.F for a list of current Science Council members.

## *II.A. Automated Reasoning for Autonomous Systems*

Deep space exploration requires significant advances in artificial intelligence to support the needed capabilities for autonomous systems. RIACS scientists and visitors have been collaborating with NASA researchers in a number of areas, as described in the following sections.

## II.A.1. Automated Software Engineering

### II.A.1.a. Task Summary and Overview

The main objective of this task in verification and validation (V&V) is to improve the reliability of software developed at NASA, which in turn will reduce the risk of mission failure. A second (and often equally important) goal is to improve software development by employing techniques from theoretical computer science, specifically Formal Methods, to reduce software development costs.

In order to accomplish these goals we are investigating the use of model checking to automatically analyze either software already in use at NASA, or new systems under development. Specifically we are analyzing autonomous systems currently being developed at NASA Ames and JPL on both the architectural and code level as well as avionics software systems for both space flight and commercial aviation.

Under this task we have been one of the pioneering research groups applying model checking to real software - a field that has been flourishing the last couple of years. Some of our most notable successes have been finding errors in the Remote Agent autonomous system before and after its space flight (as part of NASA's Deep-Space 1 mission) and rediscovering a known problem in the DEOS real-time operating system used in certain business aircraft. In both cases we applied model checking directly to the code for these systems, hence showing the viability of this approach.

### II.A.1.b. RIACS Staff

Dimitra Giannakopoulou
Charles Pecheur
Grigore Rosu (until July 2002)
Johann Schumann
Willem Visser

#### II.A.1.b.i. Visiting Staff

Howard Barringer (Manchester, UK, Professor)
Alessandro Cimatti (IRST, Italy, Research Scientist)
Jamie Cobleigh (UMASS, PhD Student)
Alex Groce (CMU, PhD Student)
Robbie Joehanes (Kansas State, PhD Student)
Sarfraz Khurshid (MIT, PhD Student)
Flavio Lerda (CMU, PhD Student)
Ny Aina Razermera Mamy (U. of Oregon, PhD Student)
Oksana Tkachuk (Kansas State, PhD Student)

### II.A.1.c. Project Description

Several efforts have been undertaken under this task as follows:

**Java PathFinder (JPF)**

JPF is a model checker for Java programs and is based on an in-house Java Virtual Machine (JVM) developed by members of the V&V task. JPF can analyze Java programs for deadlocks, assertion violations, and properties specified in linear-time temporal logic (LTL). JPF was released for worldwide use in February 2001 and since then has been downloaded by approximately 80 institutions. JPF has been used to discover a number of very subtle flaws in Java programs, most notably, it found a timing error in a Java version of the DEOS real-time operating system used in certain business aircraft. Besides the core system this project covers a number of other research areas: predicate abstraction, environment generation, heuristic search algorithms, symbolic evaluations, etc.

**Specification Centered Testing**

The goal of this project (jointly with Mats Heimdahl from the University of Minnesota) is to investigate the use of model checking of requirement specifications in order to do test-case generation. The novelty here is the use of test-coverage criteria as properties for the model checker - the model checker refutes the claim that a certain coverage cannot be achieved by giving a trace that will achieve the coverage. This trace can then be used to generate a test-case that, when used on the full system, will hopefully achieve the required coverage.

**Human Computer Interaction Finding Automation Surprises**

This joint project with Human Factors experts from NASA Ames will investigate the use of JPF to find automation surprises in prototype Java implementations of autopilots. In a preliminary study we have shown how to generate a program to be checked by JPF from an actual autopilot user-interface; furthermore a notation for describing pilot actions and the model the pilot uses to think about the state of the automation (known as the pilot's mental model) was developed. These three components were then combined to find a previously known automation surprise automatically with JPF. In the next stages of the project we will attempt to do analysis of the actual autopilot displays to detect possible problems where the displays are inadequate to show mode changes to the pilot. This will allow autopilot designers to interactively change their designs and run JPF to tell them about potential flaws, without having to run costly simulations with real pilots, etc.

**System Level Verification Technology for Autonomy Software**

Investigation of techniques that enable system-level verification of autonomous software. This is a basic research project that looks into two types of technologies for achieving its goals: compositional incremental verification techniques and software architectures. The main technical issues that must be resolved in this context are:

- How to decompose properties of a system in terms of properties of its components. We are particularly interested in identifying properties expected of components when specific architectural choices are made. In addition to that, the idea is to investigate what architectures facilitate modular (hence more scalable) verification.
- How to constrain the behavior of a component based on its environment (environment generation); this is useful for restricting the state-space of a component when verified in isolation, but it is also typically necessary for proving properties expected of the component in the context of the system.
- How to provide support for correct assume/guarantee style of reasoning. The focus here is to provide flexible assume/guarantee rules for guiding the designer to avoid incomplete proofs, or proofs that erroneously involve circularities.

**Verification of Neural Networks**

The NN V&V task is in charge of developing methods and techniques for the verification and validation of neural-network based controllers and to support the IFCS (Intelligent Flight Control system) V&V group (a collaborative project between Ames, Dryden, ISR and Boeing). The goal of the IFC project is to develop, certify, and test-fly a neural-network based flight control system on a modified F-15. In a first stage of the project, an architecture with a pre-trained neural network has been used.

**V&V of Model-Based Autonomy Software**

The goal of this task is to apply verification and validation (V&V) techniques to autonomous software, and more particularly to model-based autonomy software (MBAS) such as the Livingstone system from Ames. Charles Pecheur is leading this task for the ASE group, in collaboration with the Autonomy group at Ames and Northrop Grumman Corp. One full-time and one half-time researchers are supported by the task.

This task has produced and keeps developing two verification tools:

- MPL2SMV, a compiler that translates Livingstone models so they can be verified using the SMV model checker. This is joint work with Carnegie Mellon University (CMU).
- Livingstone PathFinder (LPF), a verification tool that applies a model checking algorithm to drive the execution of the real Livingstone program, placed in a simulated environment.

These tools were developed as part of a 3-year effort to develop and apply Livingstone V&V tools in the ISPP project at NASA KSC. This project ended in September 2001.

Recently, Charles Pecheur has been investigating the use of model checking to analyze diagnosability, that is, whether a sufficiently precise diagnosis can be achieved given the partial observability of the system.

Charles Pecheur is also involved in a project on Integrated Vehicle Health Maintenance (IVHM) for the 2nd Generation Reusable Launch Vehicle (RLV), headed by Northrop Grumman. He leads a task whose goal is to contribute V&V expertise and technologies for IVHM, including MPL2SMV and LPF.

### II.A.1.d. Accomplishments during FY2002

**JPF and related**

Java PathFinder has been released worldwide, and has already been downloaded by approximately 80 institutions. It has also attracted a number of organizations that have entered into agreement with NASA to help further develop the tool (these include: University of Minnesota, SUNY Stony Brook, Kansas State University, UC Santa Barbara, University of Liverpool, University of Freiburg and Carnegie Mellon University).

The model checker has been extended to allow symbolic execution, which in turn enables the analysis of open system, i.e., systems with unknown inputs. This is the first time that a symbolic execution based system can handle the combination of complex data structures, numeric constraints and concurrency. This symbolic system has been used for the test case generation project with University of Minnesota, where requirements specifications are automatically translated to Java and then analyzed by the symbolic execution based version of JPF to find input that will ensure a coverage measure.

The model checker returns a counterexample trace if one exists; however understanding this trace can often be difficult. We therefore developed a technique whereby errors are "explained". The model checker itself is used to find other similar traces to the counterexample as well as traces that do not cause an error, the differences in these traces are then analyzed to give a succinct representation of what went wrong.

A Java translation of the K9 Mars rover was analyzed during a benchmarking experiment. This example has been the largest attempted to date – it consists of 6 threads and a total of 8000 lines of code. The model checking of the K9 was highly successful and in the area of concurrency errors outperformed all the other technologies (testing, runtime analysis and static analysis) also benchmarked during the experiment.

A collaboration with JPL was started to analyze the Deep Space 1 fault protection unit. The idea was that we hand over the tool and the JPL engineers use it and report on their experiences. The process went relative smoothly and very few problems were encountered. The JPL engineer using JPF quickly found some interesting errors in the code not previously known, and was suitably impressed by the capabilities of the tool.

**System Level Verification Technology for Autonomy Software**

- Demonstrate the impact of compositional techniques on a running project (Rover).
  A case study has been performed at the architecture level of the Rover Executive. We used design diagrams to perform a systematic translation to FSP, the input notation for the LTSA tool. Several interaction issues (e.g. race conditions) decompose well over the system architecture, given the patterns used in the system. As a result, verification was performed successfully at the module level. Moreover, the study identified ways of improving modularity by localizing some synchronization mechanisms between threads within a subcomponent. It was possible to demonstrate several expected problems, and easily experiment with solutions. The process exposed that one planned solution cannot be implemented using Posix condition variables, and that an alternative solution has to be found.

- Extend tools to facilitate assume-guarantee reasoning.
  Defined a framework with novel algorithms that determine whether there exist contexts in which a component satisfies required properties, and if so, generate the weakest assumptions that characterize such contexts. These algorithms can also be used to decompose properties over components of a system. The approach was implemented in the LTSA tool, and was applied successfully to the Rover study.

  We are currently working on a framework for performing assume-guarantee reasoning in an incremental and automatic fashion. We are experimenting with our ideas in the context of both the LTSA and the JavaPathfinder tools.

- Translation of LTL properties to Büchi automata.
  We have optimized and extensively experimented with the translator, and shown it to perform significantly better than the state-of-the art. The translator is being released under the Java Pathfinder license, and is being used by several tools like Bandera, LTSA, and JPF.

**Autonomy Related: V&V of Model-Based Autonomy**

The project with NASA KSC wrapped up in November 2001, with the presentation of project results at the SAS'01 workshop in September and the production of a final project report and a User Guide for MPL2SMV. This project received very enthusiastic feedback from the Program managers at IV&V Center. They particularly appreciated our effort to have the V&V technology used by the practitioners at KSC, and now mention our project as exemplary in fulfilling the program's goals.

The first full version of Livingstone PathFinder (LPF) was delivered and demonstrated at Ames at the end of October 2001. In December 2001, Tony Lindsey (QSS) joined the Autonomy group at Ames to assist Charles Pecheur in the development and extension of LPF. Contributions over this year include regression tests for LPF, improved checking of diagnosed vs. actual state, generation of re-playable traces and integration with the Stanley GUI for Livingstone, interoperability with Java PathFinder (JPF) components, and support for guided search. There was also a significant effort in improving Livingstone's supporting features for LPF (checkpointing), in collaboration with the Livingstone core developers. They are also working on a continuing basis with Livingstone application developers at Ames, who are testing LPF on their models and providing useful feedback on the tool.

In September 2001, Reid Simmons (Carnegie Mellon) updated MPL2SMV to support the new syntax of Livingstone models. Charles Pecheur extended that work considerably, including new extensions to ease the specification of properties to be verified, support for backward translation of traces, and generation of duplicated models and associated properties.

In collaboration with Alessandro Cimatti (IRST, Italy), Charles Pecheur defined a formal model of diagnosis, diagnosability properties, and an approach to their verification using model checking. Alessandro Cimatti visited Ames for a week in April 2002, and they coauthored a paper laying out the foundations of their approach. The paper was presented by Charles Pecheur at the MoChArt workshop in Lyon in July 2002. Initial experiments in verification of diagnosability are underway.

In the first (baseline) period of the project with Northrop Grumman, Charles Pecheur and Stacy Nelson (NelsonConsult) produced three survey reports covering different aspects of V&V of IVHM. In the second (option I) period which started in April 2002, they are now extending and improving the MPL2SMV and LPF tools for integrating them as part of the IVHM design architecture developed by Northrop Grumman. The three reports from the Baseline period, and our work in the project, have been extremely well received by our Northrop partners.

**V&V of Neural Networks**

For V&V of the pre-trained neural-network system, a software development and V&V guide has been developed and published as a NASA contractor report (Authors: Dale Mackall, DFRC, Stacy Nelson, NelsonConsult, Johann Schumann, RIACS). This process guide is based upon a selection of major relevant international and NASA software standards that have been specifically augmented. In this guide, a classical V-shaped process model is extended by activities (testing, design, analysis, documentation) that are specific for the neural-network based controller. Appendices to the guide provide detailed mathematical prerequisites, and discuss numerical issues for pre-trained neural networks (e.g., scaling, accuracy, convergence). The process guide was accepted very well by project managers and the industry.

The second part of the project is concerned with basic research about verification and validation of neural-network based controllers (VeriCoNN). Due to manpower limitations (a second researcher only started in 08/2002), not all envisioned activities could be carried out fully. A simple Matlab implementation of a simulation of a simple, one-dimensional plant with a NN-based controller has been refined. First experiments have also been carried out using the NNet system (developed at Langley and Ames). This simulation has been used to demonstrate the operation of the "monitoring harness" based upon confidence measures. This task provided a grant to A. Kelkar (Iowa State) for research on robustness/stability of NN-based controllers. Results of this project have been presented at the ECS Project Review Meeting 2002. A scientific paper has been accepted for the international workshop on self-healing systems.

### II.A.1.e. Problems encountered

Export control issues hampered the analysis of the JPL fault protection unit. We could have had better results if the tool experts were also allowed to see the code being analyzed in order to give better guidance of how to use the tool.

### II.A.1.f. Publications and Presentations during FY2002

#### II.A.1.f.i. Publications

1. Guillaume Brat and Willem Visser. Combining Static Analysis and Model Checking for Software Analysis. Proceedings of the 16th Automated Software Engineering Conference, San Diego, November 2001.
2. Alex Groce and Willem Visser. Heuristic Model Checking for Java Programs. Proceedings of the 9th International SPIN Workshop on Software Model Checking. Grenoble, France, April 2002
3. Alex Groce (CMU) and Willem Visser. Checking Java Programs using Structural Heuristics. Proceedings of the International Symposium on Software Testing and Analysis. Rome, Italy, July 2002
4. Willem Visser, Klaus Havelund, Guillaume Brat, SeungJoon Park and Flavio Lerda. Model Checking Programs. To appear Automated Software Engineering Journal, 2002.
5. Corina Pasareanu, Matt Dwyer and Willem Visser. Finding Feasible Abstract Counter-examples. To appear Software Tools and Technology Transfer Journal, 2002.
6. John Penix, Willem Visser, SeungJoon Park, Phil Oh, Corina Pasareanu, Cindy Kong, Eric Engstrom, Aaron Larson and Nicholas Weininger. Verifying Time Partitioning in the DEOS Scheduling Kernel. Invited paper for Formal Methods in System Design journal, 2002.
7. Michael Fisher (Liverpool, UK) and Willem Visser. Verification of Autonomous Space-craft Control. Workshop on AI Planning and Scheduling for Autonomy in Space Applications, Manchester, UK, July 2002
8. Klaus Havelund and Willem Visser. Program Model Checking as a New Trend. Invited introductory paper for the special section on the 2000 SPIN Workshop. To appear Software Tools and Technology Transfer Journal, 2002.
9. Charles Pecheur, Reid Simmons, Peter Engrand. Formal Verification of Autonomy Models: From Livingstone to SMV. Extended version of an article presented at First Goddard Workshop on Formal Approaches to Agent-Based Systems, NASA Goddard,

April 5-7, 2000. Submitted for inclusion as a chapter of: Formal Approaches to Agent-Based Systems (provisional title).

10. Charles Pecheur, Alessandro Cimatti. Formal Verification of Diagnosability via Symbolic Model Checking. Workshop on Model Checking and Artificial Intelligence (MoChArt-2002), Lyon, France, July 22/23, 2002.
11. Steven Brown, Charles Pecheur. Model-Based Verification of Diagnostic Systems. Proceedings of JANNAF Joint Meeting, Destin, FL, April 8-12, 2002.
12. Stacy Nelson, Charles Pecheur. NASA processes/methods applicable to IVHM V&V. Project report, 2nd Gen RLV Program, NRA 8-30/TA-5/Northrop Grumman/ARC Task 10. Oct 2001. Available as NASA/CR-2002-211401, April 2002.
13. Stacy Nelson, Charles Pecheur. Methods for V&V of IVHM intelligent systems. Project report, 2nd Gen RLV Program, NRA 8-30/TA-5/Northrop Grumman/ARC Task 10. Jan 2002. Available as NASA/CR-2002-211402, April 2002.
14. Stacy Nelson, Charles Pecheur. Diagnostic Model V&V Plan/Methods for DME.Project report, 2nd Gen RLV Program, NRA 8-30/TA-5/Northrop Grumman/ARC Task 10. Jan 2002. Available as NASA/CR-2002-211403, April 2002
15. Dale Mackall, Stacy Nelson, Johann Schumann. Verification and Validation of Neural Networks for Aerospace systems. NASA Technical Report CR-211409, June 2002.
16. Johann Schumann, Stacy Nelson. Toward V&V of Neural Network Based Controllers. Proceedings of WOSS'02 (Workshop on Self-Healing Systems), to appear, 2002.
17. Giannakopoulou, D. and Havelund, K. "Automata-Based Verification of Temporal Properties on Running Programs", in *Proc. of the 16th IEEE International Conference on Automated Software Engineering (ASE 2001)*. November 2001, San Diego, USA.
18. Giannakopoulou, D., Pasareanu, C., and Barringer, H. "Assumption Generation for Software Component Verification", in *Proc. of the 17th IEEE International Conference on Automated Software Engineering (ASE 2002)*. September 2002, Edinburgh, UK. **Award:** Best paper, received ACM SigSoft distinguished paper award.
19. Giannakopoulou, D. and Lerda, F. "From States to Transitions: Improving translation of LTL formulae to Büchi automata". Accepted at *Proc. of the 22nd IFIP WG 6.1 International Conference on Formal Techniques for Networked and Distributed Systems (FORTE 2002)*. November 2002, Houston, Texas. Springer, Lecture Notes in Computer Science 2529.

### II.A.1.f.ii. Presentations

1. Willem Visser. Java Model Checking. Invited Presentation at the Java Verification Workshop, Portland, Oregon, January 2002.
2. Willem Visser. Analytic Verification for Aerospace. Information Technology Strategic Research Milestone Presentation to Project Manager, Dave Alfano. NASA Ames, January 2002.
3. Willem Visser, John Hatcliff and Matt Dwyer. Tutorial on Bandera and Java PathFinder at The European Joint Conferences on Theory and Practice of Software (ETAPS). April 2002, Grenoble, France.
4. Willem Visser and John Penix. Special presentation to Steve Zornetzer (Director of Information Science and Technology at NASA Ames) on Software model checking research. NASA Ames, April 2002

5. Willem Visser. Tutorial on Software Model Checking at the Automated Software Engineering Conference, September 2002, Edinburgh, Scotland.
6. Stacy Nelson, Michael Whalen, Charles Pecheur. From Research to Industry: The Role of Software Engineering Standards. Tutorial presented at the ASE'02 conference, Edinburgh, UK, September 2002.
7. Charles Pecheur. Formal Verification of Diagnosability via Symbolic Model Checking. Article presented at MoChArt 2002 Workshop, Lyon, France, July 23, 2002 (with Alessandro Cimatti).
8. Charles Pecheur. Symbolic Model Checking of Domain Models for Autonomous Spacecrafts. Presentation at University of Liège, Belgium, 12 November 2001.
9. Charles Pecheur. Symbolic Model Checking of Domain Models for Autonomous Spacecrafts. Presentation at the Dagstuhl Seminar: Exploration of Large State Spaces, Dagstuhl, Germany, 8 November 2001.
10. Charles Pecheur. Verification of Intelligent Software. Invited Presentation at San Jose State University as part of the Information Engineering Seminar Series, 17 October 2001.
11. Johann Schumann, VeriCoNN – verification of neural-network based controllers. ECS project review. NASA Ames, Aug. 2002.
12. Dimitra Giannakopoulou. "Automata-Based Verification of Temporal Properties on Running Programs". Paper presentation at the 16th IEEE International Conference on Automated Software Engineering (ASE 2001), November 2001, San Diego, USA.
13. Dimitra Giannakopoulou. "Modular Verification for Autonomous Systems". IS Program Review Workshop, Monterey, CA, USA. August 2002.
14. Dimitra Giannakopoulou. "Assumption Generation for Software Component Verification". Paper presentation at the 17th IEEE International Conference on Automated Software Engineering (ASE 2002). September 2002, Edinburgh, UK.
15. Dimitra Giannakopoulou. "Assumption Generation for Software Component Verification". Distributed Software Engineering group, Dept. of Computing, Imperial College, London, UK. September 2002.

#### II.A.1.f.iii. Other Activities

- Organization of OOPSLA 2001 Workshop on Specification and Verification of Component-Based Systems (SAVCBS), October 2001. With Prof. Gary Leavens (Iowa State) and Prof. Murali Sitaraman (Clemson). Workshop was very successful and well-attended (approx. 30 participants).

### II.A.1.g. Future Plans

**V&V of Model-Based Autonomy Software**

Livingstone PathFinder will be extended in different ways. First, we will investigate the use of heuristics to guide the simulation, re-using the guided search algorithms developed by Alex Grocefor Java PathFinder during his 2001 SSRP internship. Second, we will adapt LPF to the Titan diagnosis system from MIT. Titan offers a similar diagnosis technology as Livingstone but also provides model-based planning and execution components, thus enlarging the scope of autonomous activities to be covered. In the meantime, we will keep supporting LPF users and fostering its use on real applications.

The translator to SMV will be further improved to support verification of diagnosability, and we will investigate its adaptation to Titan models. The basic principles on verification of diagnosability laid out this year will be assessed on real-size applications, both in terms of scalability and of suitability to application needs.

The IVHM V&V project with Northrop Grumman continues until May 2003, with further tool usability improvements. The outcome will be two demonstrations of the improved tools as part of the quarterly project reviews, in October 2002 and March 2003.

**JPF related**

- We will continue support of the JPF model checker for the growing number of installations around the world. Furthermore, we will also lend support to the Engineering for Complex Systems (ECS) project where JPF is made commercial strength as well as a C/C++ front-end is being developed. This past year the project evolved from a research dominated activity to one where technology infusion was the main goal; we anticipate this trend to continue in this coming year.
- The major research activities this coming year will focus on extending the symbolic analysis capabilities for JPF. A preliminary prototype already exists, but it cannot handle all Java programs. Furthermore, we will focus on the application of the model checker, and specifically the symbolic version, to generate test cases for the joint project with University of Minnesota.
- We will also extend the current work on explaining errors to make it a more seamless addition to the tool's current capabilities.

**V&V of Neural Networks**

The major tasks in this project for FY'03 are concerned with on-line adaptive neural-network based controllers. Here, dynamic changes in the aircraft aerodynamics cause automatic on-line adaptation of the neural network. This kind of architecture poses substantial additional requirements for V&V. In particular, the traditional notion of "verification against specification" is meaningless, because an "unforeseen event" cannot be specified in advance. Here, we (in collaboration with Pramod Gupta of QSS and the NN V&V group) develop techniques which combine strong mathematical results (typically Lyapunov stability analysis, VC-dimension arguments, Bayesian statistics) with advanced techniques for testing and out dynamic monitoring harness. Besides basic research to demonstrate technical feasibility of our research approach, the major effort during FY'03 will spent on a V&V process guide for on-line adaptive NN-based systems.

**System Level Verification Technology for Autonomy Software**

- To work along with the Rover Executive project and verify early new design decisions for added functionality. We also wish to investigate the impact of different design decisions to modular verification. The aim would be to select design-patterns for autonomy that could facilitate verification of such systems.

- To extend our work on assumption/environment generation for modular verification. Our plans include: dealing with liveness/fairness properties, dealing with code level assumptions in the Java PathFinder tool, looking at timed systems in collaboration with Prof. Howard Barringer, improving our algorithms to produce partial results in an incremental setting, and finally experiment with our approach and increase its usability.

## II.A.2. Frameworks for Distributed Computing

### II.A.2.a. Task Summary

NASA has a major problem in coordinating the distribution of and collaboration about scientific and engineering data and information among widely separated scientists and engineers. This appears in activities such as spacecraft design and scientific data dissemination for space operations. Examples of current NASA projects of this kind include the Darwin distributed wind-tunnel data system (http://www-darwin.arc.nasa.gov/docs/), and the Mars Exploratory Rover/Collaborative Information Portal (MER/CIP).

The goal of this task is to research mechanisms to ease the construction of such distributed information systems. This includes developing frameworks for distributed systems, mechanisms for the creation, development, debugging, and evolution of software systems, and tools for the distributed analysis of data. Accomplishments under this activity have included design and implementation of the Object Infrastructure Framework for distributed applications, creation of a theory of aspect-oriented programming (AOP), and the current creation of a generic AOP program transformation system.

### II.A.2.b. RIACS Staff

Robert E. Filman

#### II.A.2.b.i. Visiting Staff and Students

Seminar visitors: Sri Sridharan, Infinisri Resources.

### II.A.2.c. Project Description

NASA has a difficult problem in coordinating the activities and providing data and computational services to widely separated scientists and engineers. The primary approach to simplifying such coordination is to develop distributed software applications that allow NASA employees and contractors to share data and programs. State-of-the-art technology allows building such applications, but such application development is difficult and produces fragile, one-of-a-kind solutions. Over the past year, we have worked on developing technology to simplify the task of creating distributed applications, creating this technology in the contexts of the of the DARWIN remote access to wind-tunnel data, and MER/CIP Mars Exploratory Rover/Collaborative Information Portal. A primary activity in this respect has been to explore Aspect-Oriented Programming, a new software technology that promises to simplify the creation of software systems. Activities have included

- Exploring new architectures for distributed system development, primarily in the direction of Aspect-Oriented Programming Systems.
- Working on defining and building MER/CIP system.
- Disseminating information on distributed computing and software development to NASA, academic organizations and industry.

In prior years, we worked on developing the Object Infrastructure Framework (OIF), a CORBA centered system for achieving "ilities" in distributed systems. Seeking a more flexible environment for exploring the notion of Aspect-Oriented Programming as Program Transformation for events, we started a collaboration with Klaus Havelund of Kestrel Technologies on a system that maps the descriptions of events of interest to transformations over source code that act on the occurrence of those events.

We also have been working with other Code IC staff on the development of the MER/CIP system, have created an algorithm for rapid type checking in inheritance hierarchies, and have begun exploration of the issues of data mining scientific and engineering data.

### II.A.2.d. Accomplishments during FY2002

- Designed and began implementation of an event-based, transformation-driven system for doing Aspect-Oriented Programming
- Organized several workshops and publications. Served on several program committees (see list below).
- Aided the implementation of MER/CIP.
- Created an algorithm for rapid type checking in inheritance hierarchies
- Begun work on understanding data mining of scientific data.

### II.A.2.e. Problems Encountered and Possible Resolution

None

### II.A.2.f. Publications and Presentations during FY2002

#### II.A.2.f.i. Publications


20. Filman, Robert E., Barrett, Stu, Lee, Diana D., and Linden. Ted, "Inserting Ilities by Controlling Communications," *Communications of the ACM, Vol. 45*, No. 1, January, 2002.
http://ic.arc.nasa.gov/~filman/text/oif/cacm-oif.pdf
http://doi.acm.org/10.1145/502269.502274
21. Elrad, Tzilla, Filman, Robert E., and Bader, Atef, "Aspect Oriented Programming," *Communications of the ACM Vol. 44*, No. 10, October 2001, pp. 29–32.
http://ic.arc.nasa.gov/~filman/text/oif/cacm-gei.pdf
http://doi.acm.org/10.1145/383845.383853
22. Filman, Robert E., "Polychotomic Encoding: A Better Quasi-Optimal Bit-Vector Encoding of Tree Hierarchies," Proc. *16th European Conference on Object-Oriented Programming (ECOOP-2002)*, Málaga, Spain, June, 2002

http://eprints.riacs.edu/documents/data/00/00/00/78/index.html
http://ic.arc.nasa.gov/~filman/text/alg/tr-bitvec.pdf

23. Filman, Robert E., and Havelund, Klaus, "Realizing Aspects by Transforming for Events," *Workshop on Declarative Meta Programming to Support Software Development*, Edinburgh, UK, September 2002 .
http://eprints.riacs.edu/documents/data/00/00/00/82/index.html
24. Havelund, Klaus, Goldberg, Allen, Filman, Robert E., and Rosu, Grigore. Program Instrumentation and Trace Analysis, *Radical Innovations of Software and Systems Engineering in the Future*, Venice, October 2002.
25. Filman, Robert E., Injectors and Annotations, *Workshop on Concrete Communication Abstractions Of The Next 701 Distributed Object Systems*, ECOOP'2002, Málaga, Spain, June, 2002 .
http://eprints.riacs.edu/documents/data/00/00/00/80/index.html,
http://perso-info.enst-bretagne.fr/%7Ebeugnard/ecoop/InjectorsAnnotations.pdf,
http://ic.arc.nasa.gov/~filman/text/oif/cca701.pdf
26. Filman, Robert E., and Havelund, Klaus. Source-Code Instrumentation and Quantification of Events. *AOSD 2002 Workshop on Foundations Of Aspect-Oriented Languages (FOAL)*, Twente, Netherlands, April 2002 .
http://eprints.riacs.edu/documents/data/00/00/00/77/index.html
http://ic.arc.nasa.gov/~filman/text/oif/aop-events.pdf
http://www.cs.iastate.edu/~leavens/FOAL/papers-2002/TR.pdf
27. Filman, Robert E., A Bibliography of Aspect-Oriented Software Development. Technical Report02.06,RIACS.
http://eprints.riacs.edu/documents/data/00/00/00/83/index.html,
http://www.aosd.net/TR-aosd-bibliography.pdf
28. Phoha, Vir V, *A Dictionary of Internet Security Terms*, (Dorothy Denning, Li Gong, John McLean, Robert E. Filman, Elisa Bertino, Jeffrey Schiller, Flemming Nielson, Ejiji Okamoto, John Yesberg, Alfred Menezes, Dieter Gollmann, Csilla Farkas, S. S. Iyenger, Asok Ray, Stephen Kent, and Shashi Phoha, Editorial Board), Berlin: Springer-Verlag, June 2002.
29. Aksit, Mehmet, Clarke, Siobhán, Elrad, Tzilla Holzer, and Filman, Robert E. (Editors), *Aspect-Oriented Software Development*. In preparation.

**II.A.2.f.ii. Presentations**

16. Walton, Joan, Filman, Robert E., Lee, Diana D., Mak, Ron, and Patel, Tarang, "The D3 Middleware Architecture," 2002 World Aviation Congress, Phoenix, Nov. 2002, in press
17. Filman, Robert E., "Source-Code Instrumentation and Quantification of Events." Workshop on Collaborative Software Engineering Tools, NASA Ames Research Center, August 2002.
18. Filman, Robert E., "Aspect-Oriented Programming and the Object Infrastructure Framework," Department of Computer Science, Illinois Institute of Technology, Chicago, Illinois, December 2001.
19. Filman, Robert E. "Polychotomic Encoding: A Better Quasi-Optimal Bit-Vector Encoding of Tree Hierarchies," RIACS, NASA Ames Research Center, May 2002.

### II.A.2.f.iii. Activities

- Associate Editor in Chief, IEEE Internet Computing. (Editor in Chief designate).
- Program Committee, International Conference on Aspect-Oriented Software Development 2003.
- Program Committee, Second International Workshop on Unanticipated Software Evolution, 2003.
- Program Committee, 3rd Workshop on Aspect-Oriented Software Development, German Informatics Association, 2003.
- Organizing committee, Workshop on Engineering Context-Aware Object-Oriented Systems and Environments, OOPSLA 2002, Seattle.
- Editorial Board, Journal of Software Maintenance and Evolution.
- Editorial Board, International Journal on Artificial Intelligence Tools.
- Chair, Steering Committee, International Conference on Aspect-Oriented Software Development.
- Program Committee, International Conference on Aspect-Oriented Software Development 2002.
- Organizing Committee, Workshop on the Advanced Separation of Concerns, OOPSLA 2001.
- Organizing Committee, Workshop on Engineering Complex Object-Oriented Systems for Evolution, OOPSLA 2001.
- Program Committee, Workshop on Aspect Oriented Programming for Distributed Computing Systems, 22nd International Conference on Distributed Computing Systems (ICDCS 2002).
- Guest Editor, Special Section on "Aspect-Oriented Programming," *Communications of the . ACM, Vol. 44*, No. 10, October 2001.
- Program Committee, ROBOSPHERE 2002.

### II.A.2.g. Future Plans

- Develop the system for event-based, transformation-driven Aspect-Oriented Programming.
- Continue to work in information dissemination and technology transfer, including editing IEEE Internet Computing, editing (with Mehmet Aksit, Siobhan Clarke, and Tzilla Elrad) a book of readings on Aspect-Oriented Programming, and serving on various program and organizing committees.
- Aid in the design and development of MER/CIP and successor projects.

## II.A.3. Automated Software Synthesis

### II.A.3.a. Task Summary

This research task in automated software engineering is primarily concerned with the development of advanced tools for the automatic generation of reliable, high-quality code from a high-level specification. Software synthesis is based upon formal approaches and thus enables the justified generation of programs. In this task, research focuses on the development of synthesis tools and approaches for several NASA-relevant domains, such as science data-analysis, spacecraft and aircraft state-estimation, and the generation of UML designs.

### II.A.3.b. RIACS Staff

Bernd Fischer
Grigore Rosu (until July 2002)
Johann Schumann

#### II.A.3.b.i. Visiting Staff

Kate Mullen, Bard College (returning SSRP student)
Jyoti Saboo, De Anza Intern

### II.A.3.c. Project Description

Several efforts have been undertaken under this task as follows:

#### II.A.3.c.i. AutoBayes: Automatic Synthesis of Data Analysis Programs

AutoBayes is an automatic program synthesis system for the data analysis domain. It can compile a high-level statistical model describing an analysis problem into a custom algorithm design and then further down into a fully documented, imperative program implementing that design. The system is intended to automate the application of known algorithmic principles in novel contexts; it provides much more flexibility than a fixed code repository and allows the fully automatic implementation of efficient new algorithms. AutoBayes uses a schema-based approach to program synthesis that combines aspects of Bayesian networks, automated theorem proving, symbolic-algebraic computation and generic programming

#### II.A.3.c.ii. AutoFilter: Automatic Synthesis of certifiable Software for State Estimation

AutoFilter is a tool for the automatic generation of software for state estimation using Kalman filters. Given a high-level description of the system (as a set of process and measurements differential equations), the system can automatically perform all the intermediate steps (e.g., linearization, discretization) to finally produce code for the state estimation. This tool (project team Jon Whittle, Johann Schumann, Tom Pressburger) is based on the architecture and symbolic subsystem of AutoBayes.

#### II.A.3.c.iii. Synthesis of UML designs from requirements (PecSee)

The Unified Modeling Language (UML) has become a de-facto standard for the design and development of object oriented software. Work done in this project (with Jonathan Whittle, Jyoti Saboo, and Phil Oh) supports a highly iterative software process by automatic generation of highly structured and readable designs (statecharts) from requirements (sequence diagrams).

### II.A.3.d. Accomplishments during FY2002

#### II.A.3.d.i. AutoBayes Automatic Synthesis of Data Analysis Programs

During FY02, continuing work concentrated on the maturation of the AutoBayes system. The code optimization capabilities developed in FY01 were improved and integrated with the program schemas. The code certification approach developed in FY01 was extended and partially implemented. The core graph reasoning component has been hardened and extended to provide better reasoning for matrices and tensors; this enabled us to automatically derive code for a multinomial version of the principal component analysis which has been published as an original machine learning research contribution only in 1999. The AutoBayes implementation now comprises approx. 50,000 lines of Prolog code; the system's overall TRL is 3-4. The AutoBayes development team consists of Bernd Fischer, Johann Schumann, and Jon Whittle with external collaborators W. Buntine (HIIT, Helsinki/Finland) and A. Gray (CMU, Pittsburgh).

During FY02, we also increased our efforts to apply AutoBayes to "real" science analysis problems. In collaboration with K. Knuth (ARC and GISS), we were able to synthesize and test the first programs for the analysis of images of planetary nebulae.

**II.A.3.d.ii. AutoFilter: Automatic Synthesis of Certifiable Software for State Estimation**

In FY'02 the tool was matured from handling of textbook examples toward the generation of code for realistic problems. During the year, two major case studies have been carried out:

- Modeling and code-generation of a thruster-driven space-station docking mechanism (Ed Wilson, Jon Whittle, Johann Schumann). The system (for which a hardware-based simulation system has been built within Code IC) consists of a number of IR angle sensors and thrusters. Based upon a hand-written abstract process model, an AutoFilter model was defined and Octave/Matlab code was generated. This code was then compared for quality with the original manual implementation of the filter.
- In collaboration with JPL, a major case study was carried out to demonstrate AutoFilter's capabilities. The task was to synthesize substantial parts of the Deep Space 1 state estimation code. The generated C code was required to interface smoothly with the existing, actual DS1 flight software. This code was then tested using a two board hardware simulation system (at JPL and Ames). In comparison with the original DS1 code, the synthesized code behaved exactly as required on the test cases. Project-team: Jon Whittle, Johann Schumann, Tom Pressburger, Harry Balian, JPL)

Each piece of software for a safety-critical application (e.g., in avionics) has to undergo a rigorous certification procedure – nowadays a major cost driver. Within this task, various techniques for certification support for synthesized software has been developed and extended. Besides refinements of the automatic generation of highly documented code (including design documents), research on the certification of properties has been performed: For language-specific properties, results obtained by last year's SSRP student Mike Whalen have been integrated into AutoBayes (see publications below).

Domain-specific certification is a technique designed to find domain-specific errors that would normally take intensive review by a human expert. Conceptually, the knowledge used by a domain-specific certifier has two distinct levels: an *abstract domain* specification and a *programming language* specification. These levels are linked via *symbolic evaluation* that simulates the execution of the program, and *domain-specific safety abstraction* which interprets the calculated values into the abstract domain and defines the safety policy of that particular domain. A program is *domain-specific safe* if and only if each value calculated along its execution path(s) is safe according to the defined safety policy. The two specification levels are independently reusable; e.g., once an abstract domain has been formulated it can be used to certify programs written in various programming languages, and conversely, programs can be certified for various domain-specific safety policies.

### II.A.3.d.iii. Synthesis of UML designs from requirements

During this year, the project team (Jonathan Whittle, Jyoti Saboo and Johann Schumann) mainly worked on a major case study of a central element of the CTAS system (Center Tracon Advisory System). CTAS is being developed at NASA Ames and has won a NASA software of the year award. Based upon English-language specifications of the central weather-management component, a set of scenarios (UML sequence diagrams) was easily constructed. From there, our synthesis tool generated a structured UML statechart for a commercial UML tool named Rhapsody/RT. Code that was generated by this tool was then integrated into the CTAS system and compared with the behavior of the original code to demonstrate the capabilities of our approach.

### II.A.3.e. Problems Encountered and Possible Solutions

ITAR issues substantially hampered task progress, in particular for the state-estimation program synthesis project. For example, for the DS1 experiment, most members of the project team were neither able to review the original DS1 code nor allowed to view the design documentation which contained essential information for setting up the AutoFilter model.

Certain scientific activities were discontinued due to the events of 9/11. J. Schumann organized a NIPS (12/01) workshop on V&V of Neural-network based controllers. However, the events caused restrictions in foreign travel. Thus the workshop (despite vivid initial interest) had to be canceled.

### II.A.3.f. Publications and Presentations during FY2002

#### II.A.3.f.i. Publications

30. B. Fischer, J. Schumann: Automatic Synthesis of Statistical Data Analysis Programs, Proc. NASA Goddard Science Data Processing Workshop, 2/26 – 2/28, 2002.
31. B. Fischer: Automatic Synthesis of Statistical Data Analysis Programs – Position Statement. AAAI Spring Symposium on Logic-Based Program Synthesis, Stanford, CA, 3/25 – 3/27, 2002.
32. M. Whalen, B. Fischer, J. Schumann: AutoBayes/CC - Combining Program Synthesis with Automatic Code Certification (System Description). Proc. CADE-18, Copenhagen, July 2002, Lecture Notes in Artificial Intelligence 2392, pp. 290-294. Springer.
33. M. Whalen, B. Fischer, J. Schumann: Synthesizing Certified Code. Proc. FME 2002, Copenhagen, July 2002, Lecture Notes in Computer Science 2391, pp. 431-450. Springer.
34. B. Fischer: Deduction-Based Software Component Retrieval. In German. Ausgezeichnete Informatikdissertationen 2001, Dorothea Wagner et al. eds., pp. 19-28, Lecture Notes in Informatics D-2, Köllen Verlag 2002

35. B. Fischer, J. Schumann: AutoBayes: A System for Generating Data Analysis Programs from Statistical Models. Journal of Functional Programming, 2002, to appear.
36. Gray, B. Fischer, J. Schumann, W. Buntine. Deriving Statistical Algorithms Automatically: The EM Family and Beyond. Proc. Conference Neural Information Processing Systems (NIPS*2002), to appear.
37. J. Schumann, B. Fischer, M. Whalen, J. Whittle: Certification Support for Automatically Generated Programs. Hawaii International Conference on System Sciences (HICSS-36), to appear.
38. W. Buntine, B. Fischer, A. Gray: Automatic Derivation of the Multinomial PCA Algorithm, submitted for publication.
39. J. Schumann. Automatic Synthesis of Safety-Related Software. AAAI Spring Symposium on Logic-Based Program Synthesis, Stanford, CA, 3/25 – 3/27, 2002.
40. J. Schumann. Automatic safe learning of UML Agents, Position Paper. Workshop Safe Learning Agents, 2002 AAAI Spring Symposium Series, 2002, Stanford, Mar 2002.
41. G. Rosu. Axiomatizability in Inclusive Equational Logic. Mathematical Structures in Computer Science, Volume 12, Number 5, October 2002.
42. G. Rosu. On Implementing Behavioral Rewriting. ACM SIGPLAN workshop on Rule-Based Programming (RULE'02), Pittsburgh, PA, USA, 5 October 2002.
43. G. Rosu and J. Whittle. Towards Certifying Domain-Specific Properties of Synthesized Code. Verification and Computational Logic (VCL'02), Pittsburgh, PA, USA, 5 October 2002.
44. Joseph Goguen, Grigore Rosu. Institution Morphisms. Formal Aspects of Computing, Volume 13, Issue 3-5, pages 274-307, 2002.
45. Grigore Rosu, Jonathan Whittle. Towards Certifying Domain-Specific Properties of Synthesized Code – Extended Abstract. Automated Software Engineering 2002 (ASE'02), Edinburgh, UK, 23-27 September IEEE, 2002.
46. Jose Meseguer, Grigore Rosu. Towards Behavioral Maude: Behavioral Membership Equational Logic -- [invited lecture] Coalgebraic Methods in Computer Science (CMCS'02), Grenoble, France, 6-7 April 2002. Electronic Notes in Theoretical Computer Science, Volume 65, Number 1, 2002
47. Jose Meseguer, Grigore Rosu. A Total Approach to Partial Algebraic Specification. International Colloquium on Automata, Languages, and Programming, (ICALP'02). Lecture Notes in Computer Science, Volume 2380, pages 572-584, 2002.
48. Klaus Havelund, Grigore Rosu. Synthesizing Monitors for Safety Properties -- [EASST best paper award]. Tools and Algorithms for Construction and Analysis of Systems (TACAS'02), Grenoble, France, 8-12 April 2002. Lecture Notes in Computer Science, Volume 2280, pages 342-356, 2002.
49. Michael Lowry, Thomas Pressburger, Grigore Rosu. Certifying Domain-specific Policies. Automated Software Engineering 2001 (ASE'01), Coronado Island, California, 26-29 November IEEE, 2001.
50. Klaus Havelund, Grigore Rosu. Monitoring Programs using Rewriting. Automated Software Engineering 2001 (ASE'01), Coronado Island, California, 26-29 November IEEE, 2001.

51. Klaus Havelund, Scott Johnson, Grigore Rosu. Specification and Error Pattern Based Program Monitoring. European Space Agency Workshop on On-Board Autonomy, Noordwijk, Holland, 17-19 October 2001.
52. Grigore Rosu. Complete Categorical Equational Deduction. Computer Science Logic (CSL'01), Paris, France, 10-13 September 2001 Lecture Notes in Computer Science, Volume 2142, 2001, pages 528-538.

### II.A.3.f.ii. Presentations

20. B. Fischer, Automatic Synthesis of Data Analysis and State Estimation Programs, New Millenium Program Technology Maturation meeting, ARC, 1/29/02.
21. B. Fischer, Automatic Synthesis of Data Analysis and State Estimation Programs, JPL Technology Infusion Group meeting, JPL, 2/12/2002.
22. B. Fischer, AutoBayes: A system for the Automatic Synthesis of Data Analysis Programs from Statistical Models, NASA Goddard Science Data Processing Workshop, Greenbelt, 2/28/2002.
23. B. Fischer, Automatic Synthesis of Data Analysis and State Estimation Programs, NASA/Propulsion Company Working Group Meeting, ARC, 3/12/2002.
24. B. Fischer, Automatic Synthesis of Statistical Data Analysis Programs, AAAI Spring Symposium on Logic-Based Program Synthesis, Stanford, 3/25/2002.
25. B. Fischer, Automatic Synthesis of Statistical Data Analysis Programs, Ames KDD Town Meeting, ARC, 4/16/02.
26. B. Fischer. Towards Automatic Synthesis of Machine Learning Programs, Invited talk, AI/Robotics/Vision Seminar, UC Berkeley, 4/25/2002.
27. B. Fischer. Towards Automatic Synthesis of Statistical Data Analysis Programs, Invited talk, ASTEC Semincar, Uppsala University, Uppsala, Sweden, 6/5/2002.
28. B. Fischer, Deduction-Based Software Component Retrieval, Invited Talk, Colloquium of the GI-Dissertationspreis, Mainz, Germany, 6/25/2002.
29. M. Whalen. Synthesizing certified code. Formal Methods Europe, Copenhagen, Denmark, 7/24/2002.
30. B. Fischer, Automatic Synthesis of Statistical Data Analysis Programs, Invited Talk, Estonian Business School & Institute for Cybernetics, Tallinn, Estonia, 7/29/2002.
31. B. Fischer, Automatic Synthesis of Statistical Data Analysis Programs, Invited Talk, Technical University of Helsinki, Helsinki, Finland, 8/1/2002.
32. B. Fischer. AutoBayes – Automatic Synthesis of Data Analysis Programs. IS PI Meeting, Monterey, 9/6/2002.
33. J. Schumann. Automatic Safe Learning of UML Agents. Invited lecture, Workshop Safe Learning Agents, 2002 AAAI Spring Symposium Series, 2002, Stanford, Mar 2002
34. J. Schumann. Automatic Program Synthesis and Code Certification. Automated Reasoning, Technische Universitaet Muenchen, Germany, 4/2002
35. J. Schumann. Program Synthesis @ the ASE Group. Code IC area meeting, NASA Ames, 01/2002
36. J. Schumann. The Program Synthesis System AutoBayes, CMU west-coast campus, 2002.
37. J. Schumann. Automated Program Synthesis, RIACS Science Council meeting.5/30/2002.

38. J. Schumann. Automated Synthesis of Safety-Related Software. AAAI Spring Symposium on Logic-Based Program Synthesis, Stanford, CA, 3/25/2002
39. J. Schumann, Program Synthesis @ the ASE Group, NASA Ames Code IC Area meeting, 1/17/2002.
40. J. Schumann, Automated Theorem Proving? Automated Theorem Proving!, Invited lecture, Workshop on Problems and Problem Sets (PaPS), FloC'02, Copenhagen, 08/05/2002.
41. J. Schumann, Combining Program Synthesis with Automatic Code Certification, Conference on Automated Deduction (CADE), Copenhagen, Aug. 2002.
42. G. Rosu. ``On Implementing Behavioral Rewriting," ACM SIGPLAN Workshop on Rule-Based Programming (RULE'02), Pittsburgh, PA, USA, 5 October 2002.
43. G. Rosu. ``On Lightweight Formal Methods in System Specification and Verification," IFIP W.G.1.3, 16th International Workshop on Algebraic Development Techniques, Frauenchiemsee, Germany, 23 September 2002.
44. G. Rosu. ``Experiments and Methods in Dynamic Analysis of Programs," Los Alamos National Laboratory, USA, 23 May 2002.
45. G. Rosu. ``Synthesizing Monitors for Safety Properties," Tools and Algorithms for Construction and Analysis of Systems (TACAS'02), Grenoble, France, 10 April 2002. Best software science presentation award winner at ETAPS'02 (offered by EASST).
46. G. Rosu. ``Experiments and Methods in Dynamic Analysis of Programs," University of Illinois at Urbana-Champaign, USA, 28 February 2002.
47. G. Rosu. ``Monitoring Programs using Rewriting," with Klaus Havelund, Automated Software Engineering (ASE'01), San Diego, California, USA. 28 November 2001.
48. G. Rosu. ``Certifying Domain-Specific Policies," with Michael Lowry, Automated Software Engineering (ASE'01), San Diego, California, USA. 27 November 2001.
49. G. Rosu. ``Behavioral Verification with BOBJ and Kumo," with Joseph Goguen and Kai Lin, Automated Software Engineering (ASE'01), San Diego, California, USA. 26 November 2001. Tutorial.
50. G. Rosu. ``Complete Categorical Equational Deduction," Annual Conference on Computer Science Logic (CSL'01), Paris, France, 10-13 September 2001.
51. G. Rosu. ``Interpreting Abstract Interpretations in Membership Equational Logic," Workshop on Rule-Based Programming (RULE'01), Florence, Italy, 4 September 2001.
52. G. Rosu. ``Monitoring Java programs with Java PathExplorer," with Klaus Havelund, Runtime Verification (RV'01), Paris, France, 23 July 2001.

### II.A.3.g. Future Plans

**AutoBayes**. During FY03, the system maturation process will continue. We will add more numerical algorithms as well as support for multivariate distributions; moreover, we will continue the optimization and certification efforts. We will use two application problems to drive the system maturation: (i) automating an in-depth analysis of planetary nebulae images, continuing the collaboration with K. Knuth, and (ii) synthesis of instrument placement programs in support of rover autonomy.

The effort of generating certifiable code will continue and will culminate in a major FY'03 milestone. We also will extend our collaboration with JPL to study more elaborate state estimation models.

The CTAS-UML case study will continue with the aim of maturing our synthesis tool and to develop groundwork for technology infusion.

## II.A.4. Bayesian Inference and Image Analysis

### II.A.4.a. Task Summary

NASA's EOS program, as well as other data collecting government agencies (such as NOAA, USGS, NIMA, etc.) and industry generate vast sets of observational data of the earth. This data is a basic resource that can help answer such basic scientific questions as global warming, changing land use, ocean circulation and so on. In addition this data provides raw material for answering broad policy questions such as crop assessment, forestry planning, urban planning, etc. Also, individual users may find this data helpful in providing information about their particular farm or community. Also, NASA has image processing needs for planetary rovers and for integrating information from multiple images of planetary surfaces. Given this range of users, this raises the question how can NASA best meet the information needs of these users?

Bayesian model-based data integration in principle solves the problem of how to integrate information from multiple sensors. However, there are many practical problems in constructing computationally efficient models, especially in view of the huge amounts of data involved. In the long run, it should be possible to integrate all image data into one global model as the data is received. This updated model can then be used to project particular information required by particular users. It is also very useful for spotting changes, because these are where there is a larger than expected difference between what the model expects and what is observed. The model can also be used to compress the data, since only the differences between expected and observed need to be stored. In the short term, this technology potential can be tested in demonstrations of data integration, such as integrating the images from a planetary rover as it moves around.

This problem of integrating information from different sensors to answer particular questions is a familiar one in remote sensing and other fields, and often is referred to as "data fusion". "Data fusion", as sometimes envisioned, is a fundamental misconception. Data should never be tampered with, let alone "fused". Data is what was observed, and as such cannot be changed after the event. What basic theory says is possible is to construct geophysical models of the surface and atmosphere that "best" predict the observed data. For example, a geophysical ground model would include a point-by-point description of the topography, ground cover types and spectral characteristics, soil types, moisture content, etc. From such a geophysical model, it is possible to project what a particular patch of ground would look like from a particular viewing angle, under particular lighting conditions, with a particular camera etc. That is, given the model, one can probabilistically predict what would be observed (the expected data). The difference between what is actually observed (the data) and the expected data can be used to update the model. Thus the model acts as the central repository of all the real information in the data and is constructed from our prior knowledge of geophysical processes and how radiation interacts with the geophysical system. This central model could loosely be described as the result of "fusing" the data, but it is not itself "data" or a "data product". The Bayesian probabilistic estimation approach not only allows estimation of the most probable model given prior domain knowledge and the data, but it also estimates the uncertainty associated with the model. In particular, if this model uncertainty is high, it means that there is insufficient data/prior knowledge to pin down which model actually applies.

### II.A.4.b. RIACS Staff

Peter Cheeseman
Andre Jalobeanu
Frank Kuehnel
Doron Tal
Esfandiar Bandari (50%)

#### II.A.4.b.i. Visiting staff:

Rene Vidal (Berkeley)
Bob Wang (CMU)

### II.A.4.c. Project Description

The basic theory behind our approach is that of inverse graphics. That is, if we knew what the ground is like, the lighting conditions, and the camera orientation and characteristics, etc., then we could predict what the camera would see (an image). This is the standard computer graphics problem. However, we have the inverse problem--we know what the images are, and we want to find the most probable ground truth (surface) that would have generated them, assuming we know the lighting conditions and camera characteristics. The most important (and difficult) part of this process is recovering the camera orientation and position for each image. To do this, we must register all the images with respect to each other to an accuracy of a small fraction of a pixel; this registration tells us how an image maps onto the ground truth model we are building. Our initial ground model is formed by letting each pixel "vote" on what the corresponding ground position should be depending on how much that ground position contributed to that pixel. This initial ground model is then used to project what each image should be (i.e., predict each pixel value). The differences between the predicted pixel value and the observed value are used to update the ground model until it cannot be further improved. This procedure produces an increase in both spatial resolution and gray-scale resolution (super-resolution). We are now able to generate images from a high-resolution ground model, and then reconstruct the surface from these images at a resolution much higher than the input images. We are currently trying to extend the system so that it can learn the camera parameters and lighting conditions while simultaneously inferring the unknown 3-D surface at super-resolution.

### II.A.4.d. Accomplishments during FY2002

We have been successful in reconstructing unknown surfaces from synthesized low-resolution images of that surface, even though the images are taken from different viewpoints and under different lighting conditions. The surface is reconstructed at much higher resolution than the input images (super-resolution). This is the first time this has been achieved to our knowledge. The results of this super-resolution can be seen on our web page: http://www.arc.nasa.gov/~super-res/

Our initial super-resolution surface reconstruction used artificial images generated from a digital elevation map, and LandSat images of the same area. We used artificial images because this allowed us to compare the true surface with our reconstruction. The reconstructed surface was within a fraction of one percent of the true value. However, since this initial success, we have been attempting surface reconstruction from real images. This required us to estimate both the camera parameters (internal and external) from the images, as well as reconstruct the surface. This requirement lead to considerable research on camera registration, which we recently achieved, and the surface reconstruction using the images and the inferred camera parameters, which we have also recently achieved. This approach is now being extended to use a Kalman filter to infer the location (with uncertainty) of a set of features. This approach borrows heavily from methods used in robotics, and allows the recursive estimation of the camera position with optimal accuracy.

### II.A.4.e. Problems Encountered and Possible Resolution

We encountered problems in obtaining accurate digital images, since commercial cameras "process" the data in unknown ways. This problem was solved by purchasing an industrial inspection camera. During this year, the general scope of the problem we are trying to solve has become apparent, and we require a significant increase in staff if we are to try to solve all the problems associated with robotic vision. The current staff level is sufficient to solve the 3-D surface modeling and simultaneous camera parameter estimation problem, which is the core of the overall computer vision problem.

### II.A.4.f. Publications and Presentations during FY2002

#### II.A.4.f.i. Publications

53. "(Almost) Featureless Stereo - Calibration and Dense 3D Reconstruction Using Whole Image Operations", RIACS Technical Report 01.26, October 2001
54. "High Resolution Surface Geometry and Albedo by Combining Laser Altimetry and Visible Images", R.D. Morris, U. von Toussaint, P. Cheeseman. In Proceedings of the ISPRS Workshop on Land Surface Mapping and Characterization using Laser Altimetry, Annapolis, MD, October 2001
55. "Robust Automatic Feature Detection and Matching Between Multiple Images", RIACS Technical Report 01.27, November 2001
56. "Dramatic Improvements to Feature Based Stereo", R.D. Morris, V.N. Smelyanskiy, F.O. Keuhnel, D.A. Maluf and P. Cheeseman. In Proceedings of the European Conference on Computer Vision, Copenhagen, May 2002

### II.A.4.g. Future Plans

For FY03, the Bayesian computer vision research is focused on using information from real images, using real cameras to build 3-D surface models. This requires the simultaneous estimation of the camera parameters (extrinsic) and the estimation of the surface shape and reflectance properties. We intend to be able to recover the surface shape and reflectances at super resolution for selected areas, as the full super-resolution algorithm is too expensive to be applied globally. This work requires the development of a new renderer that allows for shadows and occlusions as well as the efficient calculation of derivatives for this extended situation.

## II.A.5. Spatial Statistics and Forecasting

### II.A.5.a. Task Summary

The work in this task is focused on statistical prediction. It addresses two main application areas.

1. Issues of uncertainty in biospheric parameters derived from satellite and other observations (so-called "data products"), and the incorporation of these measures of uncertainty into biospheric predictions; and
2. Inference of nonlinear dynamical system models from observational data, and the use of these models to predict failure modes and events.

### II.A.5.b. RIACS Staff

Robin D. Morris

#### II.A.5.b.i. Visiting Staff and Students

Dr Dmitri Luchinskii (Lancaster University, UK)
Monica Mellem, Tufts University

### II.A.5.c. Project Description

The Earth Science community relies extensively on satellite observations as a primary source of data. From these observations, a large number of "data products" are derived. For example, data from the MODIS satellite are used to generate over 30 derived data products, which give information about quantities which are either impossible to measure directly, or infeasible to measure directly over the large areas of interest. These secondary data sources are used to infer the behavior of the ecological system under study.

In this work we study the uncertainty in these derived quantities. A primary derived quantity of interest to the Earth Science community is the Leaf Area Index (LAI), which gives much information about the vegetation growth in a region. Previously we have studied the uncertainty in deriving LAI from AVHRR data. We will extend this study to characterize the uncertainty in the LAI product derived from MODIS data. Once the uncertainty in the LAI is known together with the uncertainties in the other input quantities, they can be taken into account when the LAI etc. are used as inputs to an Ecological Forecasting system in order to determine the distribution over the output quantities implied by the distributions over the inputs. This will allow a measure of confidence to be determined for any predictions made by the model.

Large occasional fluctuations play a key role in a broad range of physical processes from cosmology and mutations in DNA sequences to catastrophic phenomena in nature such as flooding events and large fires. Also they are often responsible for failures of engineering systems and devices. Phenomena related to large fluctuations play a crucial role in physiology, in particular in cardiovascular systems (CVS) they are often responsible for heart failure events.

The main focus of our research is modeling, forecasting and control of large occasional fluctuations. Our approach to the problem is based on the concept of optimal fluctuational paths along which the system moves, with overwhelming probability, when it fluctuates to a given remote state. We apply statistical methods in nonlinear dynamics and catastrophe theory to develop novel statistical inference techniques in which large fluctuations away from the mean values in time-series data are used to infer the system dynamical model. Our approach is different from most of the conventional learning techniques where outliers are excluded from the data. Instead: in our approach the system dynamics can be best learned in the vicinities of the most probable (least improbable) paths that system follows during the large occasional deviations from the mean, and also in the vicinity of the "backward" motion paths - from the remote states toward the mean. By learning the underlying dynamical model we are able to use the correlations between the measured parameters to make an early prediction of the onset of the large fluctuational events as well as the most probable scenario of the system evolution during such events.

Heart failure is a very common disease, with prevalence estimates indicating that it affects 1-2% of the general population. It has long been understood that cardiovascular control is affected through rhythmic activities and in particular through the low-frequency spectral components in blood flow. Studies of their frequency and amplitude variations suggest that the circulatory system can usefully be modeled as a system of coupled, autonomous, nonlinear oscillators. They are believed to correspond respectively to cardiac rhythm, respiration, myogenic activity, neurogenic activity and endothelial signaling. Currently the Nonlinear Dynamics Group at the University of Lancaster is working in collaboration with the Department of Cardiology of the Royal Lancaster Infirmary to parameterize blood-flow data in terms of e.g. the relative spectral amplitudes, frequencies, synchronization indices and coupling constants in the underlying dynamical model of coupled nonlinear oscillators. We are working in collaboration with the Lancaster Nonlinear Group on implementing our statistical learning scheme based on large deviations to identify the dynamical model from the correlations between abnormalities in the blood-flow and respiratory parameters. Our eventual goal is to develop robust prediction algorithms that can forecast the high risk of congestive cardiac failure based on these abnormalities.

### II.A.5.d. Accomplishments during FY2002

A detailed study of the derivation of Leaf Area Index (LAI) and Available Water Capacity (AWC) from AVHRR data was undertaken. This characterized the uncertainty in the LAI and AWC estimates. It was found that the uncertainty in the estimates is significant, but not overwhelming. However, when the distribution of AWC derived from the AVHRR data was compared with that from the STATSGO database, many of the components of the distribution from STATSGO were not present in the distribution derived from the AVHRR data. Discussions are ongoing with the Earth Science partners in the project to determine the cause of these discrepancies and their significance.

The accomplishments for the analysis of the cardiovascular system model include (i) derivation of a normal-mode representation for cardiovascular-respiratory signal analysis that separates the high-frequency and low-frequency components and allows for frequency variability; (ii) the development of a Bayesian inference scheme for the interpretation of CVS data directly in terms of normal mode oscillators; (iii) the development of an initial version of the likelihood equations for the dynamical Bayesian inference in the CVS model composed of van der Pol oscillators; (iv) the analysis of the model reconstruction for the system of 2 van der Pol oscillators driven by white noise; (v) the development of software for inference in periodically driven dynamical systems.

#### II.A.5.e. Problems Encountered and Possible Resolution

Unfortunately it proved impossible to obtain data of the appropriate type and quality to apply the Bayesian learning for nonlinear dynamical systems methods developed in this task to the prediction of forest fire events. For this reason focus was shifted to work on the dynamical system behavior of the cardiovascular system.

#### II.A.5.f. Publications and Presentations during FY2002

##### II.A.5.f.i. Publications

57. Morris, Robin and Smelyanskiy, Vadim, "On the Uncertainty of Satellite-Derived Biospheric Parameters", *MODIS Vegetation Workshop*, Missoula, MT, July 2002. http://www.forestry.umt.edu/ntsg/MODISCon/

##### II.A.5.f.ii. Presentations

53. Morris, Robin, "High Resolution Soil Water from Regional Databases and Satellite Images", Combating Uncertainty with Fusion, Woods Hole, MA, April 2002.
54. Luchinskii, Dmitri, "Noise in the Life Sciences", Lancaster, UK, August 2002.
55. Morris, Robin, "Statistical and Model Based Prediction", IS Program PI Meeting, Monterey, CA, September 2002.

##### II.A.5.f.iii. Activities

- Reviewer for IEEE Trans. PAMI, IEEE Trans. Signal Processing.

#### II.A.5.g. Future Plans

- See Project Description Section

### II.A.6. Model-Based Autonomy

#### II.A.6.a. Task Summary

The Model-Based Autonomy (MBA) task is currently made up of two separate areas of work — integrated vehicle health maintenance (IVHM) and hybrid diagnosis, primarily for planetary rovers. Dr. Bajwa is involved with the IVHM effort, while Dr. Dearden is the P.I. for the hybrid diagnosis work.

PITEX (Propulsion IVHM Technology Experiment) is a collaborative project, with Northrop Grumman Corporation (NGC), in the IVHM Technology Area, under NASA's Space Launch Initiative (SLI) Program. The project has participation from three NASA Centers: Ames Research Center (ARC), Glenn Research Center (GRC) and Kennedy Space Center (KSC). The PITEX effort is a maturation and refinement of the diagnostic technology initiated in the NITEX work, last year, which demonstrated the use of model-based diagnosis of a Main Propulsion System (MPS) of a Reusable Launch Vehicle (RLV) prototype such as the X-34.

The product of PITEX is an integrated software package containing monitor algorithms and the Livingstone inference engine. It also includes a model of the MPS in order to track and diagnose its operation. During this reporting period the model has been augmented to demonstrate diagnosis of the liquid oxygen (LOX) sub-system as well as the rocket propellant (RP-1) subsystem of the MPS. This has been achieved through use of a qualitative model of the components of the system in order to determine when an off-nominal condition occurs and in order to pinpoint the failed component. The diagnostic system was run on flight-like hardware in order to assess its real-time performance.

The work on hybrid diagnosis is aimed at providing on-board diagnosis, situation estimation, and resource estimation capabilities, primarily for planetary rovers. It is funded by both the Intelligent Systems program and by the Mars Exploration program, and it is hoped that the rover on the Mars '09 mission will use the software we develop.

In contrast with the symbolic approach to diagnosis used in Livingstone, the rover diagnosis work is based on *hybrid* system models, in which the continuous nature of a system (in this case, a rover) is explicitly reasoned about to perform the diagnosis. Although this approach makes the diagnosis algorithm more complex, we believe that its greater representational power is necessary to successfully diagnose systems such as the rover in which there is a very complex interaction between the system and an unknown or only partially known environment. In addition, this approach should allow us to estimate future resource usage based on the diagnosis of the rover's internal state and its environment.

### II.A.6.b. RIACS Staff

Anupa Bajwa
Richard Dearden
Serdar Uckun

#### II.A.6.b.i. Visiting Staff and Students

Frank Hutter, Technical University Darmstadt (SSRP summer student 2002)

## II.A.6.c. Project Description

### II.A.6.c.i. Integrated Vehicle Health Maintenance

The main architectural blocks of the PITEX software are a Real-time Unit (RxU) and a Ground-Processing Unit (GPU). The RxU, which runs on a VxWorks simulator, provides real-time processing of telemetry data for tracking vehicle component status and detecting off-nominal conditions. The GPU provides visualization, explanation and archiving of vehicle data and experiment status results. The main components of the RxU are:

- Monitors, which listen in on the (simulated) telemetry stream for command and sensor data in order to detect changes in the system's state
- Real-Time Interface (RTI), which translates this information into a qualitative, discrete format and which decides when to issue a diagnosis request
- Livingstone, which is the inference engine
- Results Output System (ROS), which stores the diagnostic output and the
- Telemetry Input System (TIS), which provides status and event information for display to the ground operator using the GPU.

### II.A.6.c.ii. Hybrid Diagnosis

The overall goal of this research is to provide a diagnosis system that can be applied on-board a planetary rover (specifically the Mars Science Laboratory rover scheduled for launch in 2009), or other complex system. Such a diagnosis system must have the following characteristics:

- It should be hybrid, modelling the system as a mixture of discrete system modes, and continuous values in each of those modes. The system model may be a stochastic hybrid automaton, a hybrid Bayesian network, or another equivalent formulation.
- It should be capable of operating in real-time on-board the rover. This means that computational requirements (both time and memory) should be relatively low, and the system should preferably be anytime.
- It should cope with the types of models we expect to see in representing complex systems such as the rover. This includes non-linear models, multimodal system behaviour, very low probability faults, and high-dimensional state spaces.

Research on all these issues is ongoing. In particular, *particle filtering algorithms* are being used as the basis for the diagnosis system. These have the advantage that they easily operate with hybrid models, and are anytime. Most of the on-going research is in increasing the computational efficiency of the particle filter algorithm so that it can be used on-board, while still being able to cope with non-linear system models, low probability faults, and other characteristics of diagnosis problems that make them particularly difficult to reason about.

## II.A.6.d. Accomplishments during FY2002

The PITEX demonstration included twenty-four scenarios (one nominal operation and twenty-three failure scenarios) that were diagnosed in the MPS during the captive carry phase of flight. The software successfully tracked the state of the system through the nominal operation and it diagnosed the simulated faults for each failure scenario. Dr. Bajwa's main contribution to the project has been in the domain knowledge acquisition and modeling effort, in exercising the system to make it robust, and in demonstrating the scalability of the model.

Another task she has worked on is investigating the interoperability of the model forms between various tools such as Livingstone, eXpress (from DSI International) and BEAM/SHINE (from JPL). The idea is to be able to translate between models used by various tools, by having a common representation stored in a central warehouse.

In the hybrid diagnosis research, the main accomplishments during FY02 were the following:

- Using a model built for the Marsokhod rover, and actual data files from field trials of that rover, a number of algorithms based on particle filters have been developed and tested. We have shown that variants of particle filter algorithms can be used to correctly diagnose faults, and that for linear models with Gaussian noise, this can be done efficiently enough to use on-board the rover.
- We ordered three joints (two shoulder and the elbow) from the K-9 rover's instrument arm to be manufactured. This will provide a testbed on which we can begin experiments on using the diagnosis system for collision detection, and generate data with faults by installing faulty components in the arm. Neither of these tasks can be accomplished with K-9 itself because the rover team will not allow any activities that might damage the rover in any way.
- Using data collected from field trials of the K-9 rover, we built a model of the locomotion system that can be used to detect the slope that the rover is traversing. This model uses only the wheel current and speed information, and can be used to test the performance of the wheels against the rover's inclinometer, allowing us to detect faults in both of these.

During 2002, we also initiated dialogues with JPL and United Space Alliance (USA) for joint work on potential diagnostic applications. The JPL initiative is on diagnostics for subsurface exploration. The JPL point of contact is Brian Wilcox who has been working on a revolutionary low-cost, lightweight, low power drilling technology for subsurface exploration of Mars. The USA initiative is for fault detection and identification of the Space Shuttle payload subsystem. The USA point of contact is Matt Barry, Director of R&D at USA. At the time of writing, both initiatives are in discussion phase.

### II.A.6.e. Problems Encountered and Possible Resolution

The only significant problem encountered was access to the rover hardware for algorithm testing and modeling for the hybrid diagnosis research. At present there is an on-going difficulty with collecting telemetry data, which will hopefully soon be resolved by hardware and software changes on the rover. More seriously there is currently no way to obtain fault data since the rover is used by many different groups and faulty components cannot easily be swapped in and out. This should be partially solved by the construction of a new partial arm, which will be used only for diagnosis work.

## II.A.6.f. Publications and Presentations during FY2002

### II.A.6.f.i. Publications

58. Bajwa, Anupa and Sweet, Adam. "The Livingstone Model of a Main Propulsion System." To appear at the IEEE Aerospace Conference, Big Sky, MT, March 2003.
59. Meyer et al. "Propulsion IVHM Technology Experiment Overview." To appear at the IEEE Aerospace Conference, Big Sky, MT, March 2003.
60. Dearden, Richard, Kurien, Jim and Robinson, Peter. "Livingstone and the Remote Agent Experiment." To appear in Telematik Magazine, 2002.
61. Dearden, Richard and Clancy, Dan. "Particle Filters for Real-Time Fault Detection in Planetary Rovers." Proceedings of the Thirteenth International Workshop on Principles of Diagnosis, May 2002. (An earlier version of this paper appeared at the European Space Agency workshop on On-board Autonomy, Oct 2001).
62. Kumar, T. Satish and Dearden, Richard. "The Oracular Constraints Method." In Abstraction, Reformulation and Approximation. Aug 2002.
63. Bresina, John, Dearden Richard, Meuleau, Nicholas, Ramakrishnan, Sailesh, Smith, David, and Washington, Richard. "Planning under Continuous Time and Resource Uncertainty: A Challenge for AI." In Proceedings of the Eighteenth Conference on Uncertainty in AI, Aug 2002.
64. Dearden, Richard, Meuleau, Nicholas, Ramakrishnan, Sailesh, Smith, David and Washington, Richard. "Contingency Planning for Planetary Rovers." To appear at the 2002 NASA Planning and Scheduling Workshop, Oct 2002.

### II.A.6.f.ii. Presentations

56. Bajwa, Anupa. "Livingstone Overview" and "Propulsion IVHM Technology Experiment" presented in Poster Presentations at *World Space Congress*, Houston, TX, October 2002. (Supported the Outreach Group in Code IC in the creation of the posters)
57. Dearden, Richard. "System-level Autonomy for Robotic Exploration" AI in Space: Unique Challenges and Opportunities. A tutorial at AAAI 2002, Edmonton, AB, Aug 2002. (This talk also presented at a number of universities).

### II.A.6.f.iii. Activities

- Dr. Dearden was a member of the Space Technology 7 Autonomy technology proposal team. Along with Dr. Nicola Muscettola, he proposed the use of the IDEA agent architecture as the overall autonomy architecture of the spacecraft. The proposal was successful and was included in the autonomy mission proposal, but this was not selected for the mission.
- Dr. Dearden is a reviewer for the Journal of Machine Learning Research, the Artificial Intelligence Journal, and the Journal of Artificial Intelligence Research.

### II.A.6.f.iv. Awards

The PITEX team, including Dr. Bajwa, won an Ames Honor Award for excellence in the category of Group/Team, September 2002.

### II.A.6.g. Future Plans

In FY03 the PITEX work will be extended in order to demonstrate how various sub-system health managers can utilize complementary information in order to achieve system-level health management. Model interoperability will also be further explored and demonstrated. It is expected that when a suitable vehicle platform is available, PITEX technology will be demonstrated as an integrated experiment. One possible platform being considered is a reaction control system (RCS).

The PITEX work has also revealed the successes, and limitations, of using qualitative, model-based reasoning for diagnosis of continuous systems. Dr. Bajwa is collaborating on exploring the use an alternate technology, such as BEAM from JPL, for the detection of off-nominal conditions, and is investigating integrating BEAM with Livingstone.

For the hybrid diagnosis work, research continues into the use of particle filters for diagnosis of hybrid systems. This work can be divided into two main areas. The first is modeling and integration. Here the emphasis is on applying the algorithms to the K-9 rover (and potentially other systems), including building detailed models of various rover subsystems, and demonstrating real-time on-board diagnosis. To this end, we intend to examine possible representations, and to build models of components using these representations.

The second part of our approach is continued algorithmic development. We plan to look at ways to extend our efficient algorithms for linear models to work with non-linear systems, and looking at ways in which the structure in the model representation we choose can be exploited to reduce the state space and allow more efficient sampling methods.

We also intend to extend our work to other types of systems, including life-support systems, autonomous drilling operations, or automatic factories such as Mars in-situ propellant production.

## II.A.7. <u>Autonomous Systems for Spacecraft</u>

### II.A.7.a. Task Summary

The need for autonomous decision-making arises in many areas within NASA, including space exploration, Earth observation, rover operations, and air traffic management. While the applications are varied, the same basic problems are encountered in the different areas, namely the need for reasoning about interacting activities that share limited resources. Looking towards space operations as an example, we find that spacecraft, satellites and rovers have traditionally been operated by mission operators that build and test command sequences and then upload the result to be executed on board the spacecraft or rover in question. The drawbacks of this approach are the requirement for manpower to build and test sequences, the need for deep space communications to upload sequences at regular intervals, and the lack of flexibility during execution, enforced by the predetermined sequences. As more and more missions are active at the same time, each with limited resources and manpower, these drawbacks become significant obstacles. Furthermore, extended distant missions such as exploring Europa, and missions with severely time-constrained decision-making windows such as Mars rover explorations cannot be effectively completed using traditional command sequencing methods.

The high-level goal of this task is to continue the development of cutting-edge automated planning, scheduling and execution techniques in support of NASA mission operations. This includes autonomous operations on board spacecraft and rovers, on ground autonomous plan generation and validation as part of mission operations, and mixed-initiative planning and scheduling tools to assist mission operators in decision-making and validation. There are three key components to this task. The first is the development of advanced planning and scheduling techniques, in particular constraint-based methods. The second is the development of a tool, based on constraint-based planning techniques, that will assist Mars Exploration Rover scientists in planning daily science activities of the MER rovers during the mission. The third is the development of the infrastructure to represent and execute plans for rovers.

Constraint-based planning is a paradigm for planning and scheduling operations in complex domains. The Remote Agent, which was flight validated on board the Deep Space One spacecraft in an experiment in May 1999, used an underlying constraint representation and reasoning mechanism. Furthermore, constraint-based planning has recently been identified as one of the most promising solution candidates for providing planning technology capable of handling issues such as time and resources. In constraint-based planning, activities and states are represented as temporal intervals that are connected by constraints. The result is a network of tasks that defines a set of plans. In addition to the increased expressiveness for reasoning about time and resources, the constraint-based reasoning approach allows arbitrary constraints to be represented, and supports almost any constraint-reasoning techniques for improving the efficiency of the reasoning.

The science activity planning process in the Mars Exploration Rover mission is a challenging task. At the end of each sol (Martian day), the team of scientists will be presented with the latest information from the rovers. Then, in preparation for the next sol, they must build a plan for what science-gathering activities are to be performed. This must be accomplished in only a few hours, as the science activity plan must then be turned into command sequences that must be tested and verified before being sent to the rovers. An ongoing effort is underway to develop a mixed-initiative planning tool to assist the scientists in building those plans.

The research on autonomous operation of rovers focuses on robust and flexible plan execution mechanisms, planning and execution under uncertainty, and plan revision to adapt to changing situations. The rover domain presents new and different requirements from free-flying spacecraft in that the agent's behavior cannot be separated from the complex environment in which it operates. The work has involved a close collaboration with other researchers in the Autonomy and Robotics Area of NASA Ames, in particular in the areas of planning and scheduling, robotics, and visualization. In addition, the work has involved collaborations with universities on fundamental research. The goal of the rover autonomy work is to demonstrate, within NASA and the larger scientific community, the vital role that efficient planning and execution technologies play in intelligent control of complex systems such as rovers.

### II.A.7.b. RIACS Staff

James Crawford
Ari K. Jónsson
Kanna Rajan (until 2/1/02)
Rich Washington

### Visiting Staff

Max Horstmann, University of Massachusetts (SSRP summer student 2002)

### II.A.7.c. Project Description

The goal being addressed in this project is to develop, implement, test, and field technology that brings automated decision-making to bear on rover and spacecraft operations at NASA. There are different ways in which automated decision-making can be utilized in NASA applications. One particular scenario that motivates some of the research in this task is to use automated planning in a mixed-initiative scenario to help scientists build plans for planetary rover operations, to then execute the resulting plans on board a rover.

The core of the automated planning capability is the development of advanced constraint-based planning and reasoning techniques. The key goals of this part of the task are to advance the state-of-the-art in constraint-based planning, to improve the efficiency of the constraint-based planning framework being developed, and to make it easier to apply constraint-based planning in NASA applications. The constraint-based planning project is done in collaboration with a number of other scientists and staff at Ames Research Center, including both civil servants (ARC) and contractors (RIACS and QSS), as well as with other scientists outside Ames. A part of this project is also related to another RIACS project. Following is an outline of the collaborations and relations:

- The constraint-based planning technology development, led by Ari Jónsson, is done in collaboration with Tania Bedrax-Weiss (QSS), Will Edgington (QSS), Jeremy Frank (ARC), Conor McGann (QSS), and Paul Morris (ARC).
- The mixed-initiative planning technology development is done in collaboration with John Bresina (ARC) who leads the effort, Adans Ko (JPL) , Pierre Maldague (JPL), and Karen Myers (SRI International).
- The Earth-observing satellite scheduling work is done in collaboration with Robert Morris (ARC), who leads the work, Jeremy Frank (ARC), Lina Khatib (Kestrel), and David Smith (ARC).
- The MER prototype planning tool effort is done in collaboration with Kanna Rajan (RIACS) who leads the overall effort, John Bresina (ARC), Len Charest (JPL), Will Edgington (QSS), Bob Kanefsky (QSS), Adans Ko (JPL), Pierre Maldague (JPL), and Paul Morris (ARC). Ari Jonsson is the development lead for this effort.

The mixed-initiative planning prototype tool for the Mars Exploration Rover mission builds on the constraint-based planning technology, which has, in part, been developed as part of this task. The goal of this project is to develop, implement, and deliver a planning tool to the Mars Exploration Rover mission. The purpose of this tool is to help scientists and engineers to build activity plans for each of the two rovers by providing them with the following capabilities:

- Active enforcement of flight rules
- Selective and global completion of partial plans, satisfying applicable flight rules and resource limitations
- Explanations based on information derived from automated reasoning

Since the tool will be used in the operations of a NASA mission, additional objectives include the reliability of the provided tool, integration with the overall mission operations process, and more.

The research objectives of the rover autonomy subproject fall into two areas:

- Rover autonomy. This work centers around the command executive prototype for the K9 rover, which forms a basis for ongoing research projects as well as serving as the function command executive for rover operations. The research projects have concentrated on robust, flexible execution of plans, using a utility-based approach. More recent work has explored on-board plan adaptation and plan merging techniques, which lie somewhere between execution and full-scale planning. The rover autonomy project collaborates with other research projects at Ames, in particular Limited Contingency Planning (D. Smith, NASA ARC), and state identification and fault diagnosis techniques (R. Dearden, RIACS). The project also maintains an ongoing collaboration with S. Zilberstein at the University of Massachusetts, investigating decision-theoretic approaches for more completely autonomous exploration activities.
- Verification and validation (V&V) for rovers. This work has been conducted in collaboration with the NASA Ames V&V group, primarily D. Giannakopoulou (RIACS) and K. Havelund (QSS), to explore testing of V&V techniques on the rover executive. The eventual goal is to improve the design of the execution architecture while validating V&V techniques on large, operational systems. The rover executive was used as one of the primary testing targets for the V&V 2002 milestone.

### II.A.7.d. Accomplishments during FY2002

The key accomplishments this year were the improved performance of the constraint-based planning system developed by this project, exploration of new techniques in solving constraint-based planning problems, the continued development of the mixed-initiative planning tool prototype for the Mars Exploration Rover missions, field-test of rover execution capabilities, and a new architecture for rover execution that handles the generality of plans developed by constraint-based planning.

#### II.A.7.d.i. Constraint-based planning framework improvements

The work on constraint-based planning has continued steadily. The applicability of the technology and framework to NASA applications has been firmly demonstrated, which has led to the framework being used in a variety of projects. This has resulted in the focus being shifted somewhat to system engineering efforts and away from speculative research, but that is a necessary and useful part of developing and testing autonomy methods for NASA applications. Among the highlights:

- The stability and efficiency of the constraint-based planning system were significantly improved, in particular in terms of testing, documentation, and reasoning efficiency, which are all crucial components of fielding the system in NASA applications.

- The number of projects building on the constraint-based planning system continued to grow, as did the need for collaboration on feature enhancements, support, etc. These projects include: *IDEA*, an execution agent architecture based on using planning to control execution, *SOFIA flight planning*, a research effort to develop a tool to assist in planning observations for the SOFIA airborne observatory, *Earth observation scheduling*, a research project to develop automated planning tools for fleets of satellites, *MER planner prototype*, an effort to demonstrate mixed-initiative planning for science operations scheduling, and spoken dialog interface to automated planning.
- Advances were made in the field of domain analysis for temporal planning problems. In particular, a logical foundation was developed for reasoning about reachability and mutual exclusion in temporal planning problems. The result was published in the AIPS 2002 conference proceedings, and will provide a framework for developing domain-independent heuristics for constraint-based planning methods.

### II.A.7.d.ii. Mixed-initiative science planning tool for MER

Tremendous progress has been made in the development and fielding of a mixed-initiative science planning tool for the Mars Exploration Rover mission. In September 2001, the team demonstrated a prototype tool to a number of mission team members. In October, meetings were held with various MER mission team representatives, culminating in a decision to incorporate the tool into the mission systems. Achievements made since then include:

- Catching up with the development of necessary software requirements, design documents, test plans, etc. Also participated in overall design reviews, re-design meetings, etc. In June, we presented our own requirements and design to the mission.
- Working with members of other mission systems to develop the common interfaces. In particular, worked to develop a suitable representation of activity requests and constraints in a Rover Markup Language to allow scientists to specify constraints in the visual tool used before the planning tool in the operations process.
- Continued development of interface between planning engine and visual activity plan editor (APGEN) to improve reliability and robustness, and to provide the necessary information to the planning engine, in particular resource information.
- Development and implementation of a resource planning method in the planning engine, which uses resource profiles built in APGEN to guide choices in how to complete the plan and schedule activities, so as to provide a complete plan that satisfies resource limitations as well as flight rules.
- Participation in Thread Test C, which was the first test in which the planning tool was used as part of the uplink process. Preparations included additional development and testing, and the first delivery of the planning tool to the mission. During the thread test, team members participated and helped solve issues that came up. While the computational performance of the tool caused serious problems, the team demonstrated that it could resolve issues and deliver the required products downstream in the uplink process.
- Based on the computational issues discovered in Thread Test C, we significantly improved the efficiency of the planner. This effort focused on better methods for scheduling activities, improved constraint reasoning, and better interface for

information sharing between the two parts of the tool. This resulted in a fifty-fold speedup on a benchmark problem.

### II.A.7.d.iii. Exploration of new techniques in solving constraint problems

Driven by the computational expense of general constraint-based planning, two new research efforts were initiated during this year. Both were in part motivated by the Earth-observing satellite problem, in which the goal is to schedule a large collection of satellite image requests on a fleet of satellites. One research path is to explore evolutionary techniques in selecting satellites and times for images. Such techniques use an explicit representation of how choices are to be made, and then work towards improving the choices made by slightly modifying how decisions are made, selecting those approaches that work well and discarding those that perform worse.

The other is to develop methods for identifying planning problems that can be cast as instances of a simpler problem classes. In particular, certain planning problems can be represented as static constraint satisfaction problems, which are much easier to solve than fully general planning problems. However, transforming the problem is no always straightforward; for example, new constraints may have to be derived from the initial problem and proven to represent the same limitations as planning domain axioms.

### II.A.7.d.iv. Rover autonomy

The rover executive was used as the high-level control for a rover engineering field test in October, 2001. The rover demonstrated the ability to make decisions based on scientific criteria, selecting targets autonomously for further investigation.

A preliminary design of a revised executive architecture was completed in June, 2002, with a partial prototype working in September, 2002. The revised architecture supports concurrent activities and a richer set of temporal constraints, to support the output of the Europa planner, as well as basic plan adaptation capabilities. This is working towards an October, 2002 technology demonstration of integrated visualization, planning, execution, and autonomous arm placement.

### II.A.7.e. Problems Encountered and Possible Resolution

None

### II.A.7.f. Publications and Presentations during FY2002

### II.A.7.f.i. Publications

65. Jeremy Frank and Ari K. Jónsson, "Constraint-based Attribute and Interval Planning", to appear in *Journal of Constraints: Special Issue on Constraints and Planning*, Eds: Alexander Nareyek and Subbarao Kambhampati.
66. John Dowding, Jeremy Frank, Beth Ann Hockey, Ari Jonsson, Gregory Aist and James Hieronymus , "A Spoken Dialogue Interface to the Europa Planner", to appear in the *Proceedings of 3rd Int'l NASA Workshop on Planning and Scheduling*
67. Jennifer Dungan, Jeremy Frank, Ari Jonsson, Lina Khatib, Robert Morris, and David Smith, "Rescheduling Observations on Earth-observing Satellites", submitted to iSAIRAS 2003.

68. David Smith, and Ari Jonsson, "The Logic of Reachability", *Proceedings of the Sixth International Conference on Artificial Intelligence Planning and Scheduling (AIPS)*, 2003.
69. J. Bresina, R. Dearden, N,. Meuleau, S. Ramakrishnan, D. Smith, and R. Washington. "Planning under continuous time and resource uncertainty: A challenge for AI," *Proceedings of UAI2002*, 2002.
70. S. Zilberstein, R. Washington, D. S. Bernstein, and A.-I. Mouaddib. "Decision-theoretic control of planetary rovers," Dagstuhl Workshop on Plan-based Control of Robotic Agents, 2001. Will appear in Lecture Notes in AI.
71. J. Bresina, M. Bualat, M. Fair, R. Washington, and A. Wright. "The K9 on-board rover architecture," ESA Workshop on On-Board Autonomy, 2001.

### II.A.7.g. Future Plans

The overall goal of this work is to continue the development of advanced methods for spacecraft autonomy, in particular for rover missions and other NASA applications.

The immediate goals of the Mars Exploration Rover planning tool is to continue the development of the tool, to work with mission engineers and scientists to build a tool that can use autonomous planning methods developed at NASA Ames to help scientists and engineers build better activity plans during time-critical operations. One of the first goals is to participate in Thread Test E, which will take place in December 2002. Specific development goals include improved resource reasoning methods, more flexible plan optimization criteria, explanation facilities to support user decisions, and configuration of automatic flight rule enforcement. Responsibilities to the mission will include deliveries, reviews, participation in system tests, as well as training and fielding for operational readiness tests.

The development of advanced and efficient methods for autonomous planning will focus on two key issues, the support of missions and projects using the technology developed, and the effort to improve the performance of the planning framework. In addition, efforts will continue on the development of automated domain analysis for providing search information, as well as the development of methods to automatically transform planning problems to other representations when computational efficiency can be gained.

The near-term research emphasis in rover autonomy will be on more fully developing capabilities for on-board plan adaptation, followed by integration and evaluation of these techniques. A new activity will be to investigate the interactions of execution with a probabilistic state identification/fault diagnosis component. It turns out that this problem requires on-board response to unpredictable but identifiable failures, which in turn necessitates the development of efficient planning algorithms for this "planning under unpredictability" class of problems. In addition, we will continue extending and using the core rover execution architecture to support and evaluate the research ideas.

## II.A.8. Visual Odometry and Target Tracking

### II.A.8.a. Task Summary and Overview

Visual odometry is a significant task for rover/tool/sensor control and placement. It will allow one to correct for internal navigation units of the robots, in particularly correcting for the drift problems often associated with these hardware. A collaborative effort funded by JPL addressed the problems of feature extraction and matching. These are two fundamental subtasks of visual odometry as well as many other computational vision tasks.

### II.A.8.b. RIACS Staff

Esfandiar Bandari

### II.A.8.c. Project Description

The aim of this two years project is to merge 2D and 3D features from images and stereo. In the first year Ames provided feature extraction routines based on the eigen value analysis of the linearized auto-correlation filters, and an affine registration and matching routine for tracking features and regions between image frames.

One of the main problems in computational vision, specially in stereo motion analysis, is to extract features from the scene and track them accurately over time. These are actually two separate tasks, both of them challenging due to variations in geometry (scale, rotation, skew, occlusion, etc) and changes in imaging radiometry. While 3D features may change their resolution with depth -- and often suffer from noise -- 2D feature tracking can not use geometric invariances often used for 3D tracking.

### II.A.8.d. Accomplishments during FY2002

After researching various algorithms:

- We developed and test fast feature extraction based on linearized auto-correlation analysis.
- We tested the algorithm on a suite of simulated and real data.
- We tested the program on imagery from a moving rover in JPL's Mars yard.
- For matching and registration, we developed an affine tracking algorithm which worked in three separate modes, one augmenting on the previous technique.
    - KLT tracker (Kanade, Lucas and Tomasi). This is simple two degree of freedom tracking, based on image gradients.
    - Six degree of freedom affine tracker, which takes into geometric image distortions.
    - Eight degree of freedom tracker, which also takes into account, radiometric distortions in the form of variations in gain and offset.

### II.A.8.e. Problems Encountered and Possible Resolution

None

### II.A.8.f. Publications and Presentations during FY2002

None

### II.A.8.g. Future Plans
None

## II.A.9. Machine Learning and Knowledge Discovery

### II.A.9.a. Task Summary
The Machine Learning and Knowledge Discovery Group at NASA Ames Research Center has been newly formed to collaborate with domain scientists to answer pressing scientific questions in their fields as well as furthering the fields of machine learning, knowledge discovery, and related areas. The group hopes to build on the significant successes and reputation that previous Machine Learning researchers have had at NASA Ames.

The Group has identified three Strategic Goals to address, each of which encompasses many scientific and technical disciplines. Although goals are broad, they indicate the group's current assessment of the needs of NASA's five Enterprises. The group plans to develop research programs that address these goals assuming that appropriate resources are available. Furthermore, these goals and areas of research are subject to change and refinement as the group integrates with the NASA Enterprises.

### II.A.9.b. RIACS Staff
Ashok Srivastava
Bill Macready

### II.A.9.c. Project Description

**Strategic Goal I: Scientific Data Understanding**
With the proliferation of computers and automated data collection, most scientific disciplines now find themselves awash in data. This fact brings with it new problems but also exciting opportunity if the problems can be satisfactorily addressed. At present in many scientific projects, data is collected at disparate sites, stored in different formats, collected by experts in different disciplines, and measured at different space and time scales and at different resolutions. Moreover, the volume of data is overwhelming. Weaving this data together to form a coherent picture remains a major challenge. Some of these issues are primarily software engineering tasks (i.e. integrating data in different formats at different locations), but major conceptual challenges remain. How are different data sources to be merged together to advance scientific understanding? The goal of this research initiative is to investigate principled methods of building models from radically different data sources incorporating both empirical observations and the results from simulations and numerical computations.

Another major theme of this goal is to understand the richness and complexity of the observed data sources. In order to build models purporting to explain the data, we need to uncover and characterize the structure hidden in the data at all scales, and the manners in which these scales relate to each other. There is reason to believe that characterization of the complexity of data will also be particularly important for terrestrial data sets. Strong evidence suggests that complex ecosystems are healthy ecosystems, "simple" (e.g. periodic) trends in data are often leading indicators for problems.

Finally, we point to the importance of analyzing data in an online or near real-time fashion. If algorithms are developed which are immediately applied as raw data is collected, this should bring to the fore any unexpected discrepancies or abnormalities lurking in the data. Such occurrences can be immediately escalated to the attention of humans who can examine their cause.

Table 1 summarizes some of the anticipated impacts that scientific data understanding can have on the NASA science enterprises.

**Table 1: Enterprise Impact of Scientific Data Understanding**

| Enterprise | Selected Objectives | Impact from Scientific Data Understanding |
|---|---|---|
| Space Science | • understand the structure of the universe; understand the changing sun | • Large datasets comprising numerous sky surveys of varying energies and resolutions can be analyzed by machine learning algorithms to uncover patterns, causal structure, and offer new hypotheses. These same algorithms can also point to apparent irregularities in the data that should be investigated. Moreover, dimension reduction techniques in combination with visualization tools will help scientists come to grips with the volume of data |
| | • share the excitement of space science discoveries with the public | • dimension reduction and visualization methods might allow the public to observe first hand some of the discoveries hidden in the data |
| Earth Science | • Discern and describe how the Earth is changing | • the integration of various data sources each probing different aspects of the earth's surface and atmosphere can be assisted through Bayesian model fusion<br>• measures of complexity in data sources can help assess the health of the earth system and offer the promise to anticipate problems |
| | • Enable the prediction of future changes | • machine learning models constructed from data can be extrapolated to define likely future outcomes |
| | • identify and measure causes of change in the earth system | • causal machine learning algorithms can assist with the inference of causative relationships and the effects of intervention |

Scientific importance: This research direction will expand the bounds of machine learning since the following types of questions will need to be addressed:

- Constructing models of very complex systems based upon very different data sources;
- Integrating empirical data with numerical and simulation results;
- Identifying regions in parameter space where the model might be diverging from experimental data;
- Suggesting measurements to stress the model with the idea of either validating or invalidating features of the model;
- Generating new hypotheses from the data that might warrant exploration from the scientists.

These questions will only become more important over time as computer models become more tightly integrated into scientific methods. Given the emphasis on data and refining models over time it is quite likely that the foundational approach to this problem will be based upon Bayesian methods. However, because of the complexity of the models numerical methods will need to be applied. This research will necessarily push the frontiers of numerical methods in Bayesian model building. Moreover, the complexity of the models and the data will require novel dimension reduction and visualization schemes.

Scientific readiness: Bayesian modeling, simulation, machine learning have matured into very useful tools for the scientist but their integration into a complete toolset that can aide the scientific progress has just begun. Initial efforts will use existing tools from each of the areas and necessarily focus on the integration of these tools into a unified framework. While probably the most important research thrust for the long term, this direction is probably the least mature and offers the greatest opportunities for truly novel contributions.

**Strategic Goal II: Forecasting and Understanding Time-Based Data**

The goal is to develop new methods to forecast time based data on different time scales; to develop methods that give insight into the data generating process; to be able to build comprehensible models for spatial-time series on different time scales and spatial scales; to visualize multivariate spatial-time series; and to efficiently build methods that scale to large data sets for real-world implementations. Table 2 and Table 3 summarize some of the impact that improvements in dealing with time-based data can have on the NASA enterprises.

Time based data is ubiquitous in scientific and business endeavors. Two central problems arise in the analysis of time series: forecasting and understanding. The problem of forecasting is to generate accurate predictions given endogenous and exogenous information about the data generating process. A related problem is to gain insight into the data generating process itself.

Scientific Importance: The ability to understand and make better forecasts of time-based data will have a significant impact on NASA's Enterprises as well as on science at large. For example, the ability to forecast natural disasters such as storms and long-term climate changes based on high-resolution multivariate remote-sensing data, will impact areas such as understanding bio-geophysical processes, the spread of infectious diseases, factors that influence crop growth, and hydrologic forecasting. Each of NASA's Enterprises: Space Science, Earth Science, Biological and Physical Research, Human Exploration and Development of Space and Aerospace Technology, would significantly benefit from a comprehensive and deep capability in forecasting and understanding time based data. For each Enterprise, the table below summarizes a few of the stated objectives of the NASA Enterprise along with the potential impact from this strategic goal.

**Table 2: Enterprise Impact of Forecasting and Understanding Time-Based Data**

| Enterprise | Selected Objectives | Impact from Forecasting and Understanding Time-Based Data |
|---|---|---|
| **Space Science** | · Understand our changing Sun and its effects throughout the Solar System<br>· Improve the reliability of space weather forecasting. | · A large corpus of data regarding Solar events is available that characterizes both spatial and temporal variations on the Sun's surface. The analysis and visualization of such variations may give astrophysicists insight into solar dynamics. Current methods based on ARIMA (linear stochastic models) may be improved through the use of nonlinear models and dynamical models such has Input-Output Hidden Markov Models.<br>· Space weather forecasting requires a significant capability in multivariate time series analysis, with the constraint that the algorithms work in low-power, low-memory, and low CPU environments. These constraints arise in the case where the algorithms are deployed onboard a satellite. The development of methods to fuse multi channel data under these constraints could be a significant contribution to the field. |
| | | |
| **Earth Science** | · Discern and describe how the Earth is changing.<br>· Determine how the Earth system responds to natural and human induced change.<br>· Enable the prediction of future changes in the Earth system.<br>· Demonstrate scientific and technical capabilities to enable the development of practical tools for decision makers. | · Methods to quantify change in the Earth System completely depend on the ability to characterize changes in multivariate spatial time series. Current techniques that are employed average multivariate data over long periods of time in order to reduce atmospheric effects. Furthermore, a significant contribution could be made by building reliable algorithms that extract spatial and temporal trend information to enable analysis of human-induced change.<br>· Causal models based on features derived from such time series may be applicable in determining how the Earth System is changing due to human influence.<br>· The creation of a Decision Support System for Earth Sciences that is comprised of algorithms that can analyze and summarize the vast quantities of Earth Science time-based data that is being created using Remote Sensing technology would be a major enhancement to existing systems. It is currently estimated that Earth Science data alone will increase by a factor of 100 in the coming decade. |

Scientific Readiness: Over the last three decades, the amount of scientific data has grown at an unprecedented rate due to the development of remote sensing technology, developments in high-frequency measurement technology and the tremendous gains in computational technology. Breakthroughs in machine learning techniques that can scale to very large databases, combined with inexpensive computing power, provide the necessary capabilities to address the needs of the science community.

**Table 3: Enterprise Impact of Forecasting and Understanding Time-Based Data (cont.)**

| Enterprise | Selected Objectives | Impact from forecasting and understanding time-based data |
|---|---|---|
| **Biological and Physical Research and Human Exploration of Space** | · Conduct research on biological and physical processes to enable future missions of exploration.<br>· Invest in the development of technologies to enable safe, effective, and reliable human/robotic exploration | · Analysis of the biometric signals for the assessment of human health during space flight is a topic of research that is growing in importance. Time series analysis of biometric data, particularly focusing on EEG, cardio-pulmonary signals, blood-oxygen content, may aid in the development of safe human exploration of space. |
| | | |
| **Aerospace Technology** | · Increase safety | · Increasing safety in Aerospace Technology is an important area of research that includes the areas of Health Monitoring Systems and System Diagnosis. These technologies rely upon the analysis and understanding of multichannel vibration data. Current methods rely on conditioning vibration data and characterizing the vibration signal using the mean, variance, and higher-order moments of the signal. Recent research indicates that the development of new methods that systematically remove known sources of variation in the signals by mathematically modeling the system state may yield an improved understanding of the data generating process and an improved ability to diagnose system health.<br>· The need for increased Aviation Security is of prime concern given the recent terrorist attacks. To address this problem effectively, systems need to be created that can rapidly fuse data from a variety of sources (image, text, voice, and other biometric data) in order to make a passenger threat assessment. Analysis of the passenger records through time is critical to make a valid threat assessment. |

Areas for Research:

- Input-Output Hidden Markov Models
- Power, memory, and CPU constrained time series algorithms
- Extraction of spatial and temporal trends from multivariate spatial time series
- Creation of causal time series models
- Creation of scalable spatial time series models
- Signal decomposition algorithms
- Data fusion
- Classification of highly skewed target classes.

**Strategic Goal III: Applications of Machine Learning to Optimization**

Optimization is the task of searching through vast spaces to uncover solutions to problems having minimal cost. Whether it's scheduling experiments on missions or balancing tradeoffs on complex engineering design problems, efficient optimization is critical to the success of many scientific and business endeavors. All optimization algorithms function by assuming some structure in an objective function (which defines the model to be minimized) and exploiting this structure. The most familiar example is Newton's method which assumes a smooth quadratic fit, and improves a current guess by moving to the minimum of the quadratic. For complex problems this approach is inadequate – usually there are many local optima and Newton's method (when it converges) results in a poor solution. The key to better optimization is the ability to recognize and exploit a greater variety of structures (moving beyond quadratic). Though not widely recognized in the optimization community, machine learning is the ideal tool to identify more complex structures in an objective function. The goal of this research is to develop new efficient optimization algorithms using insights from machine learning, and test their application on real engineering design problems. Table 4 summarizes the some impact that optimization based on machine learning can have on the NASA enterprises.

Scientific importance: The marriage of machine learning and optimization would lead to insights into both optimization and machine learning and will drive new questions in each field. From the optimization perspective new methods will need to be designed to exploit the structures uncovered by the machine learning algorithms. One potential method of exploitation involves sampling at regions inferred to have low cost. This is closely connected to the currently active area of Markov Chain Monte Carlo (MCMC). In MCMC the goal is to sample from high-dimensional multi-modal probability densities and the current problem will offer new insights into how to accomplish this important task. This new application of machine learning will also drive new emphases in machine learning. One important force will be the need for models involving both continuous and discrete elements since almost all real optimization problems involve both types of variables. Moreover, the new machine learning algorithms must operate in an online fashion as they will be embedded within a larger framework. Most importantly, this research direction will force researchers to confront a difficult and important issue – how to properly strike the balance between exploiting what is known with the need to acquire new information. This exploitation/exploration tradeoff is at the frontier of many fields (optimization, reinforcement learning, distributed collectives, active learning, experimental design), and is a central step to the development of truly autonomous algorithms.

Scientific readiness: Despite the difficulty and importance of many of the above challenges the time is ripe to apply machine learning to optimization problems. Progress is expected to be rapid since mature machine learning algorithms are effective at recognizing and extrapolating structure.

Areas for research for Scientific Data Understanding (Strategic Goal I):

- Bayesian model fusion
- Online algorithms for real-time analysis
- Numerical methods for estimating likelihoods from very complex models
- Algorithms for causal inference (needed in order to suggest experiments)
- Integration of numerical and simulation models with empirical data
- Visualization techniques for high dimensional discrete and continuous data
- Empirical measures of complexity constructed from data

**Table 4: Enterprise Impact of Machine Learning for Optimization**

| Enterprise | Selected Objectives | Impact from machine learning for optimization |
|---|---|---|
| **Space Science** | • develop new technologies to enable less expensive research | • the designs of new technologies will need to be optimized to extract maximal scientific benefit at minimal cost. The new optimization methods outlined here will assist in that goal. Current optimization methods are often not applied when the objective functions of large design problems are computationally expensive, the proposed methods make maximal use of each evaluation of the objective |
| **Earth Science** | • develop advanced technologies to enable mission success | • as above, the design of any new technology benefits greatly from automated fine tuning across many dimensions by optimization algorithms. Existing technologies can be optimally tuned by algorithms and radically new technologies can be improved by optimization procedures. This applies equally to sensing technologies as well as delivery technologies. |
| **Aerospace Technologies** | • revolutionize aviation throu increased safety, reduced emissi reduced noise, increased capaci and increased mobility | • meeting the multiple (and competing) objectives defined for aerospace technologies requires sophisticated multi-objective optimization of engineering designs. The development of the methods outlined here naturally and efficiently address multiple and conflicting goals in ways that current methods cannot |

Areas for research for Forecasting and Understanding Time Based Data (Strategic Goal II)

- Input-Output Hidden Markov Models
- Power, memory, and CPU constrained time series algorithms
- Extraction of spatial and temporal trends from multivariate spatial time series
- Creation of causal time series models
- Creation of scalable spatial time series models
- Signal decomposition algorithms
- Data fusion
- Classification of highly skewed target classes.

Areas for research for Applications of Machine Learning to Optimization (Strategic Goal III):

- Machine learning over new domains with particular emphasis on models integrating both discrete and continuous elements
- Balancing of the exploration/exploitation tradeoff; a deeper understanding here also assists in the development of more effective autonomous vehicles with the ability to know when to look for new information sources and when to capitalize on what they already know
- Optimization over multi-level models starting from simple and crude models but computationally inexpensive, and progressing through to highly accurate but expensive models
- Testing on real engineering problems to understand practical constraints, and comparison to existing methods in terms of both solution quality and computational budget
- Integration of hard constraints and constraint programming methods from AI
- Sampling from high-dimensional multi-modal probability distributions (with exploration into the spinoff benefits to Bayesian learning where these methods are critical)

### II.A.9.d. Accomplishments during FY2002

This ideas expressed here are the result of numerous interviews conducted by both authors with members of domain sciences such as Earth Science and Space Science as well as conversations with Machine Learning and Information Technology scientists at NASA Ames Research Center.

Strategic Goal II (Forecasting and Understanding Time-Based Data) was discussed during a Special Interest Group (SIG) organized during the IDU meeting. The SIG was formed to create the basis of a community of multi-disciplinary researchers who are interested in the analysis of time-based data. Nearly 20 researchers from several scientific disciplines participated in the first meeting of the SIG. The strategic goals of the Machine Learning group were discussed, with significant emphasis on the goal of Forecasting and Understanding time based data.

### II.A.9.e. Problems Encountered and Possible Resolution

None

### II.A.9.f. Publications and Presentations during FY2002

#### II.A.9.f.i. Publications

Not applicable: group in formation

#### II.A.9.f.ii. Presentations

58. Presenters Macready, Bill, "Introduction to Machine Learning", *2002 Information Technology and Space Science Workshop*, NASA Ames Research Center, Oct 2002.
59. Macready, Bill, "Operations Research in Data Mining", *2002 SAS Data Mining Conference*, SAS, Raleigh NC, Oct 2002

### II.A.9.f.iii. Other Activities

- Ashok Srivastava, Chair, 2003 SIAM Workshop on Data Mining in Counter Terrorism

### II.A.9.g. Future Plans

The group intends to address the research areas that are described in the Project Description section.

## *II.B. Human-Centered Computing*

Many NASA missions require synergy between humans and computers, with sophisticated computational aids amplifying human cognitive and perceptual abilities. Areas of collaboration between RIACS and NASA scientists are described in the following sections.

### II.B.1. Work Practice Analysis

#### II.B.1.a. Task Summary:

The Work Practice Analysis Task is part of the Human-Centered Computing focus area of Ames Research Center (COE-IT). The focus of the research is on understanding how people and systems are interconnected in practice. To accomplish this, research is conducting in the areas of work systems analysis and evaluation and the development of computational modeling tools for simulating how people collaborate, communicate, and work within their environment.

The research conducted under this task is an integral part of the Work Systems Analysis and Evaluation group in Code IC. The group's research philosophy is based on the view that Human-Centered Computing is a software engineering methodology. This methodology is based on the scientific study of cognition in people and machines, especially understanding the differences between perceptual-motor/cognitive/social aspects of people and present-day computer systems with the objective of developing computer systems that fit human capabilities and practices by exploiting and improving AI programming methods.

#### II.B.1.b. RIACS Staff:

Maarten Sierhuis

##### II.B.1.b.i. Visiting Staff and Students

Alessandro Acquisti (Visiting Ph.D. Student, UC Berkeley)
Julia Brodsky (Visiting scientist, ILC Dover)
Laleh Haghshenass (Visiting MS Student, Stanford University)
Charis Kaskiris (Visiting Ph.D. Student, UC Berkeley)
Cleidson deSouza (SSRP student, UC Irvine)

##### II.B.1.b.ii. Subcontractors

Digital Space Corp.
Metrica
ILC Dover

#### II.B.1.c. Project Description:

The following projects were conducted and/or participated in under this task during FY02:

##### II.B.1.c.i. Human-Robotic Teamwork in Practice

In this project we are studying the work practices of astronauts onboard the International Space Station (ISS), using available data in the form of ISS documentation and training manuals, field trips to Johnson Space Center (JSC), and most of all ISS video data provided to us by the video library at JSC. The funding of this project comes from the NASA Cross-Enterprises and the IS programs. In this project we are collaborating with the Institute for Human and Machine Cognition of the University of West Florida (IHMC), to develop an *agent architecture for autonomous space systems* of the future that perform many tasks involving close to real-time cooperation with people and with other autonomous systems [72]. While these heterogeneous cooperating entities may operate at different levels of sophistication and with dynamically varying degrees of autonomy, they will each require some common means of representing and appropriately participating in joint tasks. Models of human-robotic "teamwork" that will support the use of autonomous systems in operational environments must be grounded in studies of actual work practice and not merely abstract theory.

In this effort, we combine the talents of members of our research team at RIACS and IHMC to develop theory and tools necessary for supporting "design to implementation" approaches for the Personal Satellite Assistant robot developed at NASA ARC that can be generalized to other space systems in a straightforward fashion.

**II.B.1.c.ii. Mobile Agents**

Mobile Agents, a collaborative project with NASA Johnson, industry and universities, is supported in part by the Intelligent Systems Program and managed by principal investigator Dr. William J. Clancey and project director Dr. Maarten Sierhuis. This project will develop a *distributed human-robotic EVA system for surface exploration.* Basic advances in information systems architecture are required to integrate the robotic systems, life support system, pressure garment, communications network and display (inside the helmet), navigation and sample collection tools, and analysis instruments. This project will develop a seamless intelligent system that integrates data from the system's agents (robots, suit, tools, etc.), provides this data on the surface to the rover and hub as well as to remote mission support (with time delay), and uses software "intelligent agents" to interpret this data to provide model-based advice pertinent to carrying out efficient and safe EVAs. The agents will nominally run onboard mobile computer systems, hence they are called "mobile agents." Agents will be implemented in the multiagent simulation system (Brahms), which will be refined and adapted to serve as a real-time system. In addition, a state-of-the-art spoken dialogue interface will be integrated with Brahms models, to support a speech-driven field observation record and rover command system (e.g., "return here later and bring this back to the habitat"). This combination of agents, rover, and model-based spoken dialogue interface constitutes *a rover assistant.*

**II.B.1.c.iii. Brahms Virtual Reality Environment**

The Brahms and Digitalspace team worked together on developing the batch-mode integration between the Brahms simulator and the virtual world system called OWorld. The integration of Brahms and OWorld together is a generic virtual reality engine that will be re-used in many other projects. This virtual reality engine is called Brahms Virtual-reality Engine (VE). One of the future applications is the development of just-in-time virtual reality training environments for space missions in which human trainees participate as avatars in the virtual environment, emerging themselves in the training scenario and interacting with simulated colleagues (i.e., intelligent Brahms bots) and other trainees (i.e., avatars representing other trainees). One of the benefits of the Brahms VE is that it runs in a standard browser on the internet and does not need expensive hardware and complicated software installations. In this way, just-in-time training can be delivered as a distributed environment in which many geographically distributed people (i.e. trainers and trainees) can participate at the same time.

### II.B.1.c.iv. Modeling and Simulation Human-Robot Mission Operations Work Systems

The Brahms team worked on a work system simulation model for the design of the MER '03 mission operations system at JPL. As part of the Ames Human-Centered Computing effort for JPL's MER'03, we are applying the Brahms tool to design the work systems for mission operations for the MER mission. As part of this project we hired a Brahms modeler responsible for modeling the mission operations work system design for the MER mission. Dr. Maarten Sierhuis manages this Brahms modeling effort. The Brahms model is being used by JPL to design the work activities of every mission operations position and the information flow requirements, as well as work spaces. The funding for this project comes from the ARC MER HCC funding.

### II.B.1.d. Accomplishments during FY02:

- *Human-Robotic Teamwork in Practice.* Funded under the IS HCC program, we continued our work practice modeling of the astronauts on the International Space Station. Our research tries to understand how well the planned ISS activities and their written procedures fit the reality of onboard life, and more specifically, to determine the work practices that have evolved on the ISS since Expedition 1 [73].

**Table 5: Types of Activities Based on Regularity and Scheduling**

| | Scheduled activity | Unscheduled activity |
|---|---|---|
| **Day-specific activity** | Maintenance activities (e.g., Replacement of urine-receptacle in Toilet)<br>Experiments (e.g., LAB PL Status/Monitor) | Emergencies<br>Job-Jar activities<br>Unexpected maintenance or repair activities |
| **Recurrent activity** | Physical exercise.<br>Daily Planning Conference.<br>Eating (lunch, dinner, breakfast) | Going to the toilet<br>Sending personal email |

To deal with unexpected events and the realities of onboard life, we categorize activities according to the degree to which the activity was scheduled (scheduled vs. unscheduled activities) and the uniqueness or repeatability (day-specific vs. recurrent activities) of the activity (see Table 5). The two activity types, recurrent and day-specific, are represented differently in Brahms.

We generalized the Brahms ISS work practice model to allow for the simulation of *any* ISS daily work plan provided to the crew as a Form-24 document [74].

**Table 6: Brahms Source Code of the *Replacementurinereceptable* Activity**

```
Object form_24_for_May_7_2001_Expedition_2 instance of DailySchedule {
 initial_beliefs:
 [...]
        (ReplacementUrineReceptable.hour_start = 9);
        (ReplacementUrineReceptable.minute_start = 50);
        (ReplacementUrineReceptable.hour_end = 10);
        (ReplacementUrineReceptable.minute_end = 20);
        (ReplacementUrineReceptable.duration = 1800); // in seconds
        (ReplacementUrineReceptable by_whom YuriUsachev);
        (ReplacementUrineReceptable.Cmd_next_activity =
ECLSSMaintenance);
 [...]
```

The Brahms agents representing the crew can execute any task on a Form-24 object by using ISS procedures represented in the model. We call this plan execution approach *Multiagent Plan Execution in Practice.*

**Table 7: Brahms Representation of Urine Collection Procedure**

```
object UrineCollectionTankReplacementProcedure instanceof Procedure {
 initial_beliefs:
        (current.main_activity = ReplaceUrinalReceptacle);
        (current.main_activity_type = composite_act);
        (current has_sub_activity PrepareTools);
        (current has_sub_activity UnstowEmptyContainer);
        (current has_sub_activity CapOnUsedUrineContainer);
[...]
        (ReplaceUrinalReceptacle first_sub_activity PrepareTools);
        (PrepareTools.type_act = primitive_act);
        (PrepareTools object_needed Screwdriver1);
        (PrepareTools object_needed 10mmAllenWrench);
        (PrepareTools object_needed 14mmWrench);
        (PrepareTools.where_performed = ToiletArea);
        (PrepareTools next_sub_activity UnstowEmptyContainer);
[...]
} // UrineCollectionTankReplacementProcedure
```

- *Mobile Agents.* This year the Mobile Agents project integrated several advanced computing and communications components with a spacesuit and a robotic rover (developed at JSC). These components include a wireless "biovest" that transmits the astronaut's physiological signs and location (developed by Stanford University's National Biocomputation Center), a "voice commanding" interface based on speech recognition (developed by Code-IC's RIALIST group) "software agents" that monitor and facilitate interactions between people and robots (implemented in with Brahms). Detailed models of these various components, as well as people and their work practices, are integrated into a "workflow" system using the Brahms "multi-agent" programming software.

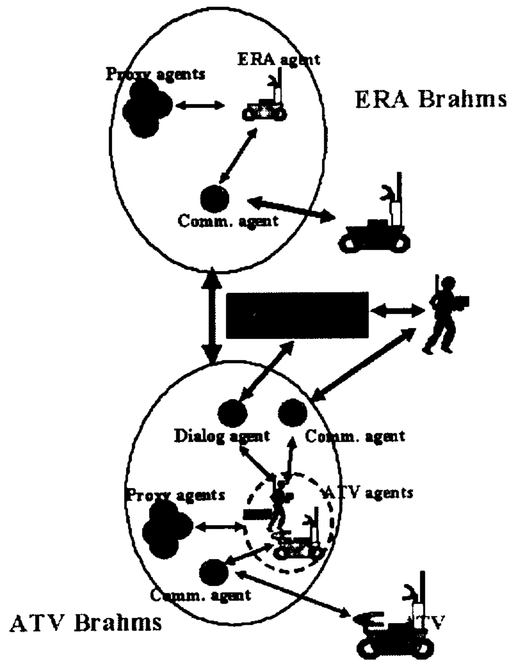


**Figure 2: Mobile Agents Architecture**

This year's accomplishments were tested in two separate field tests:

1. During the week of May 20-24, 2002 we successfully completed our second systems integration test in a week-long field test at NASA JSC in Houston, TX. This field test was in succession of our successful Ames field test from end of April. This time we included all the components tested at Ames, and added the actual ERA Robot and two separate space suits, completing the actual setup as we will test in Flagstaff, AZ in September 2002. The complexity of this field test lies in the integration of the many hardware and software systems, developed this past year, by seven different groups.

   In this field-test a "level zero" scenario proved an integration of the RIALIST speech dialogue system and a distributed Brahms multiagent runtime environment running simultaneously on the MEX computer onboard the ATV and the JSC ERA robot. The scenario was tested at JSC's Mars Yard with two separate space suits (the unpressurized space suit from ILC Dover and the pressurized JSC Mark III space suit) using a wireless microphone system connecting the space suits to the MEX computer on the ATV. The Brahms agents running on the MEX computer communicate to the Brahms ERA robot agent running onboard the ERA Robot, using the MEX wireless D-link network integrating the ERA autonomous control software running onboard the robot. The scenario showed remote commanding of the ERA by the EVA astronaut, using human speech, supported by intelligent software agents.

2. During two weeks in September (Sept. 1-13) scientists, software engineers and researchers, spacesuit and robotics engineers and communications experts trekked into the desert near Flagstaff, Ariz., to study how robots and humans can best interact using spoken language, and to gather data for comparing human and robotic performance. The goal of this field test was to verify the Mobile Agent Architecture's design conception,

implementation and integration with other agent systems in a real field environment. The tests over the two weeks included humans wearing an advanced Mark III spacesuit working alongside the ERA rover (both from JSC), humans wearing a biovest from Stanford University's National Biocomputation Center with physiological sensors, including EKG and respiration, worn by a person inside the space suit and transmitted wirelessly to a Personal Data Assistant (PDA). The speech dialogue system and Brahms were again hosted on the Mobile Exploration System (MEX), which provided computing and wireless communications on a rugged All Terrain Vehicle capable of remote field deployment. A multitude of functions, available to the human astronaut through the integration of physical systems and Brahms agents, were tested using several pre-scripted scenarios in a dessert environment:

a. Start and Stop gathering and storing biophysical data wirelessly from the human during an EVA through spoken dialogue. This was tested with and without the human wearing the Mark III space suit.
b. Start and stop tracking the astronaut's and ERA's GPS location data through spoken dialogue, for recording EVA paths.
c. Providing real-time location information through spoken dialogue.
d. Create places (geographical location objects) in real-time through spoken dialogue, associating given location names with GPS coordinates for recording visited places during the EVA.
e. Create sample bag objects in real-time, for recording when and where samples have been collected.
f. Create voice annotations using the speech dialogue system for recording voice notes for later use.
g. Have ERA take photos of the tracked human using its cameras.

- *Brahms Virtual-reality Environment.* Funded under the IS HCC program to continue the work on the Brahms VE that was started in 2000 with a Phase 1 STTR, together with Digitalspace as subcontractor, Dr. Sierhuis managed the development of a virtual world (VW) model of the Flashline Research Station. The VW model shows a dynamic simulation in Adobe Atmosphere, driven by a detailed multi-agent Brahms simulation model of three real-life scenarios from the Haughton Mars Project 2001 field season. During this field season Dr. W.J. Clancey (HCC Chief-Scientist in Code-IC) participated as a HCC researcher and one of the six crew-members living for two continuous weeks in the Mars Societies' FMARS habitat on Devon Island, in the high Arctic. The Brahms simulation model is based on actual video data from Dr. Clancey's ethnographical study of work-life on Mars, and consists of three scenarios:
  1. A planning meeting by the six crew-members in the hab.
  2. A collaborative teamwork activity in which the six crew-members collaborate together in filling the water tank of the FMARS habitat.
  3. A collaborative teamwork activity in which a crew-member assists in the EVA preparation for donning a space suit.

  The FMARS Virtual World was developed by Digitalspace, Inc. from video data provided by Dr. Clancey.

- *Modeling and Simulation Human-Robot Mission Operations Work Systems.* Based on the results from the Victoria case study, the Brahms team was asked to model the work process of the MER mission to analyze the design of the mission operations work system [61]. Initially, the analysis reports generated from a Brahms model was used to generate several analysis reports from the simulation that are being used in two ways: to develop procedures for the different organizational roles and activities and to verify these procedures during their intermediate operations field tests at JPL.

  Current tools create static representations of people, procedures and events, and do not take into consideration dynamic interactions among different elements of the work system. An integrated view of the relationship of the ground system design to the space system design is also not possible with current tools.

  A more complete representation of work system elements, simulating the interactions between people, objects and geography in a complete work system, that includes a model of the rover, has the potential to assist mission operations system designers in creating designs that are safer and more efficient. The approach is based on socio-technical systems approach from the sixties and seventies, as well as the participatory design approaches from the eighties [62].

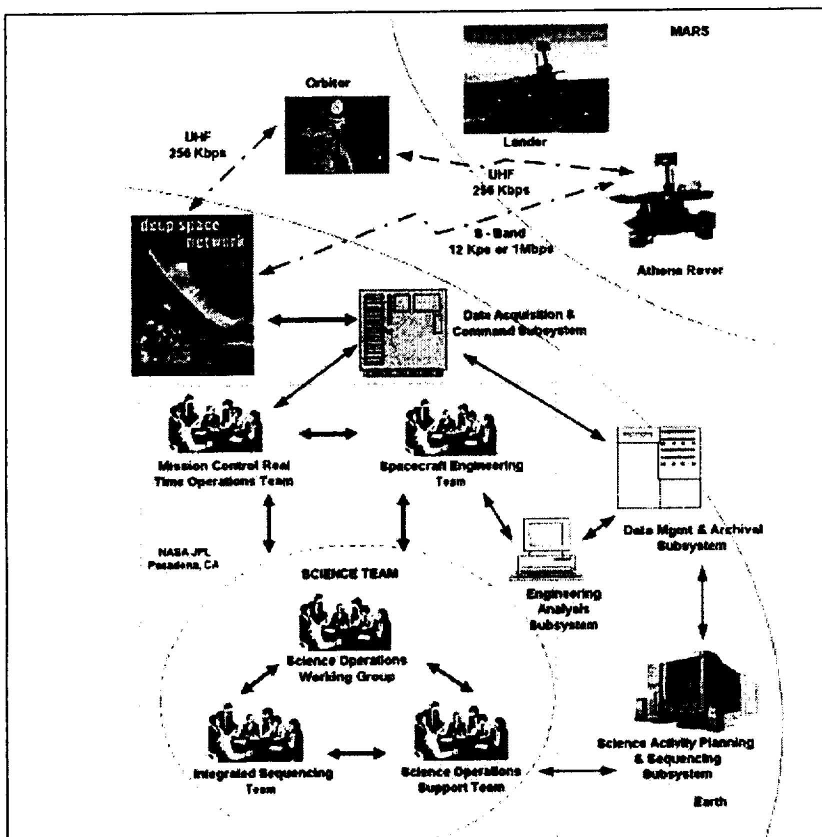


**Figure 3: MER Mission Operations Work System**

### II.B.1.e. Problems Encountered and Possible Resolution

None

### II.B.1.f. Publications and Presentations during FY2002:

#### II.B.1.f.i. Publications:

72. J. M. Bradshaw, M.Sierhuis, A. Acquist, Y.Gawdiak, R. Jeffers, N. Suri, and M. Greaves, "Adjustable Autonomy and Teamwork for the Personal Satellite Assistant," in Agent Autonomy, H. Hexmoor and R. Flacone, Eds.: Kluwer, 2002.
73. M. Sierhuis and W. J. Clancey, "Modeling and Simulating Work Practice: A human-centered method for work systems design," IEEE Intelligent Systems: Special issue on Human-Centered Computing at NASA, in press.
74. M. Sierhuis, "Modeling and Simulating Work Practice; Brahms: A multiagent modeling and simulation language for work system analysis and design," Ph.D. Dissertation. Social Science and Informatics, University of Amsterdam, The Netherlands: University of Amsterdam, 2001, pp. 350.
    *Dissertation book sales:* Over the past year RIACS has sold six copies of Dr. Sierhuis' published dissertation book "Modeling and Simulating Work Practice," outside of NASA [74].

#### II.B.1.f.ii. Presentations

60. A. Acquisti, M. Sierhuis, W. J. Clancey, and J. M. Bradshaw, "Agent Based Modeling of Collaboration and Work Practices Onboard the International Space Station," presented at 11th Computer-Generated Forces and Behavior Representation Conference, Orlando, Fl., 2002.
61. M. Sierhuis, A. Acquisti, and W. J. Clancey, "Multiagent Plan Execution and Work Practice: Modeling plans and practices onboard the ISS," presented at 3rd International NASA Workshop on Planning and Scheduling for Space, Houston, TX, 2002.
62. M. Sierhuis, W. J. Clancey, and M. H. Sims, "Multiagent Modeling and Simulation in Human-Robot Mission Operations Work System Design," presented at Hawai'I International Conference on System Sciences (HICSS-35), Big Island, Hawaii, 2002.
63. A. M. Selvin, C. Palus, and M. Sierhuis, "Knowledge Art: Integrating Compendium and Visual Explorer Methodologies to Explore Creative Sense Making," presented at Knowledge for Creative Decision-Making, Special Focus Symposium, InterSymp, Baden-Baden, Germany, 2002

#### II.B.1.f.iii. Other Activities

*Funded FY'02 Research Projects:*

- "Multi-Agent Software Design and Engineering for Human-Centered Collaborative Autonomous Space Systems". Funding awarded by NRA 99-OS-05 - Cross Enterprise Technology Development Program.

- "Teamwork in Practice: Design for Collaboration in Mixed Human-Robotic Teams". Funding awarded by NRA 2-37143 - Intelligent Systems Program in the area of Human-Centered Computing.
- "Work Practice Simulation Environment for Habitat Design and Scheduling". Funding awarded by NRA 2-37143 - Intelligent Systems Program in the area of Human-Centered Computing
- "Mobile Agent Architecture". Funding awarded by NRA 2-37143 - Intelligent Systems Program in the area of Human-Centered Computing
- 'Brahms VE: A Collaborative Virtual Environment for Mission Operations, Planning and Scheduling", Funding awarded under the "Work Practice Simulation Environment for Habitat Design and Scheduling" project by NRA 2-37143 - Intelligent Systems Program in the area of Human-Centered Computing

### II.B.1.g. Future Plans
None.

## II.B.2. Research in Advanced Language Interfaces and Speech Technology

### II.B.2.a. Task Summary
The goal of the RIALIST (Research In Advanced Language, Interfaces and Speech Technology) group is to conduct leading edge research on advanced dialogue capabilities for multi-modal dialogue systems and apply these results to applications in NASA programs. Our focus is on Spoken Dialogue Interfaces to semi-autonomous agents and training systems. There are many different types of agents being developed by NASA giving a rich variety of possibilities for experimentation. The research builds on extending and augmenting established speech and language technology such as the Nuance recognizer, state of the art speech synthesizers such as ATT Natural Voices, the SRI Gemini NL system and Open Agent Architecture systems.

The current research focuses on contextual interpretation, portability, asynchronous dialogue management, and natural dialogue designs. Present implementations include Procedure Assistant for use by astronauts on the International Space Station, a spoken dialogue interface to the Brahms multi-agent modeling system and to the Europa planner. Future research topics include prosody focus for system spoken output, dialogue move based dialogue management, language modeling based on sparse training data, rational integration of multiple knowledge sources, and evaluation of confirmation methods for collaborative dialogues.

Our group is actively exploring ways to integrate spoken dialogue interfaces into current and future NASA missions. Areas of interest include spoken dialogue interaction with robotic agents including planetary rovers, dialogue based training, augmentation of air traffic control, multi-media dialogue interfaces including eye tracking, and intelligent cockpit instrumentation.

NASA has a major issue on longer missions including the International Space Station and manned Mars exploration of proving ground assistance to the astronauts and rovers. The old models of 24/7 staff availability from Mission Control are being stretched beyond human capability by these long missions. Clearly a way to provide assistance from the ground which involves fewer people is to use computer systems with spoken dialogue interfaces to provide information and expert advice. These systems need to be so easy and natural to use that special training will not be required. The systems need to have the dialogue behavior of humans built into them. Our approach is based on empirical studies of what people say to each other, what people say to actual systems and what they say to simulations of possible future systems.

Speech recognition and natural language processing have been developed over the past ten years so that commercial dialogue systems are used providing airline information, stock quotes and directory assistance. However, NASA's needs in this area differ from the commercial sector in at least two ways:

- NASA requires interfaces for many quite diverse applications. The number and diversity of NASA applications drives the need for research in portability, that is the ability to use and /or adapt language resources and system components to multiple applications.

- Many NASA applications such as collaborative rovers will require dialogue capabilities that are likely to be beyond the state of the art for commercial systems for the foreseeable future. NASA systems need to be robust and accurate, accept any well-formed sentence and operate over special domains.

Part of our approach to addressing NASA's needs in the area of spoken language dialogue interfaces is to leverage state of the art commercial and research software and extend these systems in the areas of portability and advanced capabilities. Our relationships with many of the leading edge R&D groups allow us to use hidden features of the commercial systems to NASA's advantage. The commercial companies have hundreds of programmers who work on making their speech recognition and synthesis faster, more robust and capable of recognizing and generating natural speech. Fast and robust recognition and natural sounding speech synthesis are extremely important attributes for spoken dialogue systems, if they are to be useful as the primary interfaces to complex systems.

### II.B.2.b. RIACS Staff

Jim Hieronymus (IPA since 2/1/02)
John Dowding
Gregory Aist
Beth Ann Hockey

#### II.B.2.b.i. Visiting Staff and Students

Manny Rayner (senior consultant)
Ellen Campana, (SSRP)
Nate Blaylock (SSRP)
Dan Bohus (SSRP)
Genevieve Gorrell (student consultant)

### II.B.2.c. Project Description

This describes the specific research being undertaken under this task. This often is a bulletized list of specific research objectives.

Innovative spoken dialogue interfaces are one area in which research developments can contribute to increased efficiency, robustness and safety of these missions. An important part of creating performance enhancing systems is to have interfaces that do not in themselves create additional work or cognitive burden. Imagine how much easier it would be to use a computer system that you could simply converse with in the same way you would when delegating to a human assistant or as you would when collaborating with a colleague using speech and other modalities. Especially good payoffs for using spoken interfaces are:

- Augmenting tasks already using voice -- capturing information from conversations between pilots and air traffic controllers.
- Allowing collaboration with semi-autonomous agents that are engaged in activities that are normally accomplished using language -- astronauts in space suits interacting with groups of semi-autonomous rovers and astronauts.
- Allowing computer interaction in challenging environments or when hands are impeded by space suits or pressurization -- data entry, command and control.
- Completing complex, high workload tasks -- controlling complex computer systems where it is possible that one voice command can accomplish the equivalent of many keystrokes or mouse moves and clicks.

At the heart of the work is the question of understanding and modeling how people communicate, of exploring what representations can facilitate that communication either with other humans or with computer and robotic systems. Our approach is based on empirical study of what people say to each other, what people say to actual systems and what they say to simulations of possible future systems. We believe that one of the benefits of a spoken dialogue interface is that the closer it is to handling the types of conversations that people would have with each other, the more the interface provides a reduction in required training and in cognitive load.

Future NASA missions require these improved methods of human-computer interaction based on spoken natural language dialogues. Astronauts building the International Space Station or on planetary surfaces will need to be able to carry out many tasks without using their hands to interact with computers and robotic systems. Spoken dialogue systems are the natural way to provide this much needed interface to external devices.

One of our projects this year has been assessing JSC mission needs which could be fulfilled by spoken dialogue technology. We examined both Mission Control and ISS crew operations for the most productive initial areas for spoken dialogue applications. Through close consultation with JSC personnel we concluded that one of the most promising areas was in assisting ISS crew with executing complex procedures. Assistive spoken dialogue systems can free a crew member from some tasks such as reading a procedure to another astronaut who is performing it. The Procedure assistant project is our effort at filling this need on the International Space Station.

NASA has also been tasked with helping to improve air traffic control and cockpit efficiency. The air traffic controllers need to maintain visual contact with and attend to planes and tracking instruments, and should not be distracted by the demands of providing data to the air traffic control system. Since all of the relevant information is spoken already, a spoken dialogue system which performs helpful listening could provide the necessary data. Increasingly complex aircraft cockpits can provide much needed information to pilots, especially in emergency situations. However the user interfaces often require hundreds of keystrokes to obtain this data. An advanced dialogue interface can provide a natural way to ask for the information, and receive it in a timely fashion.

The Mobile Agents Project addresses issues in planetary exploration with teams of astronauts, scientists and robots. RIALIST has participated in this task by providing a spoken dialogue interface to the Brahms system. This system allows the astronaut to ask questions about where he or she is located, and to issue movement and task oriented commands to the robot. The system also facilitates the collection and distribution of scientific information to experts on earth or team members in a habitation module. The progress of this project is measured by the successful execution of field tests in Mars analog environments such as the Arctic and the desert.

Planning systems are an important part of semi-autonomous robotic systems such as planetary rovers, robotic aircraft, and in other complex planning and scheduling tasks such as daily MER activity scheduling. RIALIST has worked this year to refine the spoken dialogue interface to EUROPA and interface with the Ames ARA IDEA interleaved planning & execution system.

Robustness in noisy environments is an important feature for NASA spoken dialogue systems, particularly on ISS. Our current project in this area is in collaboration with Dr. T. Berger at USC investigating the use of neural net technology to recognize speech in very noisy conditions.

### II.B.2.d. Accomplishments during FY2002

**Assess JSC High Payoff Applications for Spoken Dialogue Systems**

The RIALIST group interacted with MCC, Shuttle Operations, ISS Operations, and training staff to explore where the best payoffs would be for a spoken dialogue system to aid in operations. We evaluated the practical and technical requirements for possible applications in these areas. With the collaboration of astronauts, managers and trainers we concluded that an intelligent checklist and procedure system was the best initial application.

**Procedure and Checklist Assistant Systems**

We built an initial demonstrator which showed the general form and potential capabilities of an intelligent procedure assistant. This was shown to a group of astronauts and their feedback was used to refine the design. Several suggestions were received about the best procedure to use in an initial prototype system. The procedure chosen was the Potable Water Sampling Procedure which is a long an complicated test of the drinking water quality and safety.

We built a prototype system which operates using the actual water sampling procedure. This system has dialogue features such as the ability to navigate the procedure, gracefully correct misunderstandings, take and play back audio notes, set alarms using spoken commands, keep track of the progress through the procedure (useful for interruptions) and to give more details concerning the operations being performed. These features were deemed important by the astronauts and trainers we met at JSC. The system includes a synchronized display which highlights the current step in the procedure and which can display diagrams and photographs illustrating functional information about equipment used in the procedure. We are experimenting with new more natural sounding and understandable speech synthesis systems. We have built infrastructure to support rapid changes of language and to accommodate use of the system with a variety of procedures.

Since the system is trained on examples of spoken interactions between astronauts, we have initiated recording of astronauts during training sessions. We have also arranged to have astronauts interact with the system in order to provide user feedback and additional data. This data, when annotated, can be used in an automatic training procedure to improve the performance of the system.

**Helpful Listening for Air Traffic Control**

We worked with Nancy Smith (Code IH) to improve the recording quality in the air traffic controller simulation facility. We received speech data collected in several exercises with simulated pseudo pilots and air traffic controllers. Preliminary analysis of syntactic forms were produced in simulations.

**Dialogue Interface to the Europa and IDEA**

During the year, we further developed a spoken dialogue interface to Europa and IDEA using the RIALIST PSA demonstration system. The dialogue allows the monitoring of plan execution and re-planning when the environment changes or resources become depleted. Additional capability to ask the PSA where it is located at present has been developed. These efforts were in collaboration with Jeremy Frank and Nicola Muscettola.

**Dialogue in Mobile Agents**

We refined our spoken dialogue interface to the Brahms work practices and knowledge capture modeling system which allows belief addition and revision. We collected speech data in pressurized space suits. At the field tests at JSC and Arizona, our system allowed the suit subject to ask questions about where he was located, and to issue movement and task oriented commands to the robot. The system also allowed the collection and distribution of scientific information to geological experts at JSC and team members at the field test. These efforts were in collaboration with Bill Clancey and Maarten Sierhuis.

**Eye Tracking**

We built an initial reference resolution component with an architecture that facilitates integration of multiple sources of information including eye tracking in the reference resolution process. We designed experiments for investigating the correlation of natural eye movements with spoken dialogue events. Ellen Campana was the SSRP student who worked on this project. This work was in collaboration with James Allen, Mike Tannenhaus, Lee Stone and Roger Remington.

**Noise Robust Speech Recognition**

We developed new system requirements for a robust phoneme recognizer using the neural network technology developed at USC. We met with the USC group and offered suggestions for speech data and tests to be run on the phoneme based system. The initial USC phoneme system is now functioning and ready for further development and integration. This work was performed in collaboration with Dr. Ted Berger.

**Basic Technology Development**

RIALIST continues to research and develop spoken dialogue capabilities that are applicable across tasks. There were two major areas of research and development of this type during FY02: Language Modeling using Explanation Based Learning and Targeted Help.

- Language Modeling using Explanation Based Learning (EBL) and Grammar Compilation:
  Language Models provide a necessary constraint on speech recognition by restricting recognition to the range of language needed for a particular application. In collaboration with Dr. Manny Rayner, we have made significant improvements in our capabilities for creating application specific grammars semi-automatically. We have refined our use of the machine learning technique, EBL, for creating these application specific grammars from a general grammar of English using a very small amount of training data. We have also successfully tested some new techniques for compiling the specialized grammars in to context-free language models. These new techniques are faster and more likely to stay within memory limitations than previous approaches. We anticipate releasing the tools developed in this research effort as open source in the near future.

- Targeted Help:
  Research indicates that the best speech recognition is achieved by expert users on systems using grammar-based language models. Naïve users do not understand the capabilities of the system and because of that produce many utterances that are out of the system's coverage. Targeted Help is an embedded training approach that gives feedback when the user produces an utterance that the system cannot understand and guides the user towards producing in-coverage language. As a result, users should become experts faster and spend less time producing unproductive utterances. Results show that users who receive Targeted Help are significantly more able to complete tasks with the dialogue system compared to users who do not receive Targeted Help. This work was done in collaboration with Oliver Lemon and Stanley Peters of the Computational Semantics Lab at Stanford.

**Open Source NLP (The Leo Project)**

The Regulus component of the Leo Project now provides capabilities for parsing and grammar compilation. We have also developed a much larger English grammar for use with Regulus and translation components between Regulus and Gemini grammar formats. The Regulus tools have provided valuable support for the group's Language Modeling research.

**Proposals**
Submitted 1 NSF Proposal, 1 IS proposal and 1 HEDS-CAN proposal.

**Recruiting**
Recruited two new Ph.D. Researchers and hired one of them.

### II.B.2.e. Problems Encountered and Possible Resolution

None

### II.B.2.f. Publications and Presentations during FY2002

#### II.B.2.f.i. Publications


75. Aist,G., Dowding, J., Hockey, B.A., and Hieronymus, J. "Intelligent Procedure Assistant for Astronaut Training and Support." Association for Computational Linguistics 2002 meeting, Demo session. July 2002. Philadelphia.
76. Aist, G., and Hockey, B. A. "Generating Training and Assistive Dialogues for Astronauts from International Space Station Technical Documentation." Intelligent Tutoring Systems, Workshop on Integrating Technical and Training Documentation. Biaritz, France & San Sebastian, Spain. June 2002.
77. Allen, J.F.,Campana, E. and Swift, M.D. (in press). The future of dialogue systems. To appear in J.C. Trueswell & M.K. Tanenhaus (eds.). World-situated language processing: Bridging the language as product and language as action traditions (MIT Press).
78. Baldridge, J., Dowding, J., and Early, S. LEO: An Architecture for Sharing Resources for Unification-Based Grammars, LREC June 2002, Canary Islands, Spain.
79. Blaylock, J., Dowding, J., and Allen, J. A Dialogue Model for Interactions with Planners, Schedulers, and Executives, to appear at the rd International NASA Planning and Scheduling Workshop, October 2002.
80. Brown-Schmidt, S., Campana, E. and Tanenhaus, M.K. (in press). Real-time reference resolution by naive participants during a task-based unscripted conversation. To appear in J.C. Trueswell & M.K. Tanenhaus (eds.).World-situated language processing: Bridging the language as product and language as action traditions (MIT Press).
81. Brown-Schmidt, S., Campana, E. and Tanenhaus, M.K. (2002). Reference resolution in the wild: Online circumscription of referential domains in a natural interactive problem-solving task. In Proceedings of the 24th Annual Meeting of the Cognitive Science Society. Fairfax, VA, August 2002.
82. Campana, E., Baldridge, J., Dowding, J., Hockey, B.A., Remington, R. and Stone, L. "Using Eye Movements to Determine Referents in a Spoken Dialogue System." PUI, November 2001, Florida.
83. Campana, E., Brown-Schmidt, S. and Tanenhaus, M.K. (2002). Reference resolution by human partners in a natural interactive problem-solving task. In Proceedings of the 7th International Conference on Spoken Language Processing (ICSLP): Denver, CO, September 2002.

84. Dowding, J., Frank, J., Hockey, B.A., Jonsson, A., Aist, G., Hieronymus, J., A Spoken Dialogue Interface to the EUROPA Planner, to appear at the 3rd International NASA Planning and Scheduling Workshop, October 2002.
85. Dowding, J., Frank, J., Hockey, B.A., Jonsson, A., and Aist, G. "Demonstration of a Spoken Dialogue Interface for Planning Activities of a Semi-autonomous Robot." Human Language Technology, April 2002.
86. Gorrell, G., Lewin, I., and M. Rayner. Adding Intelligent Help to Mixed-Initiative Spoken Dialogue Systems. Proceedings of the International Conference on Spoken Language Processing (ICSLP) 2002, Denver, Colorado, September 2002.
87. Hockey, B.A., Dowding, J., Aist, G., and Hieronymus, J. "Targeted Help and Dialogue about Plans." Association for Computational Linguistics (ACL) 2002 meeting, Demo session. July 2002. Philadelphia.
88. Hockey, B. A., Aist, G., Hieronymus, J., Lemon, O., and Dowding, J. "Targeted Help: Embedded training and methods for evaluation." Intelligent Tutoring Systems, Empirical Methods Workshop, June 2002.
89. Rayner, M., B.A. Hockey and J. Dowding. Grammar Specialization Meets Language Modelling. Proceedings of the International Conference on Spoken Language Processing (ICSLP) 2002, Denver, Colorado, September 2002.
90. Rayner, M., J. Boye, I. Lewin and G. Gorrell. Plug and Play Spoken Dialogue Processing. To appear in Smith, R. and J. van Kupperfeld (eds.), Current and Future Directions in Dialogue Processing, Kluwer, 2002.
91. Rayner, M, Hockey, B.A. and Dowding, J. "Grammar Specialization Meets Language Modelling.", Proceedings of the International Conference on Spoken Language Processing, Denver, Colorado, September 2002.
92. Swift, M.D., Campana, E., Allen, J.F., and Tanenhaus, M.K. (2002). Incremental referential domain circumscription during processing of natural and synthesized speech. In Proceedings of the 24th Annual Meeting of the Cognitive Science Society. Fairfax, VA, August 2002.


### II.B.2.f.ii. Presentations

64. Dowding, J., "Talking with EUROPA", Autonomy and Robotics Area Meeting, NASA Ames Research Center, May 14, 2002. Dowding, J., "A Spoken Dialogue Interface to EUROPA", SRI International, June 14, 2002.
65. Hieronymus, J., "Dialogue Systems in Space", Invited speaker for the Nordic Summer School on Uses of Spoken Language Corpora, Gothenburg University, Gothenburg, Sweden, Hieronymus, J. Research In Advanced Language Interfaces and Speech Technology (RIALIST) Research Program. Invited speaker HCC Review Meeting, Pensacola, FL.

### II.B.2.f.iii. Other Activities

- Aist, G. Reviewer, Intelligent Tutoring Systems 2002 Workshop on Empirical Methods.
- Aist, G. Panel member, Intelligent Tutoring Systems 2002 Workshop on Empirical Methods, Panel on Coding Dialogue.
- Dowding, J. Visiting Lecturer, Symbolic Systems Program, Stanford University
- Dowding, J. Program Committee, TALN 2002, Traitement Automatique des Langues

- Naturelle, Nancy, France, June 24-27, 2002.
- Dowding, J. Program Committee, International CLASS (Collaboration in Language and
- Speech Science) Workshop on Natural, Intelligent and Effective
- Multimodal Dialogue Systems, Copenhagen, Denmark, June 28-29, 2002.
- Dowding, J. Reviewer for a variety of conferences, workshops, journals, and books.
- Dowding, J. Reviewer for NASA SBIR program
- Hockey, B.A. Reviewer for Autonomous Agents and Multiagent Systems Conference (AAMAS)
- Hieronymus, J. Reviewer, NASA SBIR Phase I, NASA STTR Phase I,
- Hieronymus, J. Reviewer, IEEE Speech and Audio Processing Journal

### II.B.2.g. Future Plans

Goal: Creating natural spoken dialogue interfaces to ISS procedures, robots, aircraft, spacecraft, air traffic control and instructional systems which are flexible to use and model human behavior.

### II.B.2.h. Other Information

**Awards**

- Aist G. Fulbright Scholar grant, 2002-2003 (declined)
- Aist, G. Distinguished Finalist citation for the Outstanding Dissertation of the Year Award from the International Reading Association, 2002

**Non-RIACS students supervised**

- Susana Early (DeAnza)
- Brad Boven (Kalamazoo)
- Steven Phan (Santa Clara)

**Collaborations**

Developed collaborations with:

- The Stanford Semantics Lab headed by Stanley Peters with Oliver Lemon and Alex Gruenstein.
- EU Siridus project (Gothenburg U.(Sweden), University of Saville (Spain), Telefonica (Spain), and SRI Cambridge (UK)
- Ted Burger (USC)
- Lee Stone and Roger Remington (Code IH, NASA Ames)
- Walter Johnson (Code IH, NASA Ames)
- Nancy Smith (Code IH, NASA Ames)
- Liz Shriberg and Andreas Stolcke (ICSI and SRI )
- Alex Rudnicky (CMU)
- James Allen (Rochester)

## II.B.3. Information Structures Middleware

### II.B.3.a. Task Summary

Information Structures (Code IC) research develops technologies for the distributed and remote integration and coordination of Science and Engineering data, information, and knowledge. The key contribution of the Information Structures project is the access by NASA Science and Engineering teams to a plethora of heterogeneous data, information, and knowledge. Information Structures includes a middleware infrastructure that will permit tools, instrumentation, and processes of heterogeneous fidelity to work together to support the design, test, and analyses of data, information, and knowledge processes.

The purpose of this task is to research middleware technologies for Information Structures, and to design the architecture for, and to lead the development and evolution of, an appropriate middleware infrastructure implementation.

### II.B.3.b. RIACS Staff

Ronald Mak

### II.B.3.c. Project Description

The research objectives include:

- Analysis and specification of the requirements of the Information Structures middleware. The requirements include object caching, resource pooling, asynchronous messaging, event subscription management, and transaction control in an environment that supports scalability, reliability, security, maintainability, and reusability.
- In–depth review, evaluation, and recommendation of commercial off–the–shelf application server software that can help meet the above requirements.
- Review and recommendation of appropriate software development tools.
- Design, prototyping, and implementation of middleware that integrates an application server with custom in–house software, data, and processes. This middleware will meet the requirements and support the goals of Information Structures.

### II.B.3.d. Accomplishments during FY2002

- Completed the middleware requirements analysis and specification for version 1.0.
- Completed the application server software review, evaluation, and recommendation (of the BEA WebLogic Application Server).
- Completed the review and recommendation of software development tools.
- Completed the middleware design for version 1.0. Prototyping demonstrated that the design would meet the requirements and support the goals of Information Structures.
- Successfully led the development, implementation, and deployment of version 1.0 of the Information Structures middleware.

### II.B.3.e. Problems Encountered and Possible Resolution

Many of the Information Structures applications were already designed and partially implemented before the middleware was designed. This resulted in extra integration challenges that needed to be overcome during the "final builds". In the future, the middleware and the applications should be designed concurrently.

### II.B.3.f. Publications and Presentations during FY2002

#### II.B.3.f.i. Publications

93. Mak, Ronald, *Java Number Cruncher: The Java Programmer's Guide to Numerical Computing*, ISBN 0–13–046041–9, Prentice Hall PTR, 2002, approximately 470 pages. A technical book published in October 2002, details are at http://www.apropos-logic.com/nc/.

#### II.B.3.f.ii. Presentations

66. Mak, Ronald, "The Architecture of the InfoStructure Middleware," SAE 2002 World Aviation Congress," Phoenix, AZ, Nov. 2002.

#### II.B.3.f.iii. Other Activities

Volunteer work at the Computer History Museum located at NASA Ames, including serving as a tour docent, and the research for and writing of exhibit labels. More information is at http://www.computerhistory.org.

### II.B.3.g. Future Plans

- Research future technologies for the evolution of Information Structures, including web services, data warehousing, data mining, semantic tagging, wireless communications, etc.
- Design, lead, and implementation of version 1.1 of the Information Structures middleware.
- Coordinate with other NASA groups and organizations that wish to use or share middleware technologies.
- Technical paper on the numerical computation of $\pi$ (publication to be determined).
- Continue volunteer work at the Computer History Museum.
- Presentation of the session "Darwin, J2EE, and Mars" at the JavaOne Conference, San Francisco, CA, June 2003 (pending acceptance of proposal).
- Presentation of the session "BigInteger, BigDecimal, and a Billion Digits of Pi" at the JavaOne Conference, San Francisco, CA, June 2003 (pending acceptance of proposal).
- Reprise at Ames of the World Aviation Congress presentation on middleware (see II.B.3.f.ii), possibly as a RIACS seminar.

## II.B.4. Information Integration Issues in Aviation Data Integration

### II.B.4.a. Task Summary

We are exploring a number of issues in the area of information integration from heterogeneous sources. The work is being done in the context of the ongoing Aviation Data Integration Project (ADIP). ADIP (http://ic.arc.nasa.gov/projects/adip/) is an ongoing project on building systems that provided integrated access to heterogeneous data sources of interest to aviation safety analysts, and is one of the projects of the Information Sharing and Integration Group (http://sciencedesk.arc.nasa.gov/isig/) in the CAS area. We are working on both extending functionality as well as improving performance for the ADIP system.

An important area of interest is incorporating sources of textual data such as aviation safety reports (ASRS) into the integrated system. We want to extend the functionality of ADIP to able to extract information from textual data sources as well as integrate sources of structured (quantitative) data with unstructured (qualitative) data. We are also working on performance optimization issues for ADIP. The system integrates information from a number of remote (internet) sources and retrieving data from these different sources over the network (in response to a user query) takes a long time. We are developing schemes for intelligently materializing data from the remote sources to improve the performance of the system.

While the motivation and context for the above work has been provided by ADIP, the techniques we are developing (both for textual data management and integration and for performance optimization) are general purpose.

### II.B.4.b. RIACS Staff

Naveen Ashish

### II.B.4.c. Project Description

For providing integrated access to textual data sources in the ADIP system, the research objectives being pursued are:

- Developing techniques and using existing techniques for textual (semi-structured) data analysis tasks such as clustering, classification, summarization etc.
- Developing and using techniques for information extraction and information retrieval from textual data sources
- Developing techniques for providing integrated query access to sources of both structured and unstructured data.

The research objectives being pursued to provide performance optimization in a system such as ADIP are:

- Developing a framework for materialization of data.
- Developing smart strategies for downloading (prefetching) data from internet sources.
- Developing techniques for user interface specification and analysis to predict the probability distribution for data access.
- Investigating and developing smart polling strategies to detect that data has changed for a source whose data can changed and unspecified times and/or frequency.

### II.B.4.d. Accomplishments during FY2002

- Completed careful evaluation of existing text mining and information retrieval software such as IBM Intelligent Text Miner, Vivisimo and Qurom for the task of extracting information from textual data sources.

- Currently developing an architecture for providing integrated access to structured and unstructured data sources in ADIP.
- Completed survey of existing techniques for intelligently caching data (for performance optimization) fm remote sources (that change at unknown times and frequencies).
- Development of performance optimization materialization system (for system like ADIP) currently in progress.

### II.B.4.e. Problems Encountered and Possible Resolution

None

### II.B.4.f. Publications and Presentations during FY2002

#### II.B.4.f.i. Presentations

67. Kulkarni, Deepak, "Aviation Data Integration Project", *Presentation to Alaska Airlines*, Seattle WA, Sep 2002.

#### II.B.4.f.ii. Other Activities

- I served as an expert reviewer for 2 papers for *IEEE Transactions on Knowledge and Data Engineering* in 2002
- I also served as expert reviewer for the *International Journal of Co-operative Information Systems* in 2002

### II.B.4.g. Future Plans

- Continue work on above mentioned thrusts on semi-structured data access and performance optimization
- Build information integration applications of interest to aviation safety analysts.
- Explore synergies with proposed information integration research by our group in other domains such as earth science and disaster management information systems.

### II.B.4.h. Other Information

- Aviation Data Integration Project (ADIP) http://ic.arc.nasa.gov/projects/adip/

## II.B.5. Information Management Requirements Study

### II.B.5.a. Task Summary and Overview

A study, jointly supported by NASA, NSF, and DARPA, is being conducted into the requirements for technology advances to achieve the needed capabilities for information management in domains of particular interest to sectors of the government. This study will be carried out by a panel of experts in the related disciplines, and will produce its results within one year of the start of the effort.

### II.B.5.b. RIACS Staff

Barry Leiner

Linda Andrews

**II.B.5.b.i. Visiting Staff**

Sara Graves (Univ of Alabama at Huntsville)
Craig Knoblock (Univ of Southern California/Information Sciences Institute)
Larry Lannom (Corp. for National Research Initiatives)

**II.B.5.c. Project Description**

Advances in networking and computing technologies have fundamentally changed the problem space for information management and its various application domains. Before the availability of high bandwidth networking, networked high performance computing, and large scale storage, the major problem was the accessibility of information. Information was often stored in forms that were not easily available, and finding and retrieving the required information was difficult.

The technology advances have changed this situation. If information is on a networked computer, it can be accessed and retrieved (assuming appropriate access control permissions.) Large amounts of data can be searched and indexed, and queries can be done against such indexes to find relevant data.

The new problem space is a direct result of these advances. So much information is accessible and retrievable that the challenge is finding the right information. (As an example, a "Google" search for the term "information management" returned 1,450,000 on Nov 22, 2002) hits.) Information is stored in various formats, and the user may not have the means for dealing with the obtainable data. Turning data into actionable and coordinated information requires structure, and the large and varied amounts of information on the network make structuring that information difficult. Preserving information as the underlying infrastructure evolves becomes problematic, particularly given the amount of information being generated.

Yet information management technologies are having and will have major impact on a number of application domains of interest to multiple government sectors, such as coalition mission operations, logistics support, scientific data management, and digital libraries. Each of these application domains (and others) of information management have their own requirements as we move forward. However, there is considerable overlap of requirements, and investments in one area can have significant benefits to others.

This study will take an integrated look at the various application domains of information management (digital libraries, mission operations, logistics and maintenance databases, scientific data management) to understand the common and unique requirements of the domains, and the state of the art of the technology that can support such requirements.

**II.B.5.d. Accomplishments during FY2002**

During the course of this one-year task, we held two workshop/meetings at Ames Research Center: A kickoff meeting to introduce the panel members and determine the scope and direction of the study; and, a wrap-up meeting to solicit any outstanding comments or issues and to solidify the content of the final report.

In between these two meetings, there was an on-going electronic discussion by the co-chairs and panel members with the purpose of better understanding the synergies and differences between the three application areas covered in the study. These discussions were followed by a significant writing effort to compose and perfect the resulting final report.

### II.B.5.e. Problems Encountered and Possible Resolution

None

### II.B.5.f. Publications and Presentations during FY2002

#### II.B.5.f.i. Publications

94. Graves, Sara, Knoblock, Craig and Lannom, Larry (2002), "Technology Requirements for Information Management". RIACS Technical Report 02.07. Available at http://eprints.riacs.edu/documents/data/00/00/00/90/index.html

### II.B.5.g. Future Plans

The final report was released in November 2002. Discussions with the government points of contact will continue to make the results widely available.

## *II.C. High Performance Computing and Networking*

Advances in the performance of computing and networking continue to have major impact on a variety of NASA endeavors, ranging from modeling and simulation to data analysis of large datasets to collaborative engineering, planning and execution. RIACS and NASA collaborate in several areas as described in the following sections.

## II.C.1. High-performance Networking and Applications Project

### II.C.1.a. Task Summary and Overview

The NASA Research and Education Network (NREN) team conducts research to enable the infusion of emerging network technologies into NASA mission applications. The NREN testbed peers with Next Generation Internet (NGI) testbeds sponsored by other Federal agencies and with the university-led Internet2 testbed to provide a nationwide platform for conducting network research and for prototyping and demonstrating revolutionary applications. Emerging technologies will enable new methodologies for achieving NASA science, engineering, and education objectives. RIACS provides support for the NREN group at NASA Ames, including conducting basic networking research, participating in projects to develop new networking technology and revolutionary network applications, and assisting with software engineering and network performance engineering. For more information about NREN activities see http://www.nren.nasa.gov.

### II.C.1.b. RIACS Staff

Marjory Johnson
Jerry Toung

#### II.C.1.b.i. Visiting Staff

None

### II.C.1.c. Project Description

M. Johnson is deputy task manager for the NREN group, which is now part of the Computing, Networking, and Information Systems (CNIS) Project within the Computing, Information and Communications Technology (CICT) Program. J. Toung is a member of the NREN applications and research group, specializing in network software engineering. Components of the RIACS task include assisting in programmatic activities, participating in specified research projects for the development of new networking technology, interfacing with the university research community, participating in development of revolutionary network applications, and publishing and presenting results at conferences.

### II.C.1.d. Accomplishments during FY2002

- NASA Protocols Testbed
  M. Johnson is technical lead for the NASA Protocols Testbed project, a new task this year. The goal is to coordinate activities at NASA centers to evaluate protocols for space communications, and to ensure that test hypotheses match mission requirements.

Our group at ARC will conduct independent testing to fill in any gaps, analyze results to establish performance envelopes for various protocols, and develop and maintain a repository that organizes test information and that provides guidance to mission developers. Accomplishments this year include identification of requirements for space mission applications for various NASA Enterprises, development of the repository, and preliminary analysis of selected link-layer, transport-layer, and file-transfer protocols. The repository is located at http://www.npt.nren.nasa.gov.

- Mobile IP
J. Toung conducted a survey of Mobile IP, covering routing protocols, activities within the Internet Engineering Task Force, vendor products, and university projects. He presented this information to the NREN team, both orally and in a written report. The purpose of this activity was to prepare for future experimentation with mobile networking and the prototyping of an application that requires connection of a mobile node to the NASA Grid.

- Network Quality of Service Monitoring
J. Toung is designing and developing software for a network monitoring tool called PCMon. PCMon measures throughput of individual traffic flows, in contrast to traditional tools, which measure aggregate bandwidth. Hence, PCMon enables measurement of the effectiveness of various approaches for achieving network Quality of Service (QoS), e.g., assigning preferential treatment to specified traffic flows across a network.

  FY 2002 accomplishments include deployment of PCMon on the NREN testbed, development of a web-based graphical user interface (GUI) to facilitate the use of PCMon, and using PCMon to characterize traffic generated by selected Grand Challenge applications running on the NASA Grid.

  The current focus of this activity is integration of PCMon and a tool for making advance bandwidth reservations to create a new tool called the Resource Allocation, Measurement and Adjustment System (RAMAS). The objective of RAMAS is to enable network resource allocations to be adjusted dynamically, in response to changing network conditions. The basic functionality of the RAMAS tool has been successfully demonstrated over a wide-area network.

- Programmatic Support
  - M. Johnson is deputy task manager.
  - M. Johnson assisted in planning for the CNIS Project.
  - M. Johnson assisted in organizing the Grand Challenge Applications task and integrating it with other tasks within the CNIS Project.
  - M. Johnson assisted in preparation for a review of the CICT Program by the National Research Council.
  - M. Johnson is a member of the LSN Network Research Team, which coordinates research activities among the Federal agencies.

  - M. Johnson manages NREN research grants to universities.

- Other Activities
  - M. Johnson is co-chair of the program committee for the ICC 2003 Next Generation Internet Symposium.
  - M. Johnson collaborated with Professor Mario Gerla of UCLA to prepare an NSF proposal entitled "Efficient Transport of Bandwidth Intensive Data and Streams in Next Generation Internet Wire/Wireless Environments." The proposal has been funded, and work has begun. NASA is an unfunded participant.
  - M. Johnson helped to organize the NASA Second Space Internet Workshop, May 2002.
  - M. Johnson supported the CICT Space Communications Project by participating in a strategic planning workshop to develop guidelines for future NASA space communications research and by participating in the review of proposals submitted in response to a NASA Research Announcement
  - M. Johnson was a member of the program committee for the IFIP Workshop on Protocols for High Speed Networks, April 2002.
  - M. Johnson attended the MobiCom 2002 conference, September 2002.
  - J. Toung provided support for the NREN application at Supercomputing 2001.
  - J. Toung participated in the "Wireless Communications and Networking" workshop at JPL, September 2002.

### II.C.1.e. Problems Encountered and Possible Resolution
None

### II.C.1.f. Presentations during FY2002
68. M. Johnson, K. Freeman, and R. desJardins, "NASA Protocols Testbed," Second NASA Space Internet Workshop, May 2002.
69. M. Johnson, "Networking Technologies for Collaboration," Astrobiology Science Conference, sponsored by NASA Astrobiology Institute, April 2002.
70. J. Toung, "PCMon – A Network Monitoring Tool," RIACS Seminar, June 2002.

### II.C.1.g. Future Plans
Networking technologies that will be the focus of NREN activities next year include the integration of QoS with measurement and monitoring, mobile and ad hoc networking, and collaborative technologies. Future plans include the development of a testbed between ARC and GRC to evaluate mobile networking over IPv6, participation in a Federal agency activity to develop an architecture for space communications, application of encryption techniques to ensure secure communications between the monitoring and management elements of PCMon, and development of a Resource Decision Manager (RDM) for the RAMAS tool. The RDM would handle communications between the monitoring element and the resource-reservation element of RAMAS.

## II.C.2. Network System Support

### II.C.2.a. Task Summary and Overview

This task is to assist the network operations staff in operating the Ames Local Area Network efficiently and securely.

### II.C.2.b. RIACS Staff

David L. Gehrt

### II.C.2.c. Project Description

In performance of this task the RIACS personnel maintain and operate the authoritative name servers. This includes maintaining the security of the name servers including restricting access by unauthorized personnel and logging attempts to gain unauthorized access to the name servers and other systems and networks administered by RIACS personnel.

RIACS personnel also provide miscellaneous assistance as requested by network operations personnel on matters related to network operations.

### II.C.2.d. Accomplishments during FY2001

A major activity affecting the network at the Ames Research Center during the life of this task has been the continuation to cmpletion of the establishment of three separate Physical networks designated Public, Private and Open. Each of these networks is isolated from the Internet and other ARCLAN networks by a firewall and the Ames DMZ and possibly by an additional firewall. The three new networks each have access regulated by their own set of security rules, now all networked computers and networked peripherals are attached to one of these networks. This reconfiguration required the re-addressing of virtually all of the IP addresses used on the Ames networks. The changes were made building by building through the use of a script developed a part of this task. This script allowed the bulk conversion of the addresses for an entire building or set of buildings. As a result of this process errors are reduced and the workload on the Network Staff has been significantly reduced.

Under this task the hardware and configuration of the authoritative name servers has been modified. Six new Compaq systems were brought on line. Each of the new systems runs a version of the RedHat Linux operating system. One of these systems is configured as a hidden master name server. The only use made of this system is the maintenance of the source files from which the Centers DNS tables are generated, and the distribution of the tables to the user system accessible authoritative name servers.

The remaining Compaq systems are the Center's authoritative name servers. The authoritative name servers are physically located in room 150 in building N233, but logically these systems are connected to each of the three networks that make up the ARCLAN. Currently one of these systems is connected to the Public and Private network, with the remaining systems connected to the Open network.

With the installation of the Compaq systems the authoritative name servers in use previously, aging Sun workstations, were retired. The new systems provide a significant performance improvement for the Centers DNS. Along with the installation of the Compaq systems the configuration changes involved upgrading the BIND server software to version 9.x from version 8.x. The hidden master name server is configured to not respond to queries from user systems. Its only functions are those described above. Use of the remaining to handle recursive queries is restricted to systems with IP addresses authorized for use on the various ARCLAN networks.

Access to each of the Center's name servers is restricted to a very small set of authorized users and then only through the use of Secure Shell (SSH).

Unauthorized attempts to access any of the name servers are logged to a remote system and reported to the system administrator for investigation. Unauthorized attempts to access ARCLAN systems and networks reported to the administrative personnel and of a serious nature are reported to the Center Computer Security staff for their action. These log messages continue to be saved in a relational database.

#### II.C.2.e. Future Plans

In the future NASA plans to retire all the old authoritative name servers and replace them with new systems recently acquired by the Government. Additionally the DNS software will be maintained at the most recent and secure version.

## II.C.3. Engineering for Complex Systems

#### II.C.3.a. Task Summary

Social and organizational factors in mission design and operations are a main source of both the success and failure of complex NASA missions. This project is conducting research related to the use of information technology to address such social and organizational factors through knowledge management. Technical advances are focused on collaborative systems involving social networks and 3D graphics.

This research is being conducted in support of the System Reasoning and Risk Management (SRRM) element within the Engineering for Complex Systems (ECS) program, and the Virtual Iron Bird (VIB) activities also within ECS (and other NASA programs). SRRM is focused on developing tools and methods for risk identification, quantification and management, in support of mission design and operations. VIB is focused on developing tools and methods that support software and integrated system testing for the Space Shuttle and International Space Station.
http://ecs.arc.nasa.gov/

#### II.C.3.b. RIACS Staff

David G. Bell
Julian Gomez

### II.C.3.b.i. Visiting Staff and Students

Mohana M. Gurram

### II.C.3.c. Project Description

The following two projects were participated in under this task during FY02:

**Mishap Initiator Identification System**

Risk identification, analysis, and lifecycle decision support is an ongoing challenge within NASA. This work will test the hypothesis that: 1) systematic investigation of analysis of agency problems (e.g., mishaps, anomalies) and solutions (e.g., recommendations and lessons learned) can be used to improve future mission and program success by enabling improved risk identification, analysis and lifecycle decision support; 2) both physical (i.e., hardware and software) and managerial (i.e., human and organizational) causes of failures and problems, as well as mitigators to those causes can be addressed through such systematic analysis; and 3) that distributed web technology for such systematic investigation and analysis can provide a more effective capability for NASA to utilize knowledge of mishap causes and solutions in current and future programs and missions to improve their success.

**Virtual Iron Bird (VIB)**

As part of the VIB efforts, this project will focus on the development of advanced modeling and visualization based on intelligent diagnostic technology that will allow reliable, accurate, and autonomous assessments of system state. The project plan includes participation in the creation of a virtual space shuttle orbiter model that is based around collections of entities within the orbiter, including legacy data, engineering data, and 3D graphics models. This research includes the determination of a relevant ontology and taxonomy for human and machine agent use, and to define long-term knowledge management categorization/classification. This research also involves investigating new ways of more tightly coupling 3D-graphics applications with multiple databases of information to support knowledge management.

### II.C.3.d. Accomplishments during FY2002

- ECS-1 "Prototype Mishap Initiator Identification System" milestone was accomplished.
- 90% & 100% reviews were completed obtaining internal and external customer feedback.
- Draft ECS-7 milestone vision & planning document completed.

### II.C.3.e. Problems Encountered and Possible Resolution

None

### II.C.3.f. Publications and Presentations during FY2002

### II.C.3.f.i. Publications

95. Bier, Eric A., Conley, Ken, Pier, Ken, Liang, Tao, Chang, Bay-Wei and Bell, David G., "Sharing Information on Structured Web Pages with Multiple Authoring Roles", Paper

submitted for review to the Conference on Human Factors in Computing Systems (CHI2003).

### II.C.3.f.ii. Presentations

71. Bell, David G., Newman, Susan, Repenning, Nelson, "Process, Practice & Politics: Understanding the relationship between documentation, deployment and daily work", Academy of Management Conference, Denver, CO, August 2002.
72. Bell, David G. "Mishap Initiator Identification System", Presentation at the ECS Annual Review, August 2002.
73. Gomez, Julian. "Data from a New Perspective", Presentation to the Bay Area NeXT Group, Cupertino, CA, August 2002.

### II.C.3.f.iii. Other Activities

- Gomez, Julian. Sponsored and organized an ARC wide seminar on X3D, August, 2002.
- Gomez, Julian. Providing editorial review of academic textbook on the Java Media Framework.

### II.C.3.g. Future Plans

The following research and results are planned for the coming fiscal year:

- Initial study of mishap cause classification complete with resulting academic paper.
- A field study on mishap investigation practice and redesign of a mishap investigation tool that integrates a diagrammatic user interface with a semantic network database.
- A work practice study of formal reviews in collaboration with faculty in the Management Science & Engineering Department at Stanford University.
- Developing a spatial (volumetric) classification scheme for laser-scanned orbiter.

## *II.D. Applications of Information Technology*

In addition to the areas of research outlined above, RIACS scientists and visitors have collaborated with NASA researchers in an interdisciplinary context to apply and advance information technology in a number of application domains, as described in the following sections.

## II.D.1. NASA Astrobiology Institute

### II.D.1.a. Task Summary

The NASA Astrobiology Institute (NAI) was established in 1998 to encourage and fund collaborative multidisciplinary research in astrobiology and to foster innovative contributions to the scientific research agendas of NASA space missions. NAI represents a partnership between NASA and a number of academic institutions and research organizations and is currently composed of 15 Lead Research Teams that together represent over 700 investigators located at over one hundred institutions across the United States, plus teams in the U.K., Spain, and Australia. The NAI administrative group, known as NAI Central, is situated at Ames Research Center.

Since an overarching goal of the NAI and its members is innovative research in astrobiology with an emphasis on collaborative work, both within and among its geographically distributed teams, NAI was conceived from the beginning to operate as a "virtual institute" of collaborating scientists. To meet its virtual institute objectives NAI planned for the deployment of various communication/collaboration tools (hardware and software), the development of community building efforts including opportunities for scientific interaction potentially leading to new collaborations within and among NAI Teams, and on-going evaluation of collaboration support in order to continuously develop and improve the effectiveness of NAI.

RIACS supports the continued development of NAI by providing staffing for the role of NAI Collaborative Research Manager. The primary responsibility of this role is to provide leadership in defining, evolving, implementing, and maintaining an integrated collaborative infrastructure. The task includes a focus on both the appropriate technology architecture for the NAI as an effective virtual institute and on the premises and practices that facilitate the development of a collaborative organizational culture.

For more information about the NASA Astrobiology Institute, please visit:
http://www.nai.arc.nasa.gov.

### II.D.1.b. RIACS Staff

Lisa Faithorn

#### II.D.1.b.i. Subcontractors

Future U LLC

### II.D.1.c. Project Description

Lisa Faithorn is the Manager of Collaborative Research for NAI. A primary responsibility of this role includes guiding the research and implementation efforts of leading edge practices, tools and infrastructure that together result in a "center without walls" in which scientists can conduct collaborative research without regard to geographical location - interacting with colleagues, sharing data and computational resources, accessing instrumentation, retrieving information from digital libraries and contributing to a shared knowledge base as the field of astrobiology develops. Responsibility also includes overseeing the successful operations of the technology infrastructure components already in place.

An additional component of this RIACS task is participation in this research arena by attending and presenting at conferences, as well as publishing the results of NAI experimentation in the virtual scientific institute concept.

### II.D.1.d. Accomplishments during FY2002

- Completion of NAI Member-wide Communication/Collaboration Needs Assessment:
  In September – October of 2001, Dr. Faithorn conducted a comprehensive Communication/Collaboration Needs Assessment Survey that was made available electronically to all members of NAI. The survey was designed to collect data on a range of topics associated with extending the communication and scientific collaboration options available to NAI members. This included information about the geographical distribution of each team, the kinds of collaborative activities members would like to engage in, identification of key collaborators from the point of view of the survey respondent, and specifics about the particular IT equipment available to each member given their locations.

  Data from the survey was analyzed together with ethnographic data previously elicited through interviews and site visits with the NAI PIS and many other team members. The result was a report issued to the NAI Director in April 2002 intended as a guide to further planning and implementation of collaborative research support.

- Completion of Comparative Study of Collaborative Tools and Technologies:
  Guided by the findings of the NAI Needs Assessment regarding the types of tools and technologies prioritized as most important by NAI members, L. Faithorn worked with subcontractor C. Whitmyer of FutureU in San Francisco on a comparative study of potentially useful collaboration tools being offered by leading edge industry developers. The assessment used NAI technical and user requirements to identify the most promising solutions to two main categories of collaboration tools: real-time web-based meeting software, and document/data sharing knowledge management collaboration software. The study, completed in March 2002, recommended several tools in each category for further exploration through demonstrations by vendors and additional research regarding the range of features each tool offered.

- Completion of Collaborative Tools Demonstration Phase:

L. Faithorn, in collaboration with subcontractor C. Whitmyer of Future U, organized vendor demonstrations for NAI Senior Management and members of the NAI Collaborative Research Support Group during May, June and July of 2002. Three real-time meeting tools and four software packages offering document/data sharing, knowledge management and collaboration software were identified as potentially meeting most or all of NAI's requirements. Through the Demonstration Phase this was narrowed to one tool in each category that best met NAI's needs. Both were recommended for 6 month pilot studies with the NAI community. Both tools are currently in the procurement phase, with pilots expected to launch in October, 2002.

- NAI Video Seminar Series:
  L. Faithorn, working with others in NAI Central, led the effort to initiate two new series of NAI video seminars, the Director's Seminar Series featuring presentations by senior scientists, and the Astrobiology Research Video Forum, showcasing presentations by graduate students and Post Doctoral Fellows. Both series were piloted in the fall of 2001 and fully launched early in 2002 utilizing webcasting to the desktop, as well as the room-based Polycom systems at lead institutional sites.

- Extension of Videoconferencing beyond the NAI Lead Institutional Sites:
  The NAI membership is widely distributed beyond the Lead Institutional Sites. All 15 teams have members at many institutions. However Polycom videoconferencing systems were only deployed to the 15 lead sites and thus accessible only to the PIs of each team and those other team members who worked at these lead sites. This led to an experience of non-inclusion in the NAI virtual institute on the part of some members, and to only minimal NAI support of collaboration needs within teams.

  As a result of this feedback, L. Faithorn has been working with PIs and CO-Is to extend access to the videoconferencing system. One successful project was the deployment of additional Polycom systems to NY State University at Stonybrook and the University of Pittsburgh, two key collaborators of the Penn State University team. This extension has enabled the three institutions to offer joint courses in astrobiology with faculty and student participants from each institution.

  Another project has been to make Via Video desktop videoconferencing systems available to various NAI Co-Is. This system runs only on Windows, but does interface with the Polycom system, providing videoconferencing accessibility to distributed team members. Currently, the entire Michigan State University team, all of whom work on Windows PCs, are being equipped with Via Video systems in order to support their need for frequent all-team meetings.

- Continued Support of Scientific Collaboration Projects:
  L. Faithorn continued support of ScienceOrganizer as an experimental tool for selected NAI groups. ScienceOrganizer is a specialized web-based project information repository designed by Rich Keller of ARC Code IC to facilitate collaboration among

members of distributed science teams. L. Faithorn continued her collaboration with Rich Keller to expand the use of ScienceOrganizer by NAI members.

L. Faithorn and other members of the NAI Senior Management also recently began a series of meetings with Code IC leads to further explore additional information technology support for specialized scientific work by NAI members. The outcome of these meetings is a planning retreat now being scheduled for early December, 2002.

- Team Building and Skills Development of the NAI IT Working Group
  L. Faithorn organized and hosted a 4 day conference at ARC during the Astrobiology Science Conference in April, 2002, for the NAI IT Working Group. This group is made up of the IT Points of Contact for each of the 15 NAI teams and those at ARC from NAI Central and Code JT who work together as the NAI Collaborative Research Support Group. 10 sessions were organized for the group consisting of presentations, technology hands-on demonstrations and visits to ARC sites. This was the first time this whole group had come together in a face to face meeting and was significant in further developing a sense of a working virtual team.

  L. Faithorn proposed the NAI IT Working Group as a participant to the eNASA collaborative tools pilot and it was selected as one of 3 ARC teams for the pilot. This has resulted in access to 2 collaboration tools for the group to learn to utilize, WebEx and eRooms.

### II.D.1.e. Problems Encountered and Possible Resolution

- Videoconferencing
  A problem encountered through continued use of the Polycom videoconferencing system is the rising cost of this collaboration tool, primarily due to ISDN maintenance costs and long-distance charges. Transition to IP is now being pursued and tests are currently being conducted with all NAI Polycom sites to see where this is feasible.

  A potential problem with the Polycom system is that although the number of videoconferencing participants within NAI continues to expand, there is a limit to the number of participants the MCU can support in a multipoint meeting, This limit has not yet been reached, but research on desktop web-based videoconferencing solutions is being pursued.

- Adoption of new tools
  A number of NAI members still haven't mastered the use of the videoconferencing and data sharing systems that have been in place from the beginning at the lead institutional sites (Polycom and Netmeeting) and most rejected PostDoc as a document management system, which has since been terminated as an NAI tool. Two new collaboration tools are to be piloted beginning in the fall of 2002. Both are tools specifically identified by NAI members as important and useful to the virtual institute technology infrastructure. Nevertheless, it is always challenging to introduce new tools and technologies to a busy user group who already have established work practices.

A potential problem is that NAI members will resist learning to use these new tools. Careful plans are being developed for piloting the tools so that people will experience early success and time savings.

### II.D.1.f. Publications and Presentations during FY2001

#### II.D.1.f.i. Publications

96. Faithorn, Lisa, " The Challenges of Building Shared Knowledge and Collective Memory among a Distributed Group of Scientists ", paper to be presented at the American Anthropological Association Annual Meeting, New Orleans, November, 2002.
97. Faithorn, Lisa, "Communication and Collaboration in a Virtual Institute: NAI Needs Assessment Report", presented to NAI Director Baruch Blumberg in April 2002 and published on the web, in September 2002, as an interactive document at http://ps.pageseeder.com/ps/eval/futureu/Needs_Assessment_Final_Report/report_needs_assessment.html

#### II.D.1.f.ii. Presentations

74. Faithorn, Lisa, with Charlotte Linde, panel organizers and co-chairs, " Mechanical Memory, Mechanisms of Memory" forthcoming at the American Anthropological Association Annual Meeting, New Orleans, November, 2002.
75. Faithorn, Lisa "The Challenges of Collaboration", Presentation to the Fulbright Conference for Science Teachers, BioAstronomy Conference, Hamilton Island, Australia, July 2002.
76. Faithorn, Lisa, " Virtual Institute Directions, Solutions and Tools", Presentation to the NAI Executive Council, Ames Research Center, April 2002.
77. Faithorn, Lisa, "NAI: an Experiment in Designing a Virtual Institute" Presentation to NAI IT Working Group Conference, Ames Research Center, April, 2002.
78. Faithorn, Lisa, "Overcoming the Social and Technical Challenges of Virtual Scientific Collaboration" Presentation to Collaboration on Collaboration Working Group, Ames Research Center, April 2002 (with Claude Whitmyer and Gail Terry Grimes of FutureU).
79. Faithorn, Lisa, "Overview of Needs Assessment Research Findings" Presentation to the NAI Executive Council via videoconference, March, 2002.
80. Faithorn, Lisa, "Building Shared Knowledge", Presentation to NASA Knowledge Management Group via videoconference, February, 2002.
81. Faithorn, Lisa, "Update on the Virtual Institute", Presentation to the NAI Executive Council, San Francisco, December 2001.

#### II.D.1.f.iii. Other Activities

- Advisory Board member and ethnographic advisor to project exploring NSF funded collaboratories by academic researchers at Hybrid Vigor, a research organization in San Francisco

- Advisory Board member, CollaborateWest, an organization that promotes conferences on collaboration and the showcasing of collaborative tools and technologies
- Adjunct Faculty, California Institute of Integral Studies, San Francisco. Committee member on thesis and dissertation committees of MA and PhD students in Cultural Anthropology
- Member of Ethnography Working Group, a group of trained researchers working in private industry and government who meet monthly to share knowledge and expertise on the use of ethnographic methodologies in applied settings
- Associate, O'Neill and Associates, Sacramento, occasional consulting on healthcare-related assessment and training projects

#### II.D.1.g. Future Plans

- Full deployment to the NAI community of desktop communication and collaboration knowledge management, and collaborative work spaces.
- Creation of pilots focused on specific NAI member scientific collaboration projects in collaboration with colleagues from Code I and other Ames colleagues.
- Continued research and publication on the elements of a robust virtual institute promoting collaborative science and the technologies and practices supporting effective virtual teams.

## II.D.2. Bioinformatics Infrastructure in Support of Biological Research

#### II.D.2.a. Task Summary

The objective of this task is to provide a bioinformatics infrastructure in support biological research at NASA. The goals are to establish a physical infrastructure to support genomics research, as well as tools and data required to further NASA's lead roles in Astrobiology, Fundamental Biology, and Origins of Life research. Establishing a dedicated computational resource (including a Linux cluster, web servers, and database servers) is a primary goal. In addition, the relevant tools and data, including those specifically developed for NASA, will be deployed. Lastly, an integrated collaborative tool will be developed which will allow researchers to share data, and analyses, and provide a convenient interface to many tools and information.

#### II.D.2.b. RIACS Staff

Karl Schweighofer
Serdar Uckun

#### II.D.2.b.i. Visiting Staff and Students

Sabine Gounder (intern)

#### II.D.2.c. Project Description

- Purchase and configure Linux cluster, web servers, database server, fileserver.
- Configure Web servers, and Firewall.
- Configure PBS batch queue.

- Configure backups, password management, user accounts.
- Install genomics tools, and data.
- Write and configure automated client to regularly download genomics data from public sources.
- Write JSP or CGI interfaces to commonly used genomics tools.
- Deploy microarray database and tools.

### II.D.2.d. Accomplishments during FY2002

- Linux cluster and web servers configured and operational.
- User accounts set up and handled through Code IN.
- Backup policy and methods initiated.
- Security audits performed (Code IN).
- PBS server and clients installed.
- Development web server configured.

### II.D.2.e. Problems Encountered and Possible Resolution

- Security concerns prevent some software from being deployed as is. Remedy is being worked out to reconfigure this software to use ssh instead of rsh.
- NASA policies regarding data confidentiality require us to implement two independent web servers, one for SER data and one for public data. Some public tools must be rewritten so that they do not require a username or password. Guest or dummy virtual accounts are prohibited.
- Compute nodes have to be rebuilt because the particular configuration provided by the vendor was unsuitable.

### II.D.2.f. Publications and Presentations during FY2002

#### II.D.2.f.i. Presentations

82. Schweighofer K., Graul R., Pohorille A., "A High Availability Linux Cluster for Genomic Computation", *Astrobiology Workshop Poster Session*, Ames Research Center, 6/02.

#### II.D.2.f.ii. Other Activities

- Advised and supervised interns on bioinformatics projects.
- SBIR review panel.
- Received part of the first year funding for a DDF.

### II.D.2.g. Future Plans

- Continued development of the NASA bioinformatics web site
- Configuring genomics tools on the cluster to use PBS.
- Write grant proposals.
- Start work on a funded DDF (Schweighofer, Cullings).

# II.D.3. Advanced Visualization and Collaborative Virtual Environments for Medical and Scientific Imaging

## II.D.3.a. Task Summary

Onboard the International Space Station, complex life science experiments will address long-standing questions concerning life's ability to adapt and respond to the space environment. Many of these experiments will require astronaut intervention and the use of the Life Sciences Glovebox Facility (LSG). Within the LSG, astronauts manipulate scientific instruments, conduct experimental assays, collect tissue specimens and perform delicate dissections—all under highly controlled conditions and while physically tethered to minimize drift and within strict time constraints. The experiments often demand very detailed training and knowledge of instrumentation, surgical anatomy and specific scientific objectives. To be successful, astronauts must remain highly proficient in these experimental techniques, but due to scheduling constraints they can receive only limited Earth-based training with real LSG mock-ups and real experimental specimens. Because of mission costs, crew safety, and the immediate and far-reaching impact of the scientific results, it is imperative that novel strategies be developed to maximize the use of the LSG Facility. To this end an immersive virtual environment simulation system called the "Virtual GloveboX" (VGX) is being developed at NASA Ames Research Center. Astronauts and support crews can use the VGX for early engineering development and experiment planning, and later advanced training activities that require the LSG. The system integrates off-the-shelf hardware and software with new real-time simulation technologies to provide a realistic 3-D virtual environment, mimicking the real LSG environment. Within the VGX, the force of gravity can be turned on and off, allowing astronauts to perform experiments in a simulated microgravity environment while still on Earth. Further, a continuum of scenarios from ideal conditions to catastrophic experiment failures, such as contamination, can be introduced into the training session to allow the astronaut and principal investigator to assess critical decisions. The VGX does not eliminate the need for direct, hands-on training and experiment planning using physical mock-ups. However, it can streamline these processes, reduce the need for full-scale training sessions with live animals, increase the frequency and ease of the principal investigators input to their specific experiment, and provide astronauts with a means to keep their skills sharp on Earth and in space, thereby maximizing the efficiency and success of ground-breaking biological experiments in space.

## II.D.3.b. RIACS Staff

Xander Twombly

### II.D.3.b.i. Visiting Staff and Students

Rose Mills, M.S. Candidate, University of Utah

## II.D.3.c. Project Description

- Development of Virtual GloveboX system (VGX). This has been an ongoing development of an immersive virtual environment designed to aid in the training of astronauts to perform work in the Life Sciences Glovebox. The VGX is designed to create a virtual 3D simulation/visualization environment using hardware and software components developed from multiple sources. A physically based modeling/simulation engine is coupled with high performance graphics to render the virtual world, and multiple peripheral devices (PHANToM force-feedback device, Cybergloves for hand tracking) are used to interact with the world. The VGX display system is a hardware/software combination designed to act as a front end to different simulation engines, providing the flexibility to use the most efficient simulation engine for a specific task.

### II.D.3.d. Accomplishments during FY2002

- VGX display system with LCD projectors – created a new VGX display system to replace the original CRT projectors which failed to meet the project specifications.
- Geometry replication – streaming geometry system for moving data from a 3D simulation system to the graphics computers driving the VGX display system.
- Synchronization manager – system to link two (or more) PC's together so that their graphics scenes are sent to the display simultaneously. This is used to drive the dual projectors for the 3D stereo image in the VGX.
- Environment manager – sends internal information such as keystroke and mouse information between the systems linked with the synchronization manager. Basic bookkeeping to maintain identical graphical environments between systems.
- Display replication – internal system for rendering scenes on one computer, and then streaming them to another for display. This was used to test the initial VGX display system before the geometry replication system was built. Also used to send streaming video images across the network to allow viewing a stereo image set generated from two cameras.
- Multi-camera capture and analysis – system for capturing video from multiple cameras, and placing each frame into an analysis environment to perform segmentation and other visual analysis – based on the Intel OpenCV libraries. Basic system allowed camera calibration, segmentation based on luminance, and segmentation based on hue (for segmentation of flesh colored objects).
- Render-to-texture/display distortion – method for adjusting the image from one LCD image to get it to overlay the image from the second LCD projector in the current VGX system configuration.

### II.D.3.e. Problems Encountered and Possible Resolution

Construction of the VGX display system included the creation of a projection system and the physical superstructure resembling the life sciences glovebox. The specific construction of the system was contracted to BARCO Projection Systems. Original delivery date of the display system hardware was late July, 2001. BARCO actually delivered the system in Mid-January, 2002. After extensive in-house testing and evaluation of the system that was delivered (the 3D functionality never worked), the system was rejected by NASA and the projectors for the system were returned to BARCO for failure to meet the specifications in the contract (late March, 2002).

Since returning the projectors to BARCO, we have been attempting to incorporate off the shelf LCD projectors to perform a similar role, with limited success. The system used dual projectors whose images must be overlayed and aligned exactly pixel-for-pixel to create an appropriate 3D stereo rendering system. Unlike the CRT projectors that were returned, the LCD projectors do not allow for the spatial modification of their images to account for distortions of the image due to screen orientation and projection orientation (basically, the CRT allowed one to square the image up on the display screen whereas the LCD's do not). In order to overlay the signals, we resorted to creating a software solution that allows a limited degree of the hardware functionality of the CRT projectors to be duplicated, providing the ability to create a slight distortion in the displayed image of one projector that can be used to overlay the image with the second projector. This technique could be refined considerably to improve accuracy, but other issues with the system are more pressing at this time. The major disadvantage to the system is rendering performance – we are currently limited to ~7 frames per second, which is well below the desired rate of 15-20 fps. The bottleneck is in moving the large amounts of data from the texture buffer, and cannot be resolved other than waiting for faster hardware to become available. Mechanical alignment systems are being investigated, but it is unlikely that the current project budget will allow such solutions to be implemented in the near term.

### II.D.3.f. Publications and Presentations during FY2002

#### II.D.3.f.i. Publications

98. Smith J.D., Twombly I.A, Bruyns C, Boyle R NASA Virtual GloveboX (VGX): Training and Simulation for Biology Research Aboard the International Space Station. *AIAA Conference on International Space Station Utilization*, Kennedy Space Center, FL, October 15-18, 2001 (AIAA paper 2001-5105).

### II.D.3.g. Future Plans

The focus of the BioVIS technology center may be shifting away from the Virtual Glovebox project due to funding constraints. Work in progress is the integration of a new rigid body dynamics engine (from Arachi, Inc) to act as the underlying simulation engine in the VGX system. This simulation engine provides a physically based modeling system of rigid-body dynamics and force based interactions in the virtual world, and is ideal for the procedural tasks (experiment setup and design, manipulation of lab equipment in microgravity) that will be placed within the VGX. Simulations allowing the manipulation of rigid objects such as test tubes, slides, microscopes, etc. will be designed to allow astronauts to practice basic experimental procedures, and also allow principal investigators to test their own experimental design prior to submitting for flight testing and approval. In addition, an effort has begun to incorporate simulation of small robotic arms into the VGX environment to experiment with the feasibility of using telerobotic manipulators to perform simple experiments within the actual Life Sciences Glovebox. A grant has been submitted to incorporate the simulation of a single robotic manipulator into the VGX environment, controlled by a remote user with a device such as the PHANToM forcefeedback device. If this grant is awarded we will create a series of simulations to study the feasibility of performing telerobotic manipulations and experiment control in a microgravity environment, and study the need for human interaction with the robotic proxies that can help to increase the science yield from experiments performed in the Life Sciences Glovebox. An additional project that we will be undertaking is to begin the development of the EAR-BOT. The EAR-BOT project is in an early conceptual stage at this time, but aims to model the fine control mechanisms of the neural vestibular system, and use them to create an adaptive controller for an autonomous vehicle.

Initial efforts in this project will be to simulate the responses of afferent fibers to rotational and translational accelerations, and use these responses to control a simple feedback mechanism such as gaze stabilization. This project is subject to the development of further funding sources as well, and will be pursued on a part time basis while additional funding is sought.

## II.D.4. Astrobiology , Fundamental Biology and Bio-Info-Nano Strategic Planning

### II.D.4.a. Task Summary and Overview

Accomplished under separate tasks, the lead is engaged in strategic planning in several distinct but related biology programs. The principal efforts are concerned with Astrobiology strategic planning in the Astrobiology Integration Office (Code S), Cell and Molecular research new starts under the auspices of the Fundamential Biology Office(Code U) and the execution of the Bio-Info-Nano and NASA Missions workshop with the Directors Code D, NASA Ames.

### II.D.4.b. RIACS Staff

Kathleen M. Connell

#### II.D.4.b.i. Subcontractors

Dante Productions
JMP Associates
Toffler Associates

### II.D.4.c. Project(s) Description

For over 15 years, NASA/Ames has explored the potential of "innovation at the intersection of disciplines" as a possible strategic thrust. In 1995 Connell, in conjunction with NASA collaborators, co-created a team at Ames which gave scientific structure to this organizing principle, in the creation of the discipline now known as astrobiology. This multidisciplinary program seeks to answer fundamental questions about the origin, evolution and distribution of life in the universe, and has become a major scientific thrust for the Agency.

With Respect to Fundamental Biology, the 2002 thrust centered on the definition of a new program start in cell and molecular biology named “Generations”.

Of equal importance was the execution of the Bio-Info-Nano workshop, described in the introduction (section I.B). This workshop was multi-disciplinary, multi-organization and designed to describe the potential for ‘fusion’ as well as insertion of advanced technologies into NASA Missions.

### II.D.4.d. Major Accomplishments during FY2002

The Astrobiology Integration Office continues to serve as a focal point for NASA HQ, and as a launching point for new initiatives as well as a communications hub, via the Astrobiology Portal, which the incumbent oversees (see http://astrobiology.arc.nasa.gov). A major effort in 2002 was to bring on staff and assets in the creation of the Astrobiology magazine, which was fully incorporated into the portal this year. The ‘e-zine’ has driven traffic on the portal site to over 2 million visits per month, and provides detailed feature stories about new findings in astrobiology. The virtually managed portal includes staff from around the United States, and is constantly seeking to define the state of the art in communicating with multiple publics, including the general public, the science community and policy makers interested in this subject.

The AIO remains the home for those in the community interested in inquiring in the societal domain, and a second Astrobiology and Society conference is in the planning stages for 2003. 2002 was a challenging year, as the lead of the office assumed acting Deputy Directorship of 'Code S' at NASA Ames. In turn, as the senior professional in the office, Connell functioned as the day to day 'glue' who brought the team together on a regular basis to provide cohesion during a time of transition.

No less challenging was an effort to both assess and advocate for a new start in the Fundamental Biology arena. The proposed new start known as "Generations" was successfully included in the Presidents Budget in 2002 under the title: Generations. Connell led Toffler and Associates, and JMP and Associates thru a series of studies intended to both analyze and defend the study of cell and molecular functions in the space environment. Their eminent work assisted in the securing of the new start placement in the NASA research queue. See: http://generations.arc.nasa.gov

In addition to these complex strategic engagements with both consultants and NASA, the first film on the topic was finished in 2002 and is available on DVD to educate on the nature and potential benefits of Generations with respect to health, homeland defense and a myriad of applications to space missions and early science on the International Space Station.

Connell was awarded the NASA Public Service Medal in 2002 for "outstanding leadership in the development of broad community, state and national support of NASA's life sciences and astrobiology programs." She remains a Ph.D. candidate in good standing at the Fielding Graduate Institute, and completed 3 class modules in Organizational Leadership, Global Studies and an overview of Systems theories and practice, which she applies to her NASA projects on an on-going basis.

### II.D.4.e. Future Plans

Kathleen Connell was granted permission to devote half of her time towards completion of her Ph.D. by her NASA task sponsors with concurrence from the RIACS Director. The remainder of her time will be devoted primarily to strategic planning tasks with the above mentioned offices and RIACS duties.

## II.D.5. Integrated Human/Robotics Exploration & Astrobiology

### II.D.5.a. Task Summary

Increase USRA participation in the planning and execution of the Integrated Human/Robotics Exploration Program. Assist in the establishment of a HEDS Human Exploration Science Community to facilitate greater academia involvement in HEDS exploration efforts

### II.D.5.b. RIACS Staff

Lewis L. Peach, Jr. (USRA HQ)
Mike Duke (consultant)

### II.D.5.c. Project Description

There is a growing awareness that the integration of human exploration and robotics space sciences goals and objectives will result in a stronger, more comprehensive and productive exploration program that will lead to a time-phased, balanced set of objectives to drive future robotics missions and enable human exploration beyond low Earth orbit. This requires the collaboration of two formally somewhat diverse communities, in the development of an integrated set of objectives and requirements, and in the participation of a process to implement these

Activities undertaken under this task include:

- Overall lead for Major USRA/NASA Human-Robotic Exploration Study
- Member of various Mars Architecture Development Committees
- Member of International Mars Exploration Working Group (IMEWG)
- Member of Mars Exploration Program Analysis (Advisory) Group (MEPAG)
- Member of the Mars '01 Program Science Working Group
- Conducted Major NASA/USRA Human Enabled Science Workshop (Duke)
- Initiated and Supported NRC Study to Define Human Exploration Requirements
- Co-lead in the development of a program to effect collaboration between Russia and USA on Human Exploration (ISTC 1172)
- Formal Collaborator on Follow-on State Department Funded Russian Exploration Study (ISTC 2120)

### II.D.5.d. Accomplishments during FY2002

- Served on the MEPAG through a crucial series of science community studies to support NASA planning for Mars exploration for the decade beyond 2009, leading to a Mars sample return mission in 2013.
- Co-lead for a series of workshops and technical exchanges between the Russian Space Agency and NASA to develop an understanding of the two countries' different approaches to human exploration, resulting in a final report (ISTC 1172, Jan.'01). .Co-led U.S. effort, as formal USRA Collaborator, in definition and planning for follow-on Russian Exploration Study, which has been approved for funding by the U.S. State Department (ISTC 2120)
- Completed the final report of the USRA/NASA Human Robotic Exploration study for NASA's Revolutionary Advanced Systems Concepts (RASC) Office at LaRC. This involved an RFI, conducted by NIAC, and a major Workshop at ICASE, to bring together top planetary scientists, human exploration engineers, roboticists, and innovative advanced concepts specialists and technologist to evaluate the potential advances in the human machine partnership of the future, determine the science that would be enabled by this partnership, and identify critical technologies which should be made to make these advances possible.

- Served as members of the Steering Committee for the RASC-AL program, an academic outreach program that will facilitate collaboration between university students, and their faculty, funded by NASA's RASC Office at LaRC (RASC-AL) and conducted by the Lunar and Planetary Institute and ICASE. First university forum planned for November, 2002.
- Supported NRC study to define critical human exploration requirements which must be validated robotically at Mars, prior to the initiation of human missions (NRC report "Safe on Mars).

### II.D.5.e. Problems Encountered

- Mars Robotics Program re-planned following loss of Mars Climate Orbiter and Mars Polar Lander in 1999, resulting in loss of near-term flight opportunities for instruments selected for the '01 and '03 Mars Mission opportunities.
- NASA Budget uncertainties, resulting from recent development problems of the ISS.

Possible Resolution

- Participating in the reformulation and execution of the Mars Exploration Program, continued involvement in the on-going Mars Odyssey Mission, overseeing completion of the various technical efforts to prepare for future flight opportunities, and continuing the development of a HEDS Science Community and incorporating these efforts with the integrated Human /Robotic Exploration Programs.
- Broaden the NASA's support for these efforts by furthering collaboration between additional, new, NASA sponsors, and members of the broader academic community.

### II.D.5.f. Publications and Presentations during FY2002

Co-authored various Proceedings, Workshop Reports, and Requirements Documents resulting from the above activities.

### II.D.5.g. Future Plans

Continue these tasks as initiated this past year, with greater emphasis on:

- The development of a HEDS Science Community, with a workshop dealing with the importance of analog studies in defining future human exploration objectives and approaches.
- Stronger involvement of the academic community in these efforts (ongoing).

## II.D.6. Information Systems Technology and Program Management in Support of Life Science Enterprise

### II.D.6.a. Task Summary

This task is to provide support in the areas of information management and program strategy for the Biomolecular Systems Research program and the Fundamental Biology program.

One thrust of the Fundamental Biology program is to identify and develop technologies necessary to support disease risk models. This initiative advocates use of an eco-information approach to provide the human health research and end user community with information on which to base surveillance, risk assessment, and intervention strategies.

The initiative has two foci: development of knowledge-based models, and the technologies necessary to support those models. Models of disease may focus on clinical or medical aspects of a pathogen, its etiology, or changes in its spatial or temporal patterns. Investigators in each area have developed their own protocols for data collection, storage, and analysis, depending on the disease they wish to model and the scale at which the model operates. When attempting to combine or use existing data it is essential to understand why, where, and how they were collected. Data collected to explain a spatial pattern may be incompatible with that collected to explain a temporal pattern in the same variable. Likewise, the spatial patterns seen in specific disease indicators may represent subtle interactions among several variables. Finally, there may be multiple factors influencing a single variable at different scales in different areas. The science of informatics provides an approach to addressing some of these issues.

The related Biomolecular Systems Research program is an integrated research program focused on developing molecular level technologies to monitor cellular signals and processes with applications to crew health, safety, basic biology research, life detection, planetary protection and nanotechnology. The goal of the program is to develop biomolecular sensors that will revolutionize the practice of medicine on Earth and in space. NASA and the National Cancer Institute are jointly sponsoring and developing technologies in the areas of: biomolecular signatures, signal amplification, biomolecular sensing and manipulation, biomolecular imaging and informatics, and integrated biomolecular systems.

### II.D.6.b. RIACS Staff

Linda R. Andrews

### II.D.6.c. Project Description

Domain modeling and information management support will be provided to the cooperative research projects currently underway between NASA Ames and the National Institutes of Health under the Fundamental Biology program. The goal is to adapt and implement novel NASA information technology tools in highly distributed research environments.

Program strategy expertise will be provided to the related Biomolecular Systems Research program office to assist with selection of program scope and research investment portfolio in the area of biodefense. The goal is to gain insight into currently developed biodefense technology and determine the utility of applying appropriate Ames technology. The long-term vision is to design a targeted technology program that includes collaboration with entities designing desired biosensors.

### II.D.6.d. Accomplishments during FY2002

Although the bulk of this task will begin 10/1/02, some progress has already been made:

ScienceOrganizer, developed by Code IC, has been customized for use by a NIH malaria vector research team to unite distributed sites including Tulane University, SUNY, University of Oklahoma, Kenya Medical Research Institute, International Centre of Insect Physiology and Ecology, and field sites in Kisumu and Malindi, Kenya. Researchers at each of these sites can collect and load data on geography, socioeconomic levels (household surveys, images), mosquito and larvae sampling (including pond features), meteorological station readings, etc., onto the ScienceOrganizer server where they can immediately compare data from distant field sites. Currently the team is working to understand how the cycle of disease transmission is altered in urban environments consisting of larger human populations.

In addition to viewing field data, the tropical disease researchers use the tool's intuitive navigation features to learn about the individuals who collected the data (by reading graduate theses, for example), view images of the various research sites, read previous papers posted by the work team, and anything else the users/PIs would like to share.

The ScienceOrganizer tool was customized primarily by modeling the research domain in a fashion that allows the team to use the tool with minimal training or observation. The diverse datasets are now being uploaded by beta-testers at Tulane University and Kisumu, Kenya. Full deployment to all sites will begin this fall.

NIH is interested in the study of numerous infectious diseases in endemic countries because it is essential to understanding how to respond to the intentional introduction of a parasite into a non-resistant population.

This work has been selected for a NASA press release in the fall 2002.

### II.D.6.e. Problems Encountered and Possible Resolution

None

### II.D.6.f. Publications and Presentations during FY2002

None

### II.D.6.g. Future Plans

Tasks to begin 10/1/02:

- Participate in development of miniature, field-deployable water quality instrument.
- Perform requirements analysis of malaria vector research teams for use of various data management tools. Begin with Kenya team, then expand to use by Mali team, if appropriate.
- Explore the deployment of information sharing and analysis tools to the World Health Organization (WHO).
- Facilitate customization and training of collaborative research software tools with infectious disease research teams.
- Determine feasibility of collaborative research on risk assessment for infectious disease (including the intentional release of disease-causing agents) with the Center for

Nonproliferation Studies at the Monterey Institute, Lawrence Livermore National Laboratory, and Stanford University.

- Survey the landscape of biodefense technology development and identify relevant gaps.

## II.D.7. Earth Science Technology Office Support

### II.D.7.a. Task Summary

Serve as the Ames representative for the Earth Science Technology Office. Develop technology roadmaps and plans to support mid and far term earth science mission requirements. Communicate information and opportunities to Ames researcher. Serve as technical reviewer for various NASA Grants and Contracts. Keep Ames staff informed of Research Opportunities.

### II.D.7.b. RIACS Staff

Walter F. Brooks

### II.D.7.c. Project Description

The goal of this project is to insure that NASA Ames IT technology expertise is brought to bear on high level planning as well as specific technology projects initiated by the Earth Science Technology Office. The task requires weekly support of ESTO staff meeting, AIST staff meetings and biweekly ESTO TST meetings.

### II.D.7.d. Accomplishments during FY2002

- Serve as ESTO Alpha West representative to the ESTO Office.
  - Attended and communicated to Ames important results from Weekly staff and Monthly Office Reviews
  - Support ESTO Monthly project reviews-work with Ames and University PIs to develop project status reports
- Develop Technology Roadmaps for the ESTO Office
  - review and improve existing roadmap baseline
  - chaired session at Technology Workshop
- Inter-center Coordination in Preparation for Funding Opportunities
  - Initiated and ran a series of meetings over a 4 month period to inform the broad Ames research community of two Earth science research opportunities (AIST NRA and HQ "Reason" CAN). Brought together researchers from codes SG, SS, IN and IC to brainstorm ideas, developed a series of whitepapers. When the NRA and CAN were released analyzed and interpreted the contract vehicles and held meetings to suggest approaches and answer questions on appropriate responses.
- Technical Analysis and Assessment of Research
  - Technology Assessment of ongoing contracts managed by ARC with the University of Alabama and University of Arizona as well as one in-house effort. Arizona work is centered on the development of high density on board storage media. There are two efforts at University of Alabama. The first involves the development for SensorML and markup language for remote sensing instruments

and the second effort is the development of on board software for control and processing of science data. These involve monthly status reports yearly reviews. The Ames in-house effort is centered on on-board planning and scheduling. All 4 efforts have been approved for a 3rd year of funding based on successful completion of year 2 milestones.

- Support of Agency Technology Strategy Team TST
  - Organized Technology Strategy Team Site visit to ARC and Napa to conduct site visit of the remote sensing GRAPES project. Set up reviews and tours as well as dinner and invited speakers.
  - Attended Site visits to GRC and GSFC as part of ongoing reviews
  - Represented ARC at Monthly reviews
- Management Coordination
  - Met regularly with Branch Chiefs Division chiefs and appropriate group and project leads to maintain cognizance of earth science activities and opportunities.

### II.D.7.e. Problems Encountered and Possible Resolution

- University of Alabama effort on developing a sensor markup language required a major re-alignment of milestones for the 3rd year. Worked with the NASA and University to develop acceptable technical milestones and schedule.

### II.D.7.f. Publications and Presentations during FY2002

#### II.D.7.f.i. Presentations

83. In house seminar on Earth Science Technology Program

### II.D.7.g. Future Plans

- Continue task ARC member of the ESTO Team. Increase awareness of Ames within ESE. Help architect MOU between Ames and ARC. Assist in review of proposals and technical evaluation of the program. Continue the work of refining requirements through workshops and conferences.

## II.D.8. Integrated Thermal Protection

### II.D.8.a. Task Summary

The development of future Space Transportation Vehicles is dependent on the development of reliable thermal protection systems that require minimum refurbishment and repair. Instrumentation that can detect, diagnose and report defects in a timely manner is required. Repairs must be performed easily and quickly on the ground and preferably automatically in Space. This research program will try to identify and develop technologies to fill these needs

The goal of this task is for USRA to establish a cooperative research program with the Thermal Protection Systems development and Information Systems/Nanotechnology research groups at Ames that will lead to the developments of Adoptive Intelligent Thermal Protection Systems for future Space Transportation Vehicle systems. This effort will define and evaluate potential approaches to detecting diagnosing and repairing thermal protection systems damage on Reusable Vehicles and HEDS Space Transportation systems. It will also provide consulting on thermal protection materials and systems and other relevant technical issues.

**II.D.8.b. RIACS Staff**

Howard Goldstein

**II.D.8.c. Project Description**

- Evaluate capabilities and the potential pay-off of coupling information Systems technologies with Thermal Protection System technologies and Nanotechnology to develop Adaptive Intelligent TPS (aiTPS). Define possible approaches to detect, diagnose and repair TPS damage due to impact from micrometeorites and other on orbit debris.
- Define and investigate new concepts for aiTPS including self-healing heat shield materials, novel approaches to instrumentation and new material/instrument combinations that will diagnose and self-repair after damage from any source.
- Participate in technical reviews for Code AS and AX and other NASA organizations and programs as requested. Consult with these organizations on program related to atmospheric entry technology.

**II.D.8.d. Accomplishments during FY2002**

We participated in the Stardust program review during much of calendar year 2001. The review was completed at a technical interchange meeting with Lockheed Martin Astronautics, December 12-13, 2001. Based on the recommendations of that program review the Phenolic Impregnated Carbon Ablator (PICA) Arc jet Test Program (PAT) was initiated. Led by Al Covington, the group planned and performed an extensive arc jet test series on PICA to determine if it would perform as planned for the Stardust vehicle and to provide data which could be used to improve the analytical model used to predict PICA'a ablative performance. Howard Goldstein was a member of this team and consulted on the test articles design and on properties of the PICA material. This test series and evaluation of the data was completed in May 2002. The results, reported to Ames management and Lockheed Martin, demonstrate that the Stardust entry vehicle will meet the mission requirements. Using the arc jet data, significant revision/enhancement of the PICA analytical model was accomplished that will be used to more accurately design future vehicles.

Howard Goldstein was asked to initiate, plan and lead an arc jet test program to evaluate the thermostructural performance and reusability of Silicone Impregnated Reusable Ceramic Ablator (SIRCA) for the planned flight environment of the Kistler K-1 commercial launch vehicle. The test program was planned and initiated during September 02. Initial test results were obtained from a test series at low pressure and high enthalpy in the AHF and another series at high pressure and lower enthalpy in the IHF. The results appear to show that SIRCA is reusable for a limited number of flights in the planned K-1 environment. A report is in process. Further testing and material characterization is planned to fully understand the materials capability. A test series in the 2X 9 duct is scheduled for late October 2002.

The DDF funded program, initiated by Mr. Goldstein to develop a self healing thermal protection system for reusable launch vehicles such as the space shuttle and advanced reusable launch vehicles was continued. We have made progress developing a modified silicone rubber adhesive that would provide thermal protection to the vehicle structure by intumescing when a tile was damaged or lost during atmospheric entry. Development of a modified strain isolator pad to provide another layer of protection has been initiated.

Howard Goldstein consulted on a variety of programs and technology issues during the year including Gen–2 RLV, Gen-3 RLV, X-33, inflatable reentry vehicles, ISP proposals, arc jet testing of several materials, comparison of arc jet testing at JSC and ARC with colleagues in codes AS, ASM. AX , NASA MSFC, several companies, the Mars Society, etc. Mr. Goldstein was a member of the Aerocapture program review team at JPL in August of 2002, and he participated in a study with B. Laub and J. Balboni on Ground Test Facilities for TPS Development for reusable space transportation vehicles and planetary probes.

### II.D.8.e. Problems Encountered
None

### II.D.8.f. Publications and Presentations during FY2002
None

### II.D.8.g. Future Plans
We will continue to the program goals as originally defined with emphasis on the research program on self healing thermal protection materials. We will support development of the reusable space transportation vehicles such as the Kistler K-1, the proposed crew transfer vehicle and other high L/D vehicles by completing the SIRCA arc jet test program and as a consultant to Ames Code A Personnel and to other NASA and outside organizations as required. We will also continue efforts on evaluating the payoff of coupling Information Systems Technologies with Thermal Protection system Technologies and Nanotechnology.

## II.D.9. Multi-institutional Collaboration on Research, Development, and Demonstrations for Natural Hazards Impact Reductions

### II.D.9.a. Task Summary
This project facilitates collaborative research and development between NASA and non-NASA organizations on natural and terrorist-produced disasters and related technologies, demonstration projects, and model development and validation. This activity complements and supports research and technology development in the Earth Sciences and the Information Sciences and Technologies at Ames Research Center. The potential applications of remotely sensed data from NASA satellites and aircraft to mitigation of disasters are investigated.

The capabilities to acquire, process, and distribute data from multiple sources, including satellites and both piloted and unpiloted aircraft, to mitigate the scope and severity of disasters have continued to expand rapidly. The Earth Sciences Division (ESD) at NASA Ames Research Center has conducted extensive disaster-related research and technology development, particularly on fires and floods for more than two decades. There has been rapid growth in the capabilities of information and communications systems to process large volumes of data, to extract actionable information, and to rapidly distribute and visually display the information to the user communities. These growing systems capabilities provide new opportunities to develop and implement the needed technologies to conduct the research and related technology demonstration projects to reduce the large fiscal and human impacts of natural and terrorist-produced disasters. The national costs of natural disasters from 1992-1997 were $54 billion per year, with about 10% of these costs occurring in California.

**II.D.9.b. RIACS Staff**

Linda Andrews
Richard G. Johnson (Consultant)

**II.D.9.c. Project Description**

This project formulates, structures, and recommends architectural and programmatic approaches for collaborative disaster-related programs and information exchanges between the Ames Earth Sciences Division, and federal and state agencies, local and regional government, industries, not-for-profit organizations, and universities.

This project also reviews and assesses the applicability of advanced information systems and wireless communication systems for near-term and future research and development activities related to disasters.

**II.D.9.d. Accomplishments during FY2002**

This activity has focused primarily on formulating and structuring potential disaster-related programs for collaboration between the Ames Earth Science Division (ESD), the Information Sciences and Technology Directorate, and other non-NASA partners, including RIACS, and the Bay Area Shared Information Consortium (BASIC).

Briefing materials were prepared and a series of meetings held with the ESD Chief, Deputy Chief, Ecosystems Science and Technology Branch Chief, and the RIACS Director to explore Disaster Infosphere concepts and architectures that were focused explicitly on direct access to disaster-related information by all users. This information-centric approach is intended to provide rapid access with minimum user constraints to dynamic information that disaster managers and other users require to make effective decisions before the disaster-related information degrades in value or becomes totally obsolete. A framework for an Environmental and Disaster Infospheres Collaborative (EDIC) was formulated and presented to the Earth Sciences Division management.

A reimbursable Space Act Agreement between NASA Ames Research Center (ARC) and the California Resources Agency (CRA) was facilitated and prepared in collaboration with the ARC Earth Sciences Division, and has been signed by the CRA. This agreement is for collaboration on research and development, testbeds, and user demonstrations relevant to California resources, disasters, and natural and human-produced environments. The two initial projects structured and defined for this agreement were the Disaster and Environmental Infospheres (DEI) project and the Data Server Utilization Testbed (DSUT) project. The DSUT project was approved for funding by the CRA. The DEI project was not funded. It was structured to use information provided by NASA, State of California, and other sources to demonstrate a prototype environmental infosphere that would enable collaborating user communities to effectively access and utilize large and diverse databases to significantly improve the collaborative research process by use of the infosphere. This project would also develop and define strategic approaches and a plan for a prototype disaster infosphere for wildfires and would use a framework based on a collaborative program to utilize a long-duration Uninhabited Aerial Vehicle (UAV) for identifying and monitoring lightning-induced fires in remote areas such as the Sierra Nevada Mountains.

### II.D.9.e. Problems and Possible Resolutions

The effective acquisition, development and utilization of large and diverse databases by a large number of diverse users remains a challenging problem. The development of new approaches and appropriately monitored multiple testbeds to access effectiveness is needed.

### II.D.9.f. Publications and Presentations During FY2002

None

### II.D.9.g. Future Plans

The plan is to continue the project into FY03 along the same lines but at an increased scope of effort. Discussions are underway with the California Resources Agency and the Office of Emergency Services and other collaborators to submit a joint proposal in response to NASA CAN-02-OES-01 that would include elements of the above DEI project.

## II.D.10. AeroSpace ExtraNet Data Sharing

### II.D.10.a. Task Summary and Overview

The current aerospace environment consists of thousands of independent entities: manufacturers, space centers, airlines, airports, service providers, research organizations, and government agencies. There is no common architecture for networks, data formats, and application interfaces. Many operational systems have been in place for over 30 years. Because of the lack of a strategic aviation and aerospace information infrastructure there are thousands of incompatible data repositories, network systems and applications. These "islands" of incompatibility not only exist between organizations, but even within organizations such as air carriers and government agencies where departments are unable to efficiently share key information between maintenance, operations, and management, etc.

The extreme fragmentation of data and application non-interoperability has prevented the U.S. from developing comprehensive, system-wide applications that can monitor, track, and evaluate the overall performance of major components in the National Airspace System. The lack of comprehensive architectures not only has prevented aerospace organizations from developing system-wide applications, but it has also limited the ability of commercial third-party information technology companies to apply their extensive resources against this environment (as compared to the financial, medical or chemical industries for example).

**II.D.10.b. RIACS Staff**
David A. Maluf

**II.D.10.c. Project Description**
AeroSpace ExtraNet InfoLab:
The objective of the (AeroSpace ExtraNet) AEN InfoLab project is to develop the infrastructure to support a System-Wide Monitoring capability of the National Airspace System and to prototype Enterprise Safety applications that leverage that Infrastructure.

The extreme fragmentation of data and application non-interoperability has prevented the U.S. from developing comprehensive, system-wide applications that can monitor, track, and evaluate the overall health and performance of major components in the National Airspace System. The lack of comprehensive infrastructure not only has prevented aviation organizations from developing system-wide applications, but it has also limited the ability of commercial third-party information technology companies to apply their extensive resources against this environment.

The Data Sharing project prototyped cutting-edge information technologies to support the evolving aviation industry NAS architecture. Key middleware technologies will help integrate the data and application interfaces by providing common machine readable handles to data and application services. The middleware components coupled with next generation physical network connectivity to key data sources (airline operations, FAA flight information, maintenance and manufacturing data) will provide the core elements that will help support enterprise-wide systems and applications.

To fully conduct research that will support the far-term concepts, technologies and methods required to improve the safety of Air Transportation, a simulation environment of the requisite degree of fidelity must first be in place. The Virtual National Airspace Simulation (VNAS) provides the underlying infrastructure necessary for such a simulation system. Aerospace-specific knowledge management services such as intelligent data-integration middleware will support the management of information associated with this complex and critically important operational environment. This simulation environment, in conjunction with a distributed network of super-computers, and high-speed network connections to aircraft, and to Federal Aviation Administration (FAA), airline and other data-sources will provide the capability to continuously monitor and measure operational performance against expected performance. The VNAS will also provide the tools to use this performance baseline to obtain a perspective of what is happening today and of the potential impact of proposed changes before they are introduced into the system. National Airspace Simulation (NAS), a capability for evaluating future concepts for the Air Transportation System, is being developed and has been demonstrated for integrated aircraft models including: engines, wings, and landing gear.

Virtual Iron Bird (VIB) Centrifuge Accommodation Module (CAM) with application to the Biological Research Project:

The VIB CAM simulation features the integration of 3D models of the CAM interior, advanced information management and knowledge engineering tools, communications, and immersion technology to create a VIB for the CAM.

This project will demonstrate proof-of-concept research on system-wide infrastructure to aid in the projection of effects of decision-support when Courses-of-Action (CoA) and assessments in the National Airspace System is needed by the Federal Aviation Administration (FAA). The objective is to significantly augment FAA decision services by making RADAR TRACON information accessible as well as other information components for fast-time distributed advanced analysis to be utilized for decision making across the U.S continental FAA centers and Air Route Traffic Control Centers.

Performance Data Analysis Reporting System:
PDARS is a joint research and development project between NASA and the FAA to continue support FAA R&D needs. NASA will provide a system-wide network infrastructure (design, implementation, deployment, and support) across the continental U.S. covering FAA centers and Air Route Traffic Control Center (ARTCC) sites. The PDARS project is an FAA mission critical infrastructure requiring many applicable security features, accessibility, and restrictions—both on the network implementation and handling data. This PDARS environment is developed under Aerospace ExtraNet within NASA Information Technology Directorate, Computational Sciences Division to allow the deployment advanced networking and data analysis technology to a system-wide level capability.

NETMARK:
Object-Relational database management system (ORDBMS) is an integrated hybrid cooperative approach to combine the best practices of both the relational model utilizing SQL queries and the object-oriented, semantic paradigm for supporting complex data creation. NETMARK provides an extensible, clientless, schema-less, information on demand framework for managing, storing, and retrieving unstructured and/or semi-structured documents. It was originally developed in early 2000 as a research and development prototype to solve the vast amounts of unstructured and semi-structured documents existing within NASA enterprises. NETMARK takes advantages of the object-relational database model using very efficient keyword searches spanning both content and context.

Program Management Tool:
Within each NASA enterprise, there exists a complex hierarchical structure of program and project line management in place to supervise and oversee the progress of various definitive milestones objectives, goals, requirements, and their respective resources, such as personnel and budgetary allocation. With the increased complexity of managing, tracking, and defining milestone requirements and objectives, there are inconsistencies and uncertainties produced in the accountability associated with the research and development objectives and goals. There may be unintentional redundant requirements, goals, and objectives existing across the various NASA enterprises, resulting in inappropriate management of both funding and personnel resources. At the same time, the current existing monthly status reports for most programs are cumbersome to

generate by hand using proprietary Microsoft-based products, such as Powerpoint presentation slides, Word processing document, and/or Project planner. This usually results in tediously drawing by hand the various tables, color scheme, and icons on presentation slides and word documents. Another problem encountered is the incompatibility among the various Microsoft product versions; e.g., MS-Office 2000 is not backwards-compatible with some features of Office 97. The initial version of the PMT tool was not planned to have any automated interfaces to other center management information systems; but the PMT tool will be designed to be flexible, extensible, and be able to integrate with existing and/or new strategic center and program management information systems.

**II.D.10.d. Accomplishments during FY2002**

- **Middleware Development for Information Management.** Middleware in the context of data sharing (vs. knowledge sharing) focuses on semi-structured data packets. This research was directed XML and wrappers. One end result of this research, namely Netmark, is currently disclosed to NASA and a patent application is in progress. This work has been put to use for day to day operations at NASA.
- **Virtual National Airspace Simulation**. The work accomplished for the Virtual National Airspace System (VNAS) provides the underlying infrastructure necessary for a large airspace simulation. Smart Simulation and intelligent information in conjunction with distributed network of super-computers and a high-speed network (Information Power Grid) will provide the capability to continuously monitor and measure aircraft performance.
- **Remote Tower Sensor System.** The work done for the Remote Tower Sensor research focuses on developing virtual tower systems that integrate real-time airport data to support smart operations for towers and during low/zero visual operations. Image processing techniques were developed to aid automatically and in real time help airport towers operates under difficult weather (e.g. San Francisco Airport). The Half Moon Bay deployment will enable remote airport operations via Future Flight Central.
- **Program Management Tool** provides an integrated, coherent knowledge repository of milestone definitions, deliverables, and resource allocations to monitor and track the status of various NASA enterprises. The tool is developed to assist the program and project line managers in meeting research and development goals, objectives, and the proper allocation of resources to promote successful NASA agency missions.
- **Simulation Access Retrieval Gateway** The concept of Simulation Transaction Access and Retrieval Gateway (STARGATE) mirrors simulation access to the level of databases and employs a user-centered approach in its design, namely SQL (Structured Query Language). This modern version of information also provides remote access over distinct computing environment. This ability to access simulations as part of an information system adds a significant new capability by allowing simultaneous and seamless access to factual data and projections (e.g., numerical simulations logistics data with future planning and cost projections).

### II.D.10.e. Publications during FY2002

#### II.D.10.e.i. Publications

99. David A. Maluf, Peter Tran, Articulation Management for Intelligent Integration of Information, IEEE Systems Man and Cybernetics, Special Edition on Information Management 2001.
100. David A. Maluf et al, Secure Large Scale Airport Simulations Using Distributed Computational Resources, CrossTalk, Journal of Defense, 2002.
101. David A. Maluf, et al., An Extensible Schema-less Database Framework for Managing High-Throughput Semi-Structured Data, Applied Informatics, 2003, IASTED.
102. A. Maluf, et al., NETMARK a schema-less database paradigm, FLAIRS, 2003.

### II.D.10.f. Other Activities

#### II.D.10.f.i. Patents

- An efficient approach to handle Unstructured data in Relational Databases. (Disclosure Applied through NASA). Inventors: David Maluf, Peter Tran, Yuri Gawdiak, Tracy La, Jenessa Lin.
- A method to integrate information systems and simulations or a Simulation Transaction Access and Retrieval Gateway (STARGATE), Inventors: David Maluf.

#### II.D.10.f.ii. Program Member

- Twelfth International Symposium On Methodologies For Intelligent System ISMIS02.
- Conference on Computer Science and Information Technologies CSIT2002.

#### II.D.10.f.iii. Conference/Workshop Organizer

- PDARS User Group NASA/FAA (Feb 2002)
- PDARS User Group NASA/FAA (June 2002)
- Workshop on Human Organizational Studies for Engineering for Complex Systems, NASA (Aug. 2002)

### II.D.10.g. Future Plans

Future plans encompass the development of technologies that will dramatically reduce system/subsystem development cost and time, as well as provide a means for assessing and integrating multi-fidelity aerospace vehicle simulations.

## II.D.11. Computational Technology Applied to Human Circulatory System

### II.D.11.a. Task Summary

The objective of the task is to apply state-of-the-art computational technologies adopted from aerospace research into the hemodynamics of human circulatory system, thereby yielding a basic understanding of the impact of altered gravity on blood circulation and the development of a correlation between blood-flow patterns and arterial diseases. In the first phase, non-Newtonian behavior of blood flow and unsteady pulsatile inflow/outflow boundary conditions are to be numerically modeled based on physiological/experimental data. In the second phase, the effects of altered gravity and wall distensibility of blood vessels on blood flow patterns are also to be investigated under various gravity conditions. As a future work, a study on the correlation between blood-flow patterns and arterial diseases will be conducted by introducing a numerical model for the simulation of arterial stenotic development. This task can offer the fundamental database not only to NASA's long-duration space flight missions, but also to biomedical/biomechanical engineering industry by identifying the hemodynamic mechanism underlying the blood circulation under various gravity environments.

### II.D.11.b. RIACS Staff

Chang Sung Kim

#### II.D.11.b.i. Visiting Staff

Doyung Lee

#### II.D.11.b.ii. Subcontractors

Univ. of Eastern Illinois

### II.D.11.c. Project Description

The specific research objectives to be conducted under this task of interest can be summarized as:

- Numerical modeling of non-Newtonian behavior of blood flows
- Code validation by comparing with experimental data
- Numerical modeling for unsteady pulsatile inflow and outflow boundary conditions
- Study on wall distensibility of blood vessel
- Effects of altered gravity on blood flow patterns
- Three-dimensional reconstruction from Magnetic Resonance Angiography (MRA)
- Study on the correlation between blood-flow patterns and arterial diseases
- Parallel computation for the simulation of atherosclerosis and intimal hyperplasia development

### II.D.11.d. Accomplishments during FY2002

- Implementation of two non-Newtonian models into parallelized INS3D code
- Non-Newtonian model verification by comparing with steady/unsteady experimental data.
- Development of capillary bed model for unknown outflow boundary condition
- Implementation of a distensible wall modeling
- Suggestion of Gravity Benchmark Problems (GBP)
- Study on the effects of arterial wall distensibility and altered gravity on blood circulation

### II.D.11.e. Problems Encountered and Possible Resolution

The generalization of Newton's Law in the Navier-Stokes equations might fail in simulating the non-Newtonian behavior of blood flows having various sized blood cells. Based on physiological data, therefore, an appropriate non-Newtonian modeling should be introduced and validated. Cooperation with molecular dynamics technique is strongly recommended to identify the complex flow phenomena like fluid-fluid and fluid-wall interactions in capillaries. Another major problem is MR Angiography data acquisition for physiologically accurate three-dimensional reconstruction of complex arterial network. Co-investigation with biomedical researchers can provide a crucial breakthrough in addressing this problem.

### II.D.11.f. Publications and Presentations during FY2002

None

### II.D.11.g. Future Plans

- Three-dimensional reconstruction from MRI or MRA
- Study on correlations between blood-flow patterns and arterial diseases
- Predictive modeling of atherosclerosis and intimal hyperplasia development
- Computational simulation of intimal hyperplasia development models

## II.D.12. Development and Enhancement of Plasma Simulation Codes for Semiconductor Processing and Space Propulsion Applications

### II.D.12.a. Task Summary

Semiconductor devices are becoming increasingly smaller in size and may ultimately be sub-100nm feature-size. It is important that the processes and equipment meet the stringent demands posed by the fabrication of such small devices. The NASA Ames IPT on Devices and Nanotechnology has therefore initiated extensive efforts to model such devices and has developed comprehensive plasma simulation codes for this purpose. In addition to the focus on semiconductor processing, the IPT is also exploring the potential application of the plasma simulation codes to plasma propulsion systems such as Hall thrusters and pulsed plasma thrusters, which are key technologies being assessed for NASA's future interplanetary missions. The present research effort is directed towards assessing fundamental issues related to both the semiconductor processing and the plasma propulsion applications. Specifically, the research concerned the development and assessment of computational fluid dynamic (CFD) and neural network based tools for semiconductor process reactor design and space propulsion applications.

At the present time, reactors for CVD and plasma processing are designed by trial and error procedures. Once the reactor is in place, optimal processing parameters are determined through expensive and time-consuming experimentation. The availability of reliable and practical computational models and design tools would greatly simplify these tasks, while also being cost-effective. The idea pursued in this research is to use CFD simulations to populate the design space and to train the neural net with this data. The neural net is then used as a function evaluator to predict, evaluate and optimize the design space.

The second portion of the study involves the enhancement of plasma simulation algorithms, including both plasma-CVD reactors as well as space propulsion applications. The physical modeling of these devices shares some similarities but also some important differences. Both semiconductor and propulsion problems require the coupled solution of the fluid dynamics equations with multi-component species transport and finite-rate chemical kinetics. The multi-component species include mass, momentum and energy transport of the neutral species, ions and electrons. In addition, in the case of space propulsion applications, the plasma is characterized by different velocity scales---the neutral species continue to be low speed, but the ionic species accelerate to very high velocities (necessary for producing thrust). Moreover, the pulsed plasma thruster flowfields are also inherently unsteady, thereby introducing additional time scales into the problem. These differences necessitate the adaptation and further enhancement of the underlying numerical methodology. In the field of CFD, preconditioning methods are typically employed for handling such stiff problems. Here, we are developing a preconditioning approach suitable for the plasma simulation problems of interest.

**II.D.12.b. RIACS Staff**

**II.D.12.b.i. Visiting Staff and Students**

Sankaran Venkateswaran

**II.D.12.c. Project Description**

- Testing and validation of CFD/neural network based design methodology. The idea is to train the neural network using CFD simulations and then validate the network predictions using additional CFD simulations. Of specific interest is the degree of effort required to achieve specified levels of accuracy in the neural network predictions.
- Assessment of preconditioning techniques for enhancing the performance of plasma simulation codes. In particular, appropriate algorithmic modifications are to be devised to account for the conservation equations for neutrals, ions and electrons typically encountered in a plasma system.
- Adaptation of semiconductor plasma simulation codes for application to electric propulsion systems.

**II.D.12.d. Accomplishments during FY2002**

- Detailed validation studies of the CFD-neural network based tool has been completed and the results published in the Journal of the Electrochemical Society. Several hundred multi-dimensional CFD solutions were obtained for purposes of training and validation. The results indicate that the neural network is capable of very accurate (less than one per cent error) predictions of both average and local deposition rates in a CVD reactor.
- Preconditioning methodology has been developed for the plasma system. Preliminary studies indicate that significant performance improvements should result by incorporating the changes in plasma simulation codes.
- Progress was made in the adaptation of an existing semiconductor plasma simulation tool to handle Hall thruster plasmas with application to space propulsion.

**II.D.12.e. Problems Encountered and Possible Resolution**

None

### II.D.12.f. Publications and Presentations during FY2002

#### II.D.12.f.i. Publications

103. Venkateswaran, S., Rai, M. M., Govindan, T. R. and Meyyappan, M., "Neural Network Modeling of Growth Processes," *Journal of the Electrochemical Society*, Vol. 149 (2), G137-G142 (2002).

### II.D.12.g. Future Plans

- The present study has considered CFD-neural network based design of reactor processes. Future work will take into account the design of reactor geometry and configuration. Further, research should also be extended to process control as well.
- Algorithmic enhancements should be implemented in NASA's plasma simulation codes to improve accuracy and efficiency of the computations.

## II.D.13. Technical Leadership Recruiting

### II.D.13.a. Task Summary

This task is to help in identifying and recruiting highly qualified scientists and engineers from the university and research community to assist in providing technical leadership to NASA Ames research programs.

### II.D.13.b. RIACS Staff

Barry Leiner
Serdar Uckun
Sue Christman
Peggy Leising

### II.D.13.c. Project Description

NASA Ames plays a major role in leading and conducting state-of-the-art research in areas required for future NASA missions. This is particularly the case for information technology, where Ames is NASA's Center of Excellence. This role requires experienced and innovative leadership, both managerial and technical.

The academic and non-governmental research community plays a major role in the conduct of these research programs. Experience, both at NASA and elsewhere in the government (e.g. DARPA), indicates that engagement of this community in the planning, execution, and management of the programs is key to their success. This is a major motivation, for example, behind the Intergovernmental Personnel Act (IPA), which allows members of the university and not-for-profit research community to spend time inside the government assisting in the management of the research programs.

The objectives of this task are to

- Identify highly qualified researchers who can assist in providing leadership to the NASA Ames research programs in information technology and related areas,
- Assist in recruiting such candidates to spend time at NASA Ames, and
- Facilitate and support the engagement of identified researchers in a productive assignment at NASA Ames.

Specifically, we engaged in the following activities:

- Assist in identifying qualified researchers through advertising and word of mouth.
- Collaborate with home institutions and NASA Ames to identify the most effective method for engagement of the researcher. Potential methods include (but are not limited to) IPAs from the home institution directly, IPAs through USRA/RIACS, temporary assignments directly with NASA.
- Facilitate arrangements for the researcher to be engaged at NASA Ames.
- Support the researchers during their stay at NASA Ames.

### II.D.13.d. Accomplishments during FY2002

During FY2002, we recruited 15 scientists to work at Ames, as follows: Jan S Aikins, Gregory S Aist, Naveen Ashish, David Bell, Dan Berrios, James Crawford, Julian Gomez, Laleh Haghshenass, Chang Sung Kim, Bill Macready, Ronald Mak, Ashok Srivastava, Zhendong Su, Doron Tal, and Serdar Uckun. Of particular note are:

- Jan Aikins, who after a brief employment with USRA/RIACS, became a civil servant and deputy to the Director, Code I.
- James Crawford, who is playing an increasing leadership role in Code IC/ARA.
- Ashok Srivastava and William Macready, who were recruited to initiate an Ames research activity in machine learning for knowledge discovery.
- Serdar Uckun, who was recruited at RIACS Deputy Director.

In addition, through RIACS seminars and workshops, the visiting scientist and student programs, and the activities of the various RIACS staff, visibility has been given to the various RIACS and Ames activities and the desirability of being involved with them.

### II.D.13.e. Problems Encountered and Possible Resolution

As mentioned above, an important mechanism to engage researchers as leaders at Ames is the IPA program. Administrative problems with the IPA program have caused difficulties this past year. We hope they will be resolved early this coming year.

### II.D.13.f. Publications and Presentations during FY2002

None

### II.D.13.g. Future Plans

We intend to continue to be aggressive in advertising the opportunities that Ames/RIACS research represent, and identifying highly qualified candidates to fill those needs.

### II.D.13.h. Other Information

Open RIACS positions may be found at http://www.riacs.edu/employment

## *II.E. Visiting Students and Scientists*

A major contribution made by RIACS is the sponsoring and hosting of visiting students and scientists from universities and industry.

### II.E.1. Summer Student Research Program

Ten undergraduate and graduate students, from universities internationally, were selected in this third year of the SSRP. Five students returned from last year's program for a second session with the support of their mentors. Each student spent 10 weeks at Ames, teaming with NASA scientists on research projects in a variety of areas in IT, including automated planning and scheduling, natural language understanding, model-based autonomy for spacecraft and rovers, automated software synthesis and verification, model checking and real time fault detection.

Toward the end of the summer, students were invited to submit a proposal to continue their research at their universities on their return. René Vidal of Berkeley was chosen to receive this continuing research award for his research on the Bayesian Computer Vision project.

The following are descriptions of the research projects conducted by the SSRP participants.

#### II.E.1.a. Reinforcement Learning for Autonomous Agents
**Adrian Agogino of University of Texas**
**Mentored by Kagan Tumer**

This summer I researched methods of solving Markov Decision Processes (MDPs) with collections of agents, from a mathematical perspective. I was able to show that single agent systems that are Markovian, cease to be Markovian in the multi-agent version. However, I showed how some simple constraints can be put on the single agent problem to force it to remain Markovian, when converted to a multi-agent one. In addition I compared how different learning methods can best be applied to MDPs of collectives. I showed the relationship between simple, single-step reinforcement learners and more complex learning systems such as SARSA learning and Monte-Carlo learning. I showed how simple learners can be made to act like more complex learners, when the reinforcement functions are manipulated. This allowed me to compare relative tradeoffs between different types of complex learners that are not relevant in single agent systems, but become important in multi-agent ones. To test the theoretical results, I performed some trials on a novel time extended version of a traditional multi-agent problem, known as the El Farol Bar Problem. These experiments vindicated some of the methods used to train agents in a multi-agent MDP, as well as providing more evidence of the relative benefits and tradeoffs of some common reinforcement learning systems in multi-agent environments.

### II.E.1.b. Spoken Dialogue Systems Among Collaborating Intelligent Agents
**Nate Blaylock, University of Rochester**
**Mentored by John Dowding**

I initiated a new research project within RIALIST on spoken dialogue systems among collaborating intelligent agents. I used my background in both spoken dialogue systems and contemporary planning systems. More specifically, my contributions were to develop a theoretical model of collaborative problem solving, and building a dialogue system for an agent that can both plan and act in the real world. A paper describing this work that has been accepted for presentation at the 3rd International NASA Planning and Scheduling Workshop. On a technical level I led the integration of the RIALIST Spoken Dialogue System with the Intelligent Deployable Execution Agents (IDEA) planning and execution system, allowing it to engage in dialogues with agents that can both plan and act. In doing this, I worked closely with the developers of IDEA in the Autonomy and Robotics (ARA) group in Code IC, to make improvements in IDEA as necessary to support this collaboration, and to help bridge the gap between the spoken dialogue and planning communities. The current demonstration includes an agent carrying out a complex multi-step plan, and the same agent being directly commanded in real-time to carry out additional tasks. This forms the basis of an integrated system and test-bed in which further collaborative problem solving research can be addressed.

### II.E.1.c. Reliability of Spoken Interfaces for Space Station Procedures and Checklists
**Dan Bohus, Carnegie Mellon University**
**Mentored by Greg Aist**

My summer contribution comprised building a state-of-the-art task-oriented dialogue management component and integrating it into the Checklist system. This component lies at the core of the system, providing the dialogue control logic. It receives as input a semantic representation of the user's utterance, and, based on its current state and on interactions with other resources, decides what the system should say or do next. This dialogue management component was developed based on the RavenClaw dialogue management framework, previously developed at CMU. RavenClaw provides (1) a basic (configurable) set of core conversational skills, which stay essentially the same across various domains (i.e. handling dialogue flow, topic selection, focus shifts, non- and misunderstandings, various universal dialogue mechanisms like "repeat", "help", etc.), and (2) a framework for implementing the actual domain (dialogue task). The primary effort was therefore concentrated on building an implementation of the Checklist domain, supporting the guidance and information access functionalities described in the previous section.

A second contribution consisted of extending the conversational skills level in the core of the RavenClaw framework. Improvements addressed turn-taking issues like timeouts and barge-in capability, as well as the ability to handle requests to repeat the last system prompt, provide system help, and perform context establishment, a key feature since astronauts often pause tasks and return to them later.

Last but not least, I made contributions to the development of an XML representation for the procedures, which, apart from the contents of the task, also captures auxiliary information needed in a spoken dialogue system (i.e. lexical items, logical forms, relationships between objects and images, etc.)

### II.E.1.d. Three-dimensional Extended Kalman Filter based Simultaneous Localization and Mapping
**Chieh-Chih (Bob) Wang of Carnegie Mellon (Joint work with Doron Tal)**
**Mentored by Peter Cheeseman**

A solution to Simultaneous Localization and Mapping (SLAM) allows rovers to operate in an unknown environment and then incrementally build a map of this environment and concurrently use this map to localize rovers themselves. Over the last decade, SLAM has been successfully implemented in many applications. But most of these applications are still limited to two-dimensional environments and use precise sensors such as laser range finders. According to our previous super-resolution work, we believe that using optical sensors to accomplish three-dimensional SLAM tasks is indeed possible. This summer we modified and expended the Extended Kalman Filter (EKF) based SLAM algorithm from 2D to 3D. Exponential map representations are used instead of Euler angle in order to deal with singularity issues. A 3D visualization software is established, which would be very useful for further development. Several SLAM issues, such as filtering, computational efficiency, feature representation, and data association, were studied and discussed in this summer.

### II.E.1.e. Spoken Interface for Controlling a Semi-Autonomous Robot
**Ellen Campana of University of Rochester**
**Mentored by Beth Ann Hockey**

This summer I worked to improve the spoken interface for controlling a semi-autonomous robot that is intended to perform functions on the International Space Station (the RIALIST Spoken Dialogue Interface to the PSA). The ultimate goal of the interface is to provide a platform for natural interaction with the robot, which means that the interface must be able to handle complex language and deal with ambiguity in a natural way. Specifically, my project this summer focused on improving the way that the spoken dialogue system identifies the real-world referents of definite noun phrases. For instance, if the user directs the robot to "open the porthole" the system must be able to identify which porthole of many the user is referring to. Commonly in this case systems are forced to ask the user to be more specific, but this process can be time-consuming and unnatural. Our goal was to use information about the user's natural gaze while speaking to reduce the need for such clarification questions. There are two major components to my summer work with the RIALIST group this year. The first is a data collection effort. This is a necessary first step for system development because there is no data on the natural eye-movements that people make during interaction with a spoken dialogue system, even though there is considerable laboratory evidence that eye-movements are closely coupled with language production processes. I am currently in the process of conducting an experiment, designed this summer, in which we monitor users' natural eye-movements as they interact with the system. We systematically vary the concept recognition accuracy of the system, and the urgency of the task in order to investigate whether our findings can be generalized to environments that the intended end-users of the system are likely to be in. The second component of my summer research this year is a software agent that performs reference resolution-- the process of mapping incoming language to entities in the world. I developed this agent in such a way that it processes referential constraints in a human-like way, and it can be integrated with the general RIALIST architecture for application in many domains. This software component provides supporting architecture for a number of possible modifications to the spoken dialogue system, which will be the subject of future investigations.

### II.E.1.f. Model Checking Labeled Transition Systems
**Jamieson Cobleigh, University of Massachusetts**
**Mentored by Dimitra Giannakopoulou**

Model checking has been successful in proving a variety of properties on hardware and software systems. Model checking is limited by the number of system states that can be stored in memory. Since systems are often built from modules, it is advantageous to analyze the modules individually rather than analyze the entire system. The analyses of the individual modules can then be combined to derive global system properties, an approach known as compositional reasoning. In such a setting, the behavior of a module is dependent on the modules it interacts with, making it is necessary to construct assumptions about the behavior of the other modules to prove that a property is never violated. Our work examines two approaches to automatically generating assumptions: an incremental approach based on looking at counter-examples to a property, and a learning approach based on queries and counter-examples. We implemented both of these approaches in the LTSA tool, which analyzes Labeled Transition Systems. The learning approach was also implemented to work on Java code using Java PathFinder, where we generated assumptions for the DEOS real-time scheduling kernel.

### II.E.1.g. Knowledge Base for Data Tracking
**Judah Ben DePaula of University of Texas**
**Mentored by Keith Golden**

Dr. Keith Golden of NASA Ames is developing a system called IMAGEbot. IMAGE stands for: Information Manipulation, Archiving, Gathering, and Extraction. The purpose is to manage the huge amount of data gathered by NASA. NASA receives a stream of data from resources including satellites, telescopes, and planetary rovers. The data must be processed and delivered to scientists is a timely manner which in some contexts, such as detecting severe storms, may mean less than a day. The data must also be archived for future reference. The terabytes of archived data are worthless unless the scientists that wish to use them are able to find what they need. Traditionally, hand-written scripts process the raw data files to compress, combine, or analyze the data. There are problems with this method of data management as hand-written scripts are brittle; even a small change to the environment can cause the program to fail.

IMAGEbot will address these problems directly by automating the scripting process. A researcher will be able to request information in a given format. This query will be passed to a planner, which will try to find a sequence of data-processing commands and a set of raw data inputs that will produce the desired result. The project is new, so the core code is still under development. The first tests of the fully integrated system are months away. There are a number of modules that remain to be designed and implemented. In the design of the planner there needs to be a knowledge base that will contain all information known about the environment. It was my job last Summer to design and implement that knowledge base. I also implemented routines to populate, and retrieve, information from the database.

*Summer 2002 Development*

When the planner within IMAGEbot needs a data file representing certain information about the world, a query is given to the database module. The system must know whether the knowledge base is able to entail the query. It is fairly straightforward to check for an exact match between the query and an expression within the database. Problems with complexity arise when two or more expressions must be combined before it is clear that the knowledge base entails the query. Working in conjunction with Keith Golden, I have nearly completed development on an algorithm that checks for the entailment of a query given a set of logical expressions. The general problem, determining if any set of arbitrary first-order expressions entails another first-order expression, is provably undecidable. By placing reasonable constraints on the structure of the metadata we were able to bring down the complexity of the problem and create a solution that can be implemented in IMAGEbot.

### II.E.1.h. Work Practices Analysis of Collaborative Software Development Teams

**Cleidson R. B. de Souza, U. C. Irvine**
**Mentored by John Penix**

This summer, I conducted a field study of the CTAS software development team at NASA Ames. The goal of this field work is to understand the current practices used by the CTAS software developers with the intent of eventually providing them support for their collaboration. The team studied develops a suite of software tools called CTAS (Center-TRACON Automation System), which is designed to help air traffic controllers manage the increasingly complex air traffic flows at large airports. The CTAS development group is composed of 31 co-located NASA employees and contractors, who design, test, document and maintain ten different tools.

One important lesson from the field of Computer-Supported Cooperative Work (CSCW) is that in order to understand work practices it is necessary to adopt field studies techniques. These techniques provide an understanding that can be used to inform design in such a way that better tools and techniques are developed to fit the developers needs. This is important because as Gerson and Star observed, no matter how formal and well-defined a process may seem, there will always and inevitably be a set of informal practices by which individuals and groups monitor and maintain the process, keep it on track, recognize opportunities for action and the necessity for intervention or deviation, etc.

In short, a work practice analysis project brings together the fields of CSCW and Software Engineering to provide a foundation for software engineering tools which enable people to collaborate effectively.

### II.E.1.i. Formalizing State Estimation Domain Knowledge Using UML
**Emanuel Grant of Colorado State University**
**Mentored by Jonathan Whittle**

The NASA Ames system that was the subject of the internship work (AutoFilter) is a code generator that synthesizes a fully executable program (a Kalman Filter). AutoFilter works by taking as input a mathematical specification of a set of stochastic equations, along with descriptions of the noise characteristics and filter parameters. AutoFilter then generates code that implements a variant of a set of equations, which were derived form the input. These equations are representations of many possible configurations of filters that may be synthesized. The internship work involved the creation of a set of UML models along with OCL constraint statements of the AutoFilter system. The UML models and OCL constraint statements are intended for use in validating the output from the system against the system input. In conducting the work AutoFilter was treated as a black box, and the work focused on developing static models of the input and output of AutoFilter. The task faced in this project was centered on being able to determine whether the structure of a filter program that was an output from AutoFilter conformed to a specified framework. This issue was born out of work done on modifying AutoFilter to provide additional functionality. It then became a concern as to whether the synthesis process was negatively affected by the modifications, i.e., was AutoFilter still generating synthesized filters in the same manner after the modifications as it was before the modifications?

### II.E.1.j. Analyzing Counterexamples Produced by Model Checking Programs
**Alex Groce, Carnegie Mellon**
**Mentored by Charles Pecheur**

This summer I worked on analyzing counterexamples produced by model checking programs. Just an error trace is often insufficient to allow a programmer to easily diagnose and understand a bug in a program. This work introduces a search technique to find suitable "near" traces that either exhibit different versions of the error or reach the same control state without error. Once the sets of negative and positive traces are generated, various analysis methods are used to describe common elements of the traces to explain the bug. In particular, invariants, program points covered, and means of transforming positive traces into negative traces are explored. Some results are not approximate; for instance, errors in which only thread scheduling is important may be identified.

### II.E.1.k. Formalizing the MDP approach for Planetary Rovers SSRP 2002
**Max Horstmann, University of Massachusetts**
**Mentored by Rich Washington**

Planetary Rovers such as the 1997 Sojourner Rover on Mars have to work under tough resource and time constraints. Long communication times make teleoperation impossible and lead to the need of autonomous control. Current approaches for autonomous decision-theoretic control are based on heuristic search in the space of contingency plans, but suffer from a low flexibility and lead to high downtimes due to plan failures. A promising approach is the formulation of the Rover problem as a Markov Decision Process (MDP). MDPs are handled by control policies which potentially provide an optimal action in all of the modeled states and therefore could lead to a higher degree of efficiency. The goal of my project was to find an appropriate action model for the MDP approach and to implement a framework for a high-level control agent. Experience and data has been gathered by working with the K9 rover, and several issues arose. For instance, the MDP formulation needs a fully observable state. But it turned out that the K9 rover lacks information about several state features such as its own position. Possible solution: Actions include localization methods, so there's no or only little uncertainty about the state every time a decision has to be made. As another example, some actions require anticipation, such as the warmup of instruments. This does not naturally fit into the MDP framework. Possible solution: Analyze control policies during execution and use stochastic prediction methods. I developed a program to automatically generate Rover instructions, based on the Contingency Rover Language (CRL) and could successfully run it on the K9 rover. This framework can be used in further research. One of the long-term goals is to test decision-theoretic methods developed in my research group at the University of Massachusetts on the K9 rover at NASA Ames. Also, I formulated an MDP state- and action space description and a realistic scenario for the application in a future mission. To give an idea about potential improvements to the contingency plan approach, I formulated a test problem and ran it in a simulation, showing that an optimal solution would need many branches if it were formulated with the current methods. All in all, the results were encouraging and showed the potential of the MDP approach, so my future research will be based on results of my work during the SSRP.

### II.E.1.l. Model-Based Fault Diagnosis in Hybrid Systems
**Frank Hutter, University of Darmstadt**
**Mentored by Richard Dearden**

My work at NASA Ames during the summer 2002 focused on model-based fault diagnosis in hybrid systems. In this context, a hybrid system is defined as a system with both discrete modes and continuous parameters. Our primary domain of interest were planetary rovers, where the discrete component is given by various operational and fault modes. The continuous parameters in this domain include (amongst others) measured quantities like wheel current and speed. The inference task at discrete time step t is to reason about the rover's belief state (a probability distribution over possible states) given the sensor measurements up to time t. While in the case of linear models Rao-Blackwellised Particle Filtering can be used for efficient diagnosis, the case of non-linear models has received little attention in the community so far; standard Particle Filtering is usually used but infeasible in real-time scenarios. In my joint work with Richard Dearden, I have developed a new algorithm (UKF-PF) for efficient, biased inference in non-linear hybrid systems. The basic idea is to replace the Kalman Filter in Rao-Blackwellised Particle Filtering with an Unscented Kalman Filter to approximate sufficient statistics for the continuous parameters. Although through the bias we lose theoretical guarantees like convergence in the limit of an infinite number of particles, experiments show that in practice UKF-PF is far superior to standard particle filtering (PF): UKF-PF vastly outperforms PF and enables real-time fault diagnosis in non-linear hybrid systems. Application areas are not limited to the planetary rover domain. The technique is applicable for state estimation in hybrid systems in general whenever transition and observation models are available.

### II.E.1.m. Symbolic Execution for Concrete Test Case Generation
**Sarfraz Khurshid, MIT**
**Mentored by Willem Visser**

The model checker has been extended to allow symbolic execution, which in turn enables the analysis of open systems, i.e., systems with unknown inputs. This is the first time that a symbolic execution based system was able to handle the combination of complex data structures, numeric constraints and concurrency. I assisted in the extension and testing.

The model checker returns a counterexample trace if one exists; however understanding this trace can often be difficult. We therefore developed a technique whereby errors are "explained". The model checker itself is used to find other similar traces to the counterexample as well as traces that do not cause an error. The differences in these traces are then analyzed to give a succinct representation of what went wrong.

A Java translation of the K9 Mars rover was analyzed during a benchmarking experiment. I participated in this experiment, along with several other students. This example has been the largest attempted to date – it consists of 6 threads and a total of 8000 lines of code. The model checking of the K9 was highly successful and in the area of concurrency errors outperformed all the other technologies (testing, runtime analysis and static analysis) also benchmarked during the experiment.

### II.E.1.n. Certifying Kalman Filters
**Laurentiu Leustean, University of Bucharest**
**Mentored by Grigore Rosu**

Code certification is a lightweight approach to demonstrate software quality on a formal level. Its basic idea is to require code producers to provide formal proofs that their code satisfies certain quality properties. Our project this summer consisted in building a certifier for three Kalman Filters: Simple Kalman Filter, Information Filter and Extended Kalman Filter. Kalman Filters are stochastic, recursive algorithms which provide statistically optimal estimates of the state of a system based on noisy sensor measurements. They are the most common way of solving a state estimation problem, an important problem found in spacecraft, aircraft and geophysical applications. To generate and automatically validate such proofs, we need to formalize the domain knowledge, which includes: matrices, random matrices, functions on matrices, matrix differentiation. The language we chose for this purpose was Maude, a high-performance executable specification system in the OBJ family. For the optimality proofs for the three Kalman filters we used ITP, an inductive theorem prover implemented in Maude. We used many domain-specific (matrix theory) lemmas, which were formulated and proved independent of the Kalman Filters, so they can be stored in a database together with their proofs. With the help of these lemmas, we automated the proofs, minimizing the number of hints. Thus, we had only 4 hints in the optimality proofs for all three Kalman Filters, while in previous work there were used more than 400 hints only for the Simple Kalman Filter.

### II.E.1.o. Initialization of Model-Based Clustering Algorithms
**Kate Mullen, Bard College**
**Mentored by Bernd Fischer & Johann Schumann**

Clustering is a ubiquitous data analysis task. Its goal is to group together subsets of a data set such that elements assigned to the same subset have the least amount of variance. In my work this summer, I concentrated on cases where a model exists for the data set such that the number of classes (i.e., data subsets) and the probability density function followed by each of each of the classes is known. In such cases, the Expectation Maximization (EM) and K-means algorithms may be applied. EM and K-means are iterative algorithms that refine the model parameters ineach iteration and that converge monotonically to a local maximum in the likelihood of the parameters given the data. However, most data analysis tasks are multimodal and hence the algorithms are not guaranteed to converge to the global maximum. Because the likelihood reached is deterministic after initialization, refinement of the initialization process is important. I developed and implemented initializations that improve on the random initialization strategy currently employed in the EM algorithm as it is implemented by the AutoBayes system for automated software synthesis. A suite of tests was developed and run on univariate and multivariate artificial data in order to quantify the improvements in terms of maximum likelihood reached, iterations required, and stability of solutions reached. A number of initializations were shown to be preferable as compared to the random strategy, and are being considered for incorporation within AutoBayes.

### II.E.1.p. Integration of CONFIG into Livingstone Pathfinder
**Aina Razermera-Mamy, University of Oregon**
**Mentored by Charles Pecheur**

The work was to integrate CONFIG in Livingstone Pathfinder. Livingstone Pathfinder, a tool maintained by Charles Pecheur and Tony Lindsey, is a structure to control and automate the execution of Livingstone, a model-based diagnosis system developed at Ames. For its operation, it incorporates a testbed which in turn includes a simulator for physical system. Prior to the work, the simulator was another instance of Livingstone. The goal is to replace it with CONFIG.

CONFIG is a simulation framework from the Johnson Space Center and the development of which is headed by Jane Malin. The bulk of the work was to interface CONFIG, a Lisp program, with Livingstone Pathfinder, a Java program. Two interface tools were separately used. The first one is IPC, an inter-process communication package developed at the Carnegie Mellon University by Reid Simmons. The second tool is JLinker from Franz Inc.

### II.E.1.q. Human-Machine Interaction Project
**Oksana Tkachuck, Kansas State University**
**Mentored by Guillaume Brat**

This work is a continuation of the work I did last summer. We are working on the paper for the DASC conference. The paper is based on the work and the report I've done last summer. The paper and the report are entitled "Using Code Level Model Checking to Discover Automation Surprises". As a side work, I implemented a tool that extracts logic tables from avionics software. This is a very useful exercise since currently this is being done by hand. The tool takes as an input a block of heavily nested IF-THEN-ELSE code and identification of inputs and outputs of the block. The tool extracts a Control Flow Graph from the block, and for each feasible path in the system records a separate mode (situation). Feasibility of paths is identified with the help of a theorem prover. The output of the tool is an SGA (Situation-Goal-Action) table. We are hoping to continue working with Human Factors on automation of their tasks.

### II.E.1.r. Finding Optimal and Pareto Optimal Solutions of Simple Temporal Problems with Preferences
**Brent Venable, Univ. of Padova, Italy**
**Mentored by Lina Khatib**

The STP (Simple Temporal Problem) formalism is already used in some projects here at NASA. However it deals only with hard constraints and that means that it is useful only for problems where knowledge is crisp and precise. There are many environments, on the other hand, where regarding everything as either "allowed" or "not allowed" can be impossible or simply meaningless. For the former reasons a new framework has been proposed for adding flexibility to STPs, that is: Simple Temporal Problem with Preferences. This new type of constraint satisfaction problem allows the addition of preferences locally, to the time points involved in the constraints and globally rating the solutions. My task for this summer at Ames was to implement optimal and pareto-optimal solvers for STPPs. .Addressed problems are problems that involve manipulation of temporal information and problems that involve having preferences associated with local temporal decisions

During my stay here at NASA I have been supervised mainly by Lina Khatib, my mentor, who has given me an essential help both in choosing the problems on which to focus and in finding solutions for them. Robert Morris has helped me see where my project could fit in, highlighting connections to other areas as scheduling and planning. Paul Morris has given some great ideas, that turned out to be essential in overcoming difficulties both in setting the conceptual basis and in the actual implementation.

### II.E.1.s. Bayesian Computer Vision - BAYVIZ
**Rene Vidal, U.C. Berkeley**
**Mentored by Peter Cheeseman**

This summer I used my insight into degeneracies in the error surface, coming from a prior understanding of the computer vision problem and previous related research, which allowed the group to avoid having to rediscover this problem themselves. This saved us very considerable effort, and lead to a much-improved algorithm for finding the true error surface. Also, I implemented this improved error surface fitting program which the group is now testing.

## II.E.2. Visiting Scientists' Research Project Descriptions

RIACS regularly sponsors and hosts visiting scientists from universities and industry to provide highly specialized expertise to various NASA projects. The following are descriptions of the research conducted by these visitors.

### II.E.2.a. Howard Barringer

*Verification and validation (V&V) to improve the reliability of software – Willem Visser, Dimitra Giannakopoulou*

If formal, or even rigorous, specification and development methods are going to play a significant role in either software or hardware development practices then these highly specialized methods must be embedded within tools and techniques that will be easy to use for the user. Two main areas were addressed:

1. Assumption Generation for Software Component Verification
   This work defines a framework that brings a new dimension to model checking of software components. When checking a component against a property, our model checking algorithms return one of the following three results: the component satisfies a property for any environment; the component violates the property for any environment; or finally, our algorithms generate an assumption that characterizes exactly those environments in which the component satisfies its required property. Our approach has been implemented in the LTSA tool and has been applied to the analysis of a NASA application.

2. Application of ESC/JAVA to the "Rover" Java Code
   Four, bottom level, class files, TimePoint, TimeInterval, TimeDate and TimeCondition were selected for initial analysis. My approach to application of the tool was to add the minimal number of annotations necessary to get the files "passed" by the checker.

After many runs, it is concluded that the tool is best designed for rapid discovery of common programming faults such as null object de-referencing, array bound checking, etc. Weaknesses include the specification language for function and predicate definitions, underlying theories are difficult to build for abstractions corresponding to classes, and a great deal of annotation is required for anything other than the simplest constraints.

### II.E.2.b. Bjorn Sjogreen, Neil Sandham

*Computational Fluid Dynamics – Helen Yee*

Our interest has focused on numerical methods for computing astrophysical phenomena, in particular problems where magnetic fields play an important role. The equations of magneto-hydrodynamics (MHD) is a reasonable model for many such phenomena, for example solar wind / Interstellar plasma interaction, or the accretion disk around a black hole. The MDH equations are obtained by adding forces from a magnetic field to the compressible Navier-Stokes equations, and by adding equations for the magnetic field components.

We extended our high order method for hydrodynamics equations to the ideal MHD for both the conservative and non-conservative form. Our results show that it is possible to implement the Roe method for the conservative equations in such a way that good divergence conservation holds.

We have developed two sixth order accurate methods, WAV66 and ACM66 for the MHD equations. An interesting aspect of the MHD equations is the existence of so called intermediate waves. We solve the supersonic flow past a perfectly conducting cylinder, where, for sufficiently strong magnetic fields, the bow shock normally occurring in front of the cylinder, breaks up into several different waves, some of them intermediate.

### II.E.2.c. Pavel Bochev

*Computational Mathematics and Algorithms – Tim Barth*

I visited RIACS from 05/19/02 to 05/22/02. I am from Sandia National Laboratories and I was hosted by Dr. Tim Barth. My visit had two main objectives which are summarized below. 1. Presentation. I gave a talk on "Computational electromagnetics and differential complexes". This talk familiarized the audience with the research and development work that is being currently done at Sandia National Laboratories. My talk focused on the close connections between edge elements, staggered finite differences and co-volume methods in electromagnetics with some ideas from differential geometry. 2. Discussions. I had several meetings with NASA Ames staff members to discuss the work done at Sandia Lab on discretization and numerical solution of PDE's. During these meetings I described the research I am involved with at Sandia and the methodology that we are using to solve problems in electromagnetics, device simulation and fluid mechanics. I also had an extensive working session with my host, Tim Barth. Our discussion focused primarily on the Discontinuous Galerkin method and its application to Hamilton-Jacobi equations.

### II.E.2.d. Marsha Berger

*Algorithms for inviscid flow in complex geometries – Michael Aftosmis*

In collaboration with Michael Aftosmis and Scott Murman, we are continuing the development of algorithms for inviscid flow in complex geometries. Our approach is based on the use of Cartesian grids with embedded boundaries that are not aligned with the mesh. In previous work, we have developed a mesh generator, watertight surface triangulation tool, multigrid mesh coarsener, and a highly scalable steady state flow solver based on the use of space filling curves. This summer, the major focus was the beginning of development of a time dependent flow solver that incorporates moving geometry. We derived the basic equations that moving geometry satisfy for this kind of mesh, where the object simply moves through a stationary mesh. We developed various levels of approximation to this, the simplest being staircased in time (akin to the staircase meshes formerly used with Cartesian methods and embedded boundaries in space). Implementation has only just begun. Since part of the implementation will involve our previously developed multigrid and other tools, we spent part of the time this summer speeding those steps up.

### II.E.2.e. Julia Brodsky

*ISS Modeling Scenarios – Maarten Sierhuis*

Julia Brodsky had consulted the ISS modeling group on the ISS flight scenarios, mission timelines and onboard flight procedures. She also conducted documentation research and provided the group her astronaut training expertise. In addition, she served as a liaison with the Johnson Space Center groups (Training Department, astronaut schedulers, flight controllers, etc) Her expertise allowed the group to identify the additional possibilities for Brahms modeling applications by various groups of customers( such as astronauts, astronaut instructors and flight planners)

### II.E.2.f. S. Charles Dey

*Three Dimensional Mapping and Medical Imaging Simulator – John Ziebarth*

I optimized PFI Immunization solving 8 equations in 25x25x25, 50x50x50, and 100x100x100 matrices in C++ ad developed a 3 dimensional mapping and medical imaging simulator. This model and simulation can be customized through variable settings to represent how cancerous cells may spread and how the immunization system may fail under different settings and different scenarios. This simulation shows the cancer cells growing and mutating as well as spreading to neighboring cells and attacking lymphocytes. Under other variable scenarios, it also depicts the lymphocytes surrounding and harboring the cancerous cells, forbidding the mutation to spread to healthy cells.

The simulation was written in OpenGL API with cancer and lymphocyte cell behavior being dictated by the mathematical model being solved by PFI, and general physics equations on particle theory, collision and gravitation models.

### II.E.2.g. Genevieve Gorrell

*Grammar testing Tools – Beth Ann Hockey*

A tool was created that enabled a series of grammar test runs to be done. A second tool was also implemented that allows pairs of test runs to be compared, and the significance of the difference in their performance to be assessed using McNemar sign test. The grammar testing tool takes a set of Nuance grammars and a set of transcribed corpora, and batch recognizes the corpora using grammars. The results of the batch recognition are then compiled into a Prolog readable database.

The McNemar sign test is most predominantly used in the field for comparing data of this kind. The Prolog readable database created by the grammar testing tool is used to produce a set of figures indicating the statistical significance of the difference between the performance of pairs of batch recognition test runs.

Progress was made on a research task currently being done in the RIALIST group. The task is to establish whether a useful class n-gram language model can be created from a Gemini grammar using part-of-speech information contained within it.

- The program used to create a class list from Gemini grammar was improved for the PSA domain and the Witas domain.
- Required additions to the class list were added.
- The resulting class list was used to create a trigram language model.

### II.E.2.h. Mohana Gurram

*Data Integration and Relational Databases – David Maluf*

- Developed an application to monitor server transactions from a remote client.
- Understood the functionality of Netmark (a schema-less object relational database management system by AEN group).
- Under the guidance of Dr. David Maluf, we were able to solve the problem of data integration from different data domains. The technologies associated to accomplish this task were XML, Sparrow, NETMARK and JSP. The goal of this task was to provide a user with an environment to build models to query any source of information and the ability of server side applications to understand those models and provide templates to

query the required source of information. Ultimately the extracted information is stored on server side file system and is monitored through Webdev.

- In the process of developing a tool to accomplish the task of object and relational transactions in a user-friendly fashion. This tool will be able to perform many database related operations from a user-friendly interface. In order to accomplish the task of object transactions, the tool relies on Netmark technology.

### II.E.2.i. Roby Joehanes

*Integrate Bandera with JPF – Willem Visser*

I mainly worked with Willem Visser and Flavio Lerda to integrate Bandera with JPF. Among the things needed for the integration are conformance with the JPF's new interface, fixing the LTL support, adding Promela-style formula in the Buchi automaton, and some bug fixes in both Bandera and JPF side. I also added some support of saving/loading JPF's counterexamples in Bandera.

### II.E.2.j. Charis Kaskiris

*'Brahms', an agent-based language – Maarten Sierhuis*

**Introduction**

I have been a visiting student at the Research Institute for Advanced Computer Science (RIACS), NASA Ames Research Center since May 2001. I have been working with the Brahms Team under Dr. Maarten Sierhuis and Dr. William Clancey on the Mobile Agents project. The group falls under the Human Centered Computing (Code IC), whose research involved modeling and simulation of work processes of humans and systems in organizations.

My work involves the study of human-robotic-systems interactions during extravehicular (EVA) activities on planetary surfaces and how agent-based technologies, namely Brahms, can be used in aiding the work practice through the use of intelligent mobile agents. Once we have a choreographed scenario of human EVA activities, then my role is to model the interactions in a simulation using the Brahms Agent-Oriented Language.

Simulations are critical at this stage as EVAs on planetary surfaces have not been performed since moon exploration and there is an important need of visualization of the processes involved. The interaction processes and protocols are designed during and through the model design phase. Our research objective is, however, even more ambitious; we have been working on a concept of building simulation code that through minimal integration effort we can plug-n-play agents into real-time systems. That is, we build the simulation with specific characterization of what is a real artifact/agent in the world, and then by specifying the interfaces, we can substitute them with their real-life representations and have the humans, robots, systems, and agents interact and perform the simulation scenarios in real life.

Under the auspices of the project I have also participated in a week long field study where our simulation was performed in real-life with astronauts, robots, and the artificial distributed agents we have built. The experiment was highly successful and it was a great experience as it involved on-site empirical observation of work practice and had elucidated technical issues with regards to the project. An even bigger field test is planned for September 2002 in Arizona.

**SIMS Research Relevance**

Why is this of interest to me? Well, this is a methodology that directly fits into the infrastructural aspects of my dissertation proposal. What I am interested in studying is what I would call Economic Systems Design (ESD). The main focus of ESD is the design, development, testing and implementation of exchange systems, including such applications as:

- auctions to allocate public resources (e.g., FCC spectra)
- network markets (e.g., electric power, natural gas, and water)
- scheduling systems (e.g., scientific observations in space missions)
- price patterns and overreaction in financial markets, and
- distributed software agents for computational market systems.

ESD utilizes experimental methods in economics as a way to testbed new exchange systems. More specifically we use "smart" (computer-assisted) markets to develop computerized systems and algorithms that allow participants to execute complex trades, subject to feasibility constraints.

The particular approaches I want to use include the realistic simulation representation of new types of electronic trading systems and the utilization of distributed agent architectures and personal agents in trading. Hence, it is critical for me to have a grounding in how distributed real-time and simulation agent oriented systems work. This is what the experience at NASA is offering to me in conjunction with the opportunity to work on the development of an Agent-Oriented Language.

**Professional Development**

The experience has also allowed me to interact with other simulation oriented and agent-based systems researchers within NASA and while attending conferences. Furthermore, I have also started extending my knowledge on distributed systems, object-oriented technologies, web services, and multi-agent systems. All of the above have direct bearing into the issues I want to explore in my dissertation and future research career. It is also a great pleasure to work with Maarten and the rest of the Brahms team as they are a high performance group with regards to the work we do.

I also had the opportunity to work on two amazing projects, one regarding visualization of simulations and the other on distributed multi-agent systems. In the process I have collaborated with virtual reality researchers, robotics researchers, space suit researchers, and agent-based simulation scientists.

**Publications & Work in Progress**

In the process of working on the two projects we have submitted an initial short paper, written technical reports, and have papers in progress. The areas we span are Artificial Intelligence, Human Factors, Space Exploration, Distributed Systems, and Simulations. The following list includes documents generated by myself and co-authors while at RIACS:

Brahms TM 02-0009 *"Mobile Agents: Distributed Human Robotic EVA System for Surface Operations; Scenario and Field Test Design For the September 2002 Field Test in Arizona"* - Version 0.3 Draft (July 28, 2002).

> This document serves as the main design documentation for the FY'02 field test in September in the Arizona dessert for the Mobile Agent Architecture (MAA). Its purpose is to be used amongst the MAA teams and the Principal Investigators of the MAA for collaborative design and interaction. It also serves as a technical specification on the real-time features of the distributed architecture. [with Maarten Sierhuis]

*"Brahms Mobile Agents: Architecture and Field Tests"* - Submitted HRI AAAI Symposium, (Fall 2002)

> We have developed a model-based, distributed architecture that integrates diverse components in a system designed for lunar and planetary surface operations: an astronaut's space suit, cameras, rover/All-Terrain Vehicle (ATV), robotic assistant, other personnel in a local habitat, and a remote mission support team (with time delay). Software processes, called "agents," implemented in the Brahms language run on multiple, mobile platforms. These "mobile agents" interpret and transform available data to help people and robotic systems coordinate their actions to make operations more safe and efficient. The Brahms-based mobile agent architecture (MAA) uses a novel combination of agent types so the software agents may understand and facilitate communications between people and between system components. A state-of-the-art spoken dialogue interface is integrated with Brahms models, supporting a speech-driven field observation record and rover command system (e.g., "return here later and bring this back to the habitat"). This combination of agents, rover, and model-based spoken dialogue interface constitutes a "personal assistant." An important aspect of the methodology involves first simulating the entire system in Brahms, then configuring the agents into a run-time system. [with William J. Clancey, Maarten Sierhuis, and Ron van Hoof]

*"Brahms -Virtual Worlds Integration"* - Working Draft (November 2001)

> This is a preliminary draft on the integration of the Brahms work-practice simulation environment with the Adobe Atmosphere Virtual Reality environment. It describes how the Brahms simulation environment has been integrated with the VR environment and used to visualize a simple model. [with Bruce Damer, Bruce Cambell, Maarten Sierhuis, and Ron Van Hoof]

Proposal 000104, *"Brahms VE: A Collaborative Virtual Environment for Mission Operations, Planning and Scheduling"* (October, 2001) [HTML]

> This document describes the integration between Brahms and Adobe Atmosphere. [with Bruce Damer, Maarten Sierhuis, Ron Van Hoof, Bruce Campbell, Dave Rasmussen, Merryn Neilson, Stuart Gold, Galen Brandt]

Brahms TM01-0007 *"STTR VRE/FMARS Demonstration Model"* - Version 0.45 Draft (29 November 2001)

This document serves as the background implementation manual for the Brahms model to be used as demonstration of the integration of the Brahms Simulation Engine with a Virtual Reality Engine and the demonstration of Brahms models in Virtual Worlds, using Adobe Atmosphere [with Maarten Sierhuis]

Brahms TM01-0008 *"Mobile Agents: Distributed Human-Robotic EVA System for Surface Operations: Scenario and Field Test Design"* - Version 0.3 Draft (20 December 2001) [*]

This document is still work in progress for the Mobile Agents Project. It involves the design of models for integration with the robotic systems, information appliances, and space suits. It also involves the design of field experiments to be contacted in the summer of 2002 in analog environments. [with Maarten Sierhuis & Michael Sims]

*"Interface between Brahms and the OWorld 3D API: Brahms VE OWorld exporter and JavaScript Writer"*, Provisional Statement of Invention, NASA Ames

Provisional statement of invention regarding the integration of the Brahms language with Adobe atmosphere and the representation of the model in 3D OWorlds. [with Bruce Damer, Maarten Sierhuis, Ron Von Hoof]

**Future**

The Mobile Agents project is funded for a third year and I intent to continue working on it. I also intent to initiate a research project in conjunction with that work that directly

### II.E.2.k. Dohyung Lee

*Multi-Dimensional Supercompact Wavelets for Fluid Dynamics And Data Compression with Hybrid Supercompact Wavelets – Dochan Kwak*

Highly compact wavelets named supercompact multi-wavelets are extended to higher dimensions for the purpose of numerical fluid solution data compression. Due to the compact nature of the wavelets, the proposed multiresolution approach with the supercompact wavelets offers numerous benefits such as high data compression. Supercompact wavelets are well suitable for gas flow in which some local complex features are embedded in an overall smooth solution. Supercompact wavelets provide advantageous benefit that it allows higher order accurate representation with compact support and therefore, it avoids unnecessary interaction with remotely located data across singularities such as shock and vortices. The numerical implementation is relatively easy and less expensive than the comparable interpolatory multiresolution scheme. The practical demonstration of the multidimensional supercompact wavelets is presented. Several numerical tests demonstrate large data compression ratios for the outputs of flow field simulation Fluid dynamics and wavelets applications.

When solving fluid dynamics equations and manipulating solution data of the fluid dynamics equations, there are two ways to utilize wavelets: direct and indirect methods. The direct method is to implement Fast Wavelet Transform (FWT), similar to Fast Fourier Transform (FFT), which computes solutions in the transformed space. In the indirect method, the solution is computed in the physical domain using wavelets to reduce the computational work load and/or to improve the accuracy of solutions.

For direct methods, Engquist B, S. Osher and S. Zhong [1] presented a class of fast wavelet based algorithms for linear evolution equations with time-independent coefficients. O. Vasilyev and S. Paolucci [2] developed a dynamically adaptive multilevel wavelet collocation method which allows a simple method of local computational refinement.

For indirect methods, A. Harten [3] presented a method of computing compressed solutions with multiresolution representation rather than placing the solutions on a given grid structure. The efficiency of this method is proportional to the rate of data compression. M. Gerritsen [4] used wavelet theory in detecting shocks, sharp gradients and local oscillations. L. Jameson [5] presented a method to use wavelets for grid refinement and order selection.

**Wavelet bases for fluid dynamics**

Typical photographic images and fingerprint images are very unlike the numerical solutions of the Navier-Stokes equations. In general, the former images have very little inherent smoothness with discontinuities almost everywhere. On the other hand, the numerical solutions are smooth almost everywhere and have discontinuities only in local areas (shock, vortices, and shear layers). Wavelet bases determine the efficiency of wavelet transformation. The wavelet bases should be representative of the solution to the problem at hand and applicable to constraints such as numerical accuracy and boundary conditions.

Harten [3] recognized that interpolatory multiresolution methods are well suited for representing solutions of the partial differential equations. Harten's interpolatory algorithm for a uniform grid and periodic boundary conditions can easily be extended to non-uniform grids and non-periodic boundary conditions because the requisite interpolation formula does not depend on uniform spacing and can be modified at or near a boundary. This is in contrast to the classical theory of discrete wavelets where it is not possible to expand a function by translation and dilation of a single function on a non-uniform grid, and where non-periodic boundary conditions are difficult to implement.

Harten's algorithm is closely related to second generation wavelets that are not necessarily translates and dilates of a single scalar function. Harten's interpolatory wavelets and their generalization to multiwavelets have this property when applied on non-uniform grids. Sweldens describes a lifting scheme for constructing biorthogonal wavelets by generalizing a basic interpolatory multiresolution algorithm. Meanwhile, the support of the wavelet increases while the order of accuracy remains the same as the basic algorithm.

**Wavelet compactness:**

In this paper, we consider compact schemes that have least support for higher order of accuracy. Before this work, several authors have attempted to develop wavelets that may have compact support with higher order of approximation and, if possible, additional beneficial features such as orthogonality and analytic (nonfractal) form. The most fundamental Haar wavelets (first order of approximation) were observed by Strang [8] to be compact, piecewise analytic and orthogonal.

Daubechies [9] extended the order of approximation while retaining the orthogonality, but losing compactness. Geronimo, *et. al* [10] used multi-wavelets, but the results are similar to those found in [9]. Strang and Strela [11] extended the multi-wavelet work of Geronimo, *et. al* [10] for general orders of approximation. Donovan, *et. al* [12] used multiresolution analyses for the construction of multi-wavelets, which have arbitrary regularity and orthogonality, and are also symmetric and piecewise polynomial. Both Strang [11] and Donovan [12] wavelets have compact support to some degree but not to the level of the Haar compactness. As an early trial, Beam and Warming [13] extended Harten's interpolatory multiresolution to include Hermite interpolation, which leads to vector interpolation and multi-wavelets. These wavelets are nonfractal and accurate; however, they are not orthogonal (but biorthogonal) and are less compact than Haar wavelets.

Finally, based on the multi-wavelet basis of Alpert [14], Beam and Warming [15] constructed supercompact multi-wavelets which meet all the above conditions.

As far as supercompact support of the basis is concerned, the advantages of the multiresolution algorithm are two-fold. First, the application to functions defined on a finite interval does not require special treatment at the boundaries of the interval. Second, the application to functions which are only piecewise continuous between internal boundaries, can be efficiently implemented.

In this paper, Beam's supercompact wavelets are recasted in the structure of multiple wavelets and are also generalized to higher dimensions. Demonstration of actual data compression for 3D fluid simulation data with multi-dimensional wavelets is presented.

**Supercompact Wavelets**

As the Haar wavelets can exactly represent any piecewise constant function, the supercompact wavelets can exactly represent any piecewise polynomial functions. It is because supercompact wavelets are based on multi-wavelets (family of wavelets). Multiple wavelets conduct decomposition and reconstruction processes using more than single mother wavelet. They offer fundamental advantages such as orthogonality, symmetry, short support, and a higher degree of accuracy [16]. Because of these beneficial characteristics, the multiple wavelets in general allow better data compression and feature extraction than a single wavelet.

The use of multiple wavelets essentially entails additional pre- and post- processing for transforming given discrete scalar data ( $u(x)_j$ ) to vector quantities $(\vec{\alpha}_j)$. If the original given data has more than one dimension (vector, or matrix), the transformed data has one more additional dimension. Hence, the actual decomposition and reconstruction are performed on this transformed form $(\vec{\alpha}_j)$ of data set. In the following subsections, we begin with a description of pre- and post- transformations (interpolation) by the use of polynomial functions. Next comes the decomposition and reconstruction based on orthogonal polynomial transformation. Although explicit knowledge of wavelet and scaling functions is not necessary for the supercompact wavelet implementation, exact forms of scaling and wavelet functions are also described.

- Data Compression with Hybrid Supercompact Wavelets

Data compression technique is proposed that combines 3D, 2D and 1D versions of supercompact. Computation domain is divided into subsets for each dimensional version of multiresolution application. Compared to Harten's interpolatory multiresolution, supercompact wavelets require small number of support points, leading to quite simple and less expensive numerical implementation. The hybrid multiresolution method allows high actual data compression for fluid simulation data since the method is based on highly effective wavelets and subdivision methods. Several numerical tests substantiate large data compression ratios for flow field simulation.

Recently, the data size from fluid dynamic simulations has sharply increased for many practical applications.

The applications are full-scale integrated aero-vehicle simulation and turbo machinery analysis that consists of many small components such as stator and rotator, etc. Grid size and simulation fidelity levels in governing equations are the crucial factor that determines the size of the yielding data sets.

These big data sets, sometimes, do not fit on main memory, local disks, and possibly even on a remote storage disk. It also creates tremendous rendering time as well as technical difficulties in interactive post-processing/visualizing the data. It also demands a corresponding huge I/O time in solution and post-processing algorithm. Due to limited network bandwidth constraints, transmission time between different computer systems is sharply increasing.

**References**

1. B. Engquist, S. Osher, S. Zhong, Fast wavelets based algorithms for linear evolutions. ICASE Report, 1992.
2. O. Vasilyev, S. Paolucci, A dynamically adaptive multilevel wavelet collocation method for solving partial differential equations in a finite domain. JCP 1996;125:498-512.
3. A. Harten, Multiresolution algorithm for the numerical solution of hyperbolic conservation laws. Comm. Pure Appl Math 1995;48:1305-42.
4. M. Gerritsen, Designing an efficient solution strategy for fluid flows. PhD Thesis, Stanford University, 1996.
5. L. Jameson, T. Adachi, O. Ukai, A. Yuasa, Wavelet –based numerical methods. Int J Computat Fluid Dyn 1998;10:26 7-80.
6. W. Sweldens, The lifting scheme: a new philosophy in biorthogonal wavelet constructions. Proc SPIE 1995;2569:68-79.
7. W. Sweldens, P. Schroder, Building your own wavelets at home. ACM SIGGRAPH Course Note # 13, 1996.
8. G. Strang, Wavelets and dilation equations: a brief introduction. SIAM Rev 1989;31:614-27.
9. I. Daubechies, Orthonormal basis of compactly supported wavelets. Comm Pure Appl Math 1988;41:909-96.

10. J. Geronimo, D. Hardin, P. Massopust, Fractal functions and wavelet expansions based on several scaling functions. J Approx Theory 1994;78:373-401.
11. G. Strang, V. Strela, Orthogonal multiwavelets with vanishing moments. J Opt Engng 1994;33:2104-7.
12. G. Donovan, J. Geronimo, D. Hardin, Interwining multiresolution analyses and the construction of piecewise polynomial wavelets. SIAM J Math Anal 1996;27(6):1791-815.
13. R. M. Beam, R. F. Warming, Discrete multiresolution analysis using hermite interpolation: biorthogonal multiwavelets. Technical report, SIAM J Sci Comput 2000: 22(4):1269-317.
14. B. Alpert, A class of bases in $l^2$ for the sparse representation of integral operators. SIAM J Math Anal 1993;24(1):246-62.
15. R. M. Beam, R. F. Warming, Multiresolution analysis and supercompact multiwavelets. SIAM J Sci Comput 2000: 22(4):1238-68.

### II.E.2.l. Daniel Hammerstrom

*Neural network and associative memory computational models – Barry Leiner.*

Two projects are being worked on collaboratively with NASA groups. The first is "Biological computing for robot navigation and control," with Marwan Jabri, Chris Assad, Dan Hammerstrom, Misha Pavel, Terrence Sejnowski, and Olivier Coenen. This is part of NASA's Intelligent Systems program. We are now entering the 2nd year. For the first year we have build basic simulations which are simplifications of a number of different parts of the brain. Also we developed a virtual robotic like environment that allows us to address sensing and control issues without having to first create a working robot.

The second project is unfunded, but is a program in nanotechnology that we are doing jointly with NASA's Nanotechnology Team at Ames. In particular we have been working with Dr. T.R. Govindan and Manoj Samanta. We have been developing circuit and architectural models of molecular scale circuits for neural network and associative memory computational models. Our group at OGI just recently submitted a large proposal to NSF, "Design and Fabrication of Nanoscale Resonant Tunneling Diode Based Associative Memory," with John L. Freeouf, John Carruthers, Dan Hammerstrom, Raj Solanki, Raphael Tsu. Although the NASA team will not be getting funding they were listed as a major collaborator.

### II.E.2.m. Flavio Lerda

*Java PathFinder – Willem Visser*

During the summer at NASA Ames, I worked on applying the model checker Java PathFinder to programs written in C/C++. Most of the applications relevant to NASA mission are currently written in C or C++, while only a limited number are in Java.

During my stay, I came up with a mapping between C and Java, which can be further extended to handle C++. This mapping allows to define a source-to-source translation from C to Java. The obtained Java source can be compiled and executed to obtain an equivalent behavior.

Moreover, in defining the translation, I paid special attention to analyze the impact the translation will have on the successive model checker run, in particular with regard to the state explosion problem.

In order to make the approach feasible, it is necessary to reduce to the minimum the overhead introduced by the translation, or at least its effects during model checking. To do so, I developed a translation that uses a class library and calls to class methods: these calls can be trapped by the current implementation of the model checker to perform on-the-fly optimization of the model checked code.

### II.E.2.n. Oliver Lemon

*Project: Multi Modal Dialog – Beth Ann Hockey*

We currently work with RIALIST in 3 areas of multi-modal dialogue systems research: grammar development, dialogue management, and "targeted help" for users. This summer we ran user tests of a dialogue system for collaboration with autonomous mobile robots, which showed that use of our "targeted help" module results in 20% better task completion and 33% faster task completion times. Our research in intelligent dialogue management and grammar specialization is central to the ongoing development of portable and modular spoken dialogue systems for support of human-machine collaboration.

### II.E.2.o. David Smith

*Rover autonomy. – Richard Washington, Richard Dearden*

This work centers around the command executive prototype for the K9 rover, which forms a basis for ongoing research projects as well as serving as the function command executive for rover operations. The research projects have concentrated on robust, flexible execution of plans, using a utility-based approach. More recent work has explored on-board plan adaptation and plan merging techniques, which lie somewhere between execution and full-scale planning. The rover autonomy project collaborates with other research projects at Ames, in particular Limited Contingency Planning (D. Smith, NASA ARC), and state identification and fault diagnosis techniques (R. Dearden, RIACS). The project also maintains an ongoing collaboration with S. Zilberstein at the University of Massachusetts, investigating decision-theoretic approaches for more completely autonomous exploration activities.

### II.E.2.p. David Zingg

*Gradient algorithms for optimization – Thomas Pulliam*

The project with Dr. Thomas Pulliam continues our collaboration on the evaluation and characterization of evolutionary and gradient-based algorithms for multi-objective airfoil optimization based on the Reynolds-averaged Navier-Stokes equations governing turbulent viscous flow. The evolutionary algorithm studied is a genetic algorithm developed at the NASA Ames Research Center, while the gradient-based algorithm is a discrete adjoint technique developed at the University of Toronto. The two approaches use a common design space, and the objective function value calculated for a given set of design variables is also identical, permitting direct comparisons of efficiency and accuracy in the computation of Pareto fronts for competing objective functions. A paper based on this collaboration will be presented at the 41st AIAA Aerospace Sciences Meeting.

## *II.F. Inventions*

Based on the work described above as well as work performed in prior years of this cooperative agreement, the following invention disclosures have been filed with NASA.

**Table 8: Invention Disclosures**

| Date | Disclosure No. | Title | Inventors |
|---|---|---|---|
| Oct-01 | ARC-14725-1 | Livingstone 2/ Skunkworks: A System for Automated Diagnosis & Discrete Control of Complex Systems, & a Suite of Supporting Dvlpment & Runtime Tools | P. Pandurang Nayak, RIACS Contractors: QSS Group, Inc., NASA Code GE, and Caelum Research Corp. |
| Oct-01 | ARC-14741-NP | An Algorithm for Automatically Generating Statechart Designs from Requirement Scenarios | Johann Schumann, RIACS and Jonathan Whittle, Computer Scientist |
| Request 8/28/02 | ARC -141517-1CU | Remote Agent Planner/Scheduler: A System for Generating Complex | Paul Morris & Kanna Rajan Contractors: Recom, Caelum, JPL |
| Request 1679 on 8/28/02 | ARC-14441-1 | Method & Apparatus for Virtual Interactive Medical Imaging by Multiple Remotely-Located Users | Ian Twombly and Steven Senger |
| Request 1679 on 8/28/02 | ARC-14489-1 | Ultrasound Segmentation & Echocardiogram Using Stochastic Relaxation | Esfandiar Bandari |
| Jan-02 | ARC-14768-1NP | Polychotomic Encoding: A Better Quasi-Optimal Bit-Vector Encoding of Tree Hierarchies | Robert Filman |
| Jan-02 | ARC-14775-1 | Livingstone to SMV translator | Charles Pecheur & Reid Simmons |
| Mar-02 | ARC-14780 | Livingstone PathFinder: An automated simulation tool for Livingstone applications | Charles Pecheur |
| Sept-02 | Pending ARC Response | A Method to Integrate Information Systems & Simulations or a Simulation Transaction Access & Retrieval Gateway (STARGATE) | David Maluf |
| Sept-02 | ARC-14756-1 | NASA Virtual Glovebox (VGX): Advanced Astronaut Training & Simulation System for Life Science Experiments Aboard the International Space Station | Ian Twombly, RIACS Jeffrey Smith and Richard Boyle, NASA Code SLR |
| May-02 | ARC-14797 | AutoClass: A Bayesian Classifications Systems | Dick White & Justin Black, UC Berkeley |
| May-02 | Pending ARC Response | A Bayesian Method for Super-Resolved 3-D Surface Reconstruction from. Multiple Images | Peter Cheeseman & Robin Morris |
| Request 1679 on 8/28/02 | Pending ARC Response | Polychotomic Encoding: A Better Quasi-Optimal Bit-Vector Encoding of Tree Hierarchies | Robert Filman |

# III. Seminars ,Workshops and Technical Reports

RIACS holds both regularly scheduled and special seminars which bring in speakers from the Ames scientific community, university researchers, and high-tech industry researchers. These seminars provide an excellent forum for education and collaboration with leaders in the field.

RIACS also holds and sponsors workshops that address computer science issues in depth by bringing together NASA and RIACS scientists with leading researchers in the field from universities and industry.

## *III.A. RIACS Seminars*

(#150) Date: Sept. 19, 2002:
Dr. Cynthia Thompson, University of Utah's School of Computing
Title: " Adaptive Conversational Item Recommendation"

I discuss an adaptive, conversational, speech interface system for recommendation and item search tasks, using destination selection as an example. Our focus has been on making the interaction between the system and user more efficient over time due to adjustments on the part of the system using simple machine learning techniques. We view the search for items to recommend as an interactive process of constraint satisfaction, with the advisory system proposing attributes and the human responding. Viewing the conversation in this way allows us to simplify dialogue management, so that most of the work can be performed by a domain independent task management component. The system unobtrusively collects user preferences from conversation logs, and uses them to build user models which guide future conversations. I will present an experimental analysis of the system's ability to change its behavior based on past interactions with a user, and will conclude the talk with a discussion of our future research plans.
Bio Sketch:
Cindi Thompson received her Ph.D. in Computer Sciences from the University of Texas, Austin in 1998 after working with Professor Ray Mooney. After that, she held a postdoctoral research position at the Center for the Study of Language and Information at Stanford University, where she worked in a position created jointly by Stanford and DaimlerChrysler Research and Technology. Her focus was on conversational interfaces that adapt their behavior to the preferences of a user, and she collaborated with Pat Langley, Stanley Peters, and Mehmet Goker. Cindi Thompson joined the University of Utah's School of Computing in August 2000 as an Assistant Professor. Her research and teaching interests include artificial intelligence, machine learning, and natural language processing. She is spending the 2002-2003 academic year as a Visiting Assistant Professor in the Computer Science Department at Stanford University.

(#149) Date: Aug. 27, 2002:
Alan Hudson, Yumetech, Inc
Title: "X3D: The Next Generation Open 3D Graphics Standard"

The X3D Graphics Working Group is designing and implementing the Extensible 3D (X3D) Graphics specification. In this new standard, the geometry and behavior capabilities of the

Virtual Reality Modeling Language (VRML 97) are encoded using the Extensible Markup Language (XML). This presentation will introduce the X3D specification and detail its new features including XML integration, multi-texturing, NURBS and a new scripting interface. Moreover, the specification has been divided into profiles and components to make implementations more market-focused and easily extended. The Xj3D toolkit is an Open Source implementation of the VRML97 and X3D specifications. It is a Java API that makes implementing the new specification easier for application developers. Utilizing a component architecture, it can be used to provide simple features and is designed to render content on devices ranging from mobile phones to fully immersive systems like CAVES or domes. This presentation will show the features of the current system and provide a roadmap of future development.
Bio Sketch:
President: Yumetech, Inc. http://www.yumetech.com/
Web3D Open Source Chair http://www.web3d.org/TaskGroups/source/

(#148) Date: Aug. 22, 2002: SSRP 2002 Student Researchers
Rene Vidal, U.C. Berkeley
Title: Bayesian motion estimation and surface reconstruction
Kate Mullen, Bard College
Title: "Initialization of Model-Based Clustering Algorithms"
Brent Venable, University of Padova
"Finding Optimal and Pareto Optimal solutions of Simple Temporal Problems with Semi-convex Preference Functions"

(#148) Date: Aug. 22, 2002:
Dr. N.S. Sridharan, TrustNet
Title: "TrustNet"

Vision: To advance trust in organizations, communities and societies at large. At present TrustNet is a consortium in its formative stage and the founders are engaged in discussions with trusted colleagues to get feedback and advice on the TrustNet vision, scope of activities, and sources of revenue. Trust Technology for Socio-technical systems. Socio-technical systems can be designed with explicit attention to trust. To do this well, we need to understand the many dimensions of Trust and the dynamics of how trust is eroded or supported by technology. We need to start with a grounding of the core concepts developed about Trust from disparate fields such as psychology, philosophy, social sciences, economics and political science. Many current systems are designed with a technical focus and hence concentrate on creating systems that can be trusted by people. This is one-sided trust. Complementing this is bilateral trust of people (people to people and system to people). Such thinking may be applied to any number of socio-technical systems (Next generation Internet, Air Transportation, Large scale factory systems, ocean going platforms, hospitals etc). Trust is often confused with reliability, safety, security and privacy protection. Those of us who were early users of Arpanet remember the pre-boom Internet as a trusted Internet; our intellectual crown jewels all resided on this network - data, models, code and reports. We examine the possibility of rebuilding a Trusted Internet based on sound

understanding of how trust is generated, maintained and supported. Implications for designing other socio-technical systems will be explored in this seminar.
Bio Sketch:
Dr. Sri (pronounced Shree) Sridharan was the chief architect for knowledge management at Intel until 2000. At Intel he was widely known and recognized for his vision and strategy that balances business savvy with pragmatism and people sensitivity. He is a trained storyteller and is the initiator of various storytelling initiatives for Intel. He also participated in Manufacturing Systems Architecture, Enterprise Application Integration, and Strategic Information Systems. Prior to Intel he spent many years as an academic (Stanford, Rutgers, TU Munich) and managing R&D groups in several industries (BBN, FMC, Intel). His specialty is Artificial Intelligence and his PhD is in Computer Science. His work in AI explored applications in organic chemistry, psychology, genetics, manufacturing and legal reasoning. He was Program Chair of IJCAI-89 and on the editorial board of Artificial Intelligence. He currently is developing a workshop for corporate executive teams on Leadership, Strategy and Trust. He is working a book on the same subject. He is co-founder of a consortium called TrustNet that brings technology to bear constructively on trust, making sense and listening. He is actively involved in several initiatives to improve education and community-based leadership in poor regions of the world.

(#147) Date: Aug. 13, 2002:
Robert Eklund, Telia Research, Sweden
Title: "Disfluencies in Spontaneous Speech - A Speech Technology Challenge"

In recent years, automatic speech recognition (ASR) systems have attained accuracy levels on constrained tasks that are sufficient for many commercial purposes. However, for more open-ended speech input, robustness remains an important issue. One important question to be addressed is the processing of disfluencies (DFs), i.e. phenomena like pauses (filled or silent), repetitions, truncated words, repairs, and so on, which occur frequently in spontaneous speech, which is why modeling of DFs is crucial for automatic systems that interact with humans in spoken language. Disfluencies have been studied both within and across languages, as well as within and across domains and modalities. This talks has two different parts. First, a general introduction to DFs is given, including a typology of different, as well as different implications for different fields of research, as well as an introduction to different data collection methods, and various ways to transcribe and annotate the data collected. Second, disfluency research at Telia Research AB, Sweden, is described in some detail. Four travel dialogue corpora are described, with four different settings: Human--"Machine"--Human (Wizard-of-Oz); Human--"Machine" (Wizard-of-Oz); authentic Human--Human and authentic Human--Machine. Results on five different kinds of disfluencies are presented: filled and unfilled pauses, prolonged segments, truncations and explicit editing terms. Also, a brief discussion on cross-language aspects of DFs is given, with examples from American English, Swedish and Tok Pisin.

(#146) Date: Aug. 8, 2002: SSRP 2002 Student Researchers
Cleidson de Souza,U.C Irvine
Title: "A Field Study of Collaborative Software Development Teams"
Sarfraz Kurshid, MIT
Title: "Symbolic Execution for Concrete Test Case Generation"

Frank Hutter, Darmstadt Univ. Of Technology
Title: "Model-based fault diagnosis and state estimation for planetary rovers"
Alex Groce, CMU,
Title: "What Went Wrong?"

(#145) Date: Aug. 1, 2002: SSRP 2002 Student Researchers
Laurentiu Leustean, University of Bucharest
Title: "Certifying Kalman filters"
Flavio Lerda, CMU
Title: "Towards Model Checking of C Programs"
Dan Bohus, CMU
Title: "Spoken Dialog Management for an Astronaut's Procedure Assistant"
Jamieson Cobleigh, University Massachusetts
Title: "Assumption Generation for Compositional Model Checking"

(#144) Date: July 25, 2002: SSRP 2002 Student Researchers
Max Horstmann, Univ. of Massachusetts
"The Markov Decision Process Approach on Planetary Rover Contingency Planning"
Judah De Paula, University of Texas
"Partitioning Algorithm for Data Tracking"
Nate Blaylock, University of Rochester
"Integrating Planning and Execution for a Dialogue Interface with Autonomous Agents"
Ellen Campana, University of Rochester
"Eye Tracking Technology in Speech Recognition Interfaces"

(#143) Date: July 18, 2002: Dr. Gautam Biswas, Vanderbilt University
"An Integrated Approach to Fault Detection and Isolation in Complex Hybrid Systems"

The need for reliability and robustness in present day systems requires that they possess the capability for accommodating faults in the controlled plant. Fault accommodation requires tight integration of online fault detection, isolation, and identification with the system control loop. This thesis presents a model-based approach to online fault detection, isolation, and identification in complex systems. The plant models for such systems are necessarily hybrid, i.e., their behavior evolution combines continuous operating regions (modes) interspersed with discontinuous changes that model mode transitions (the result is a discrete change in the continuous model). Hybrid models form the natural representation for embedded systems in avionics, automotive, and robotics domains. The wide applicability of hybrid systems has inspired a great deal of research from both control theory and theoretical computer science. A model-based approach to fault detection and isolation of hybrid systems presents an interesting set of challenges that mostly revolve around interactions of the continuous and discrete components of the system. The tracking of the system behavior evolution has to be performed across modes of operation. This requires continuous tracking, identification of discrete changes, and updating the model and state after a discrete change. The fault isolation has to reason across modes of operation to identify hypotheses that can explain all deviant observations. This may involve rolling back in the mode space to generate hypotheses and then rolling forward in the mode space to catch-up to the current system mode of operation. For fault accommodation, the quantitative value of the fault has to be determined so that appropriate corrective action may be taken. We present an integrated diagnosis architecture that tracks the hybrid system, and detects, isolated and identifies the fault. We use hybrid bond graphs as a comprehensive modeling framework from which models for the individual components of our diagnosis architecture are derived. We use these models to develop diagnosis algorithms that combine hybrid behavior tracking with mode detection and combined qualitative-quantitative reasoning techniques in the continuous domain. The effectiveness of the approach is demonstrated for a realistic example: the fuel transfer system of aircraft. These techniques are now being applied to a NASA application, the Water Recovery System (WRS) of the Bio-Plex system for long-term manned missions.

Bio Sketch:

Gautam Biswas is an Associate Professor of Computer Science and Engineering, and Management of Technology at Vanderbilt University and a faculty associate at the Institute for Software Integrated Systems (ISIS) in the School of Engineering. He has a Ph.D. degree in Computer Science from Michigan State University in E. Lansing, MI. Prof. Biswas conducts research in Intelligent Systems with primary interests in hybrid modeling and analysis of complex embedded systems, and their applications to diagnosis and fault-adaptive control. As part of this work, he is working on fault-adaptive control of fuel transfer systems for aircraft, and the water recovery system of the Bio-Plex system. He is also initiating new projects in distributed monitoring and condition-based maintenance. In other research projects, he is also involved in developing simulation-based environments for learning and instruction, decision-theoretic planning and scheduling techniques for intelligent manufacturing systems, and Hidden Markov Model techniques for clustering of temporal data sequences. His research is currently supported by funding from NASA, DARPA, and the NSF. Dr. Biswas has served on the Program Committee of a number of conferences. He was chair of the 1997 IJCAI Workshop on Engineering Problems for Qualitative Reasoning, co-chair of the 1996 Principles of Diagnosis Workshop, the 1999 AAAI Spring Symposium on Hybrid Systems and AI, the 2001 Workshop on Qualitative Reasoning, Senior Program committee for AAAI-97 and AAAI-98, and Technical Committee co-chair for the 2000 IEEE SMC conference. He is currently an Associate Editor for the IEEE Transactions on Systems, Man, and Cybernetics and the International Journal of Applied Intelligence. He is also a guest editor of a IEEE Transactions on Systems, Man, and Cybernetics Part B special issue on "Diagnosis of Complex Systems: Bridging the methodologies of the FDI and DX Communities." Currently, Dr. Biswas is a Senior member of the IEEE Computer Society, ACM, AAAI, and the Sigma Xi Research Society.

Contact Information:

Gautam Biswas
Dept. of Electrical Engineering and Computer Science
Box 1824, Station B
Vanderbilt University

(#142) Date: July 11, 2002: Mark Tuttle - Apelon
" Explaining Drugs to a Computer: Building NDF-RT - A Reference Terminology for Medications"

Drugs are dangerous, expensive and revolutionizing healthcare. Computers need to help us reduce errors, improve quality and manage costs, but they cannot do with this productively without an ability to "interoperate by meaning." Specifically, if a patient's chronic care medication list is in one computer, and his or her acute care medications are being entered into another, the latter computer's decision support can act on both lists only if there is sufficient interoperation by meaning. At present, the best way for two computers to agree that they "understand drugs" is for them both to use the same Reference Model. The VHA (Veterans Health Administration) and Apelon are building such a Model. Beginning with NDF (National Drug File), the current VHA Formulary, we are creating an ingredient-centric, semantic, i.e., computer-understandable, definition for each drug based on chemical structure, mechanism of action, physiologic effect, pharmaco-kinetics/dynamics, and therapeutic intent. The Model allows the more than 80K orderable drugs at VHA to "inherit" the properties of their active ingredients, except when the particular "dose form" dictates otherwise. The initial version of NDF-RT has been computed from NDF and the UMLS Metathesaurus. It can be browsed from the Web. Human editors are reviewing the next version. Of potential relevance to Astro-Biology and Bio-Terrorism is the instantiation of this model on all "biologically active" substances. Attendees at a recent meeting on "Drug Informatics" agreed that the relevant portion of the schema for a semantic definition of medications could and should be applied to all other biologically active substances.
Bio Sketch:
Mark S. Tuttle is Vice President of Strategy at Apelon (www.apelon.com): Currently, Mark directs a number of efforts to create, deploy, maintain, and use biomedical Reference Terminologies. Mark is a co-founder of Lexical Technology - which later merged with Ontyx to become Apelon - and he proposed and then led the initial development of the NLM UMLS (Unified Medical Language System) Metathesaurus. Prior to founding Lexical, he helped the University of California, San Francisco win and execute a first round UMLS award, while teaching computer science at University of California, Berkeley. He studied computer science, applied mathematics, and information theory at Harvard University and Dartmouth's Thayer School of Engineering. Mark was elected to be a Fellow of the American College of Medical Informatics (FACMI) in 1993.

(#141) Date: July 3, 2002: Dr Dominique Estival, Defence Science & Technology Organisation, Edinburgh, AUSTRALIA
" Developing Automated Spoken Language Dialogue Systems"

The main focus of my talk will be the work of the NLP group at Syrinx Speech Systems, and in particular I will describe Sylan, the automated dialogue system we developed. Sylan was fully integrated with the Sycon speech recognizer into a platform combining speech recognition, natural language processing, dialogue management, telephony and database integration and permitting the deployment of natural language dialogue systems in automated call centres. The aim was to produce a framework to build applications where the structure of the dialogue can be less constrained than current commercial directed systems, and which allow users to input more natural, multi-token utterances which are interpreted and processed in several stages. I will first describe the architecture of Sylan, which, being modular, allowed us to build a system with domain-independent components reusable from application to application. I will then present

those components from the point of view of application developers, describing the data structures used by the system and the utilities to build them, and drawing examples from the two prototypes we developed. I will conclude by discussing the constraints imposed on the development of commercial systems, and by drawing the lessons learned along the way to point to further research directions.
Bio Sketch:
Human Systems Integration Group
Information Technology Division
Defence Science & Technology Organisation
PO Box 1500
Edinburgh SA 5111
AUSTRALIA
Phone: +61 (08) 8259 6485 Fax: +61 (08) 8259 5589
Email: dominique.estival@dsto.defence.gov.au

(#140) Date: July 3, 2002: Dr. Kamal Ali, Vividence
" On explaining the degree of error reduction by combining classifiers"

In the past decade, one of the most interesting developments in Machine Learning has been that combining classifiers in various ways produces ensembles that have lower classification error rates. Looking one step further, I will present results that explain why combining classifiers to produce more accurate classifications helps in some problems much more so than in other problems. The work examines how the noise inherent in the problem limits the lift due to ensembles, how the dimensionality of the problem affects error reduction and the effect that diversity of classifiers has on ensemble error rate. I will also briefly present results from two applications. The first involves recognizing mineral signatures from infrared reflectance spectra using learning expert systems. The second application researched the efficacy of active learning to speed-up human labeling of satellite images and compared a naive Bayes classifier to within- and cross-human labeling variance.
Bio Sketch:
Dr. Kamal Ali received his Ph.D. in Computer Science from the University of California, Irvine, in December 1996. His dissertation was on learning probabilistic first-order theories but his later work was on learning classifier ensembles. He has published numerous papers in Data Mining and Machine Learning conferences, reviewed for ICML, KDD, PAMI and MLJ. He is the recipient of a NSF scholarship, UC Regents fellowship and the Horner Exhibition prize in Math at Sydney University. After graduation, at IBM Almaden, he published papers and served as a data-mining consultant. At TiVo he led the team to implement collaborative filtering for TV recommendation used in 400,000 homes. At Vividence he developed a divide-and-conquer system for clustering clickstreams and developed a system for clustering free text. His current research interests include combining classifiers, developing learning algorithms for image analysis, Bayesian statistics, web-mining and text-mining. He is interested in getting involved in mining of scientific datasets and applying learning to space images.

(#139) Date: June 27, 2002: Dr. Venkataram Sundareswaran, Rockwell Scientific
" Human-Computer Interaction and Augmented Reality Research at Rockwell Scientific"

Human Computer Interaction (HCI) work at Rockwell Scientific Company (RSC) includes basic research in developing new HCI tools and integration of multiple modalities to create intuitive interfaces. Basic research includes Speech Recognition in noisy environments and Augmented Reality (AR). In multimodal interaction, research is focused on developing "point and speak" and "look and speak" interfaces, and in developing an Integrated Displays Testbed that includes various levels of displays (handheld, tablet, large screen) and interaction among them. The talk will begin with a broad introduction to the HCI work at RSC and a video presentation that summarizes this work. The second part of the talk will be focused on Augmented Reality technology, with an introduction to AR followed by recent results achieved by the HCI team at RSC. In particular, an approach to solve the "registration" problem, namely that of aligning real world view with virtual graphical objects, based on the technique of visual servoing, will be presented.
Bio Sketch:
Dr. Sundareswaran leads a team of eight HCI researchers at RSC's Information Sciences division. In this role, he manages Government and corporate research projects. His background is in Computer Vision and Graphics, and has authored numerous publications in these areas. He graduated from NYU in 1992, and after a brief stint as a post-doctoral fellow at INRIA (France), he spent three years at Boston University in Neural Network modeling of visual motion perception. He joined RSC in 1996.

(#138) Date: June 13, 2002: Jerry Toung, RIACS NASA Ames Research Center
" PCMon - A Network Monitoring Tool"

PCMon is a network measurement and monitoring tool that is being developed and deployed on the NASA Research and Education Network (NREN) testbed, a research testbed used for prototyping emerging networking technologies. A new tool was required to enable evaluation of the effectiveness of these new technologies. For example, Quality of Service (QoS) is an emerging technology that is designed to provide different levels of service for different traffic flows, a capability that enables efficient sharing of network resources among multiple users while providing preferential treatment to selected applications when network resources become scarce. Traditional network monitoring and measurement tools capture statistics on aggregate traffic only, rather than individual traffic flows, and hence would be unable to verify that "preferential treatment" is actually delivered as promised by the QoS mechanism. During this presentation, I will to talk about how PCMon has evolved as a functional tool and how it has already been used to monitor QoS requirements on some NASA's applications. I will also do a live demonstration of its capabilities and close with few notes on future directions.
Bio Sketch:
I hold a Master Degree in Electrical Engineering and specialize in network software engineering. I have been with RIACS at the NASA Research and Education Network since January 2000. My interests include network programming and device drivers development for FeeBSD.

(#137) Date: June 7, 2002: Mohammed El-Beltagy, BiosGroup, Inc., in Santa Fe, NM
"Dialog systems that learn"

In many design optimization problems, the designer is faced with the dilemma of how to simulate the problem at hand using a number of different models. Some models maybe quite elaborate in their representation of the problem and hence tend to be computationally expensive. Other models may be far less elaborate and hence computationally cheaper. The computationally cheap models tend to be less accurate than the expensive ones. The designer uses his/her experience, and understanding of the problem domain to switch between different models. S/He goes through a few iterations till a satisfactory design is found. Designs created in such a fashion are not necessarily optimal and they could be improved upon, given more design iterations and an adequate search technique. It is hence important to develop techniques that make maximal use of the many models available within a limited computational budget. Conducting search on such an environment where there are multiple models for evaluating fitness is what is meant by the term Multilevel optimization (MLO).<?xml:namespace prefix = o ns = "urn:schemas-microsoft-com:office:office" /> Suitable methods for conducting MLO maybe sought using algorithms and techniques gleaned from natural process, mainly Evolutionary Algorithms and Artificial Neural Networks. In this presentation, an exposition will made of the issues to be considered when carrying out multilevel optimization. This will be followed by a comparison of how various optimization algorithms perform on a multilevel problem using three simple model selection strategies. Having established that evolutionary inspired search methods work well in such an environment, a topological mapping based model selection approach is then presented. Finally, Gaussian processes based metamodeling and model fusion approaches are explored. We will show that there are significant gains to be made in the synthesis between optimization and machine learning techniques for MLO.
Bio Sketch:
Mohammed El-Beltagy received the B.Sc degree in Mechanical Engineering from the American University in Cairo, Cairo, Egypt, in 1994, the M.Sc. degree in Mechatronics from Lancaster University, Lancaster, U.K., in 1996, and the Ph.D. degree in Mechanical Engineering from the University of Southampton, Southampton, U.K., in 2000. His PhD. was on multilevel optimization (MLO), where the goal was to optimize an engineering design, the performance of which can be evaluated by a number of simulation models that a tradeoffs between computational speed and accuracy. During the course of his PhD. work, he devised a number of successful strategies to MLO that involved a tight coupling between machine learning and optimization. His work at Southampton was funded by the Engineering and Physical Sciences Research Council (EPSRC), and British Aerospace plc. He is currently a Senior Scientist at BiosGroup, Inc., in Santa Fe, NM. Since joining BiosGroup, he has generated IP around automated financial trading and high dimensional matching, lead a project that devised a new algorithm for the optimization of trading networks for a natural gas pipeline operator. In an extended distributed control problem, Mohammed developed methodologies for self-organizing, self-healing data networks for a large mass storage software provider. His current research interests include search theory, and utilizing machine learning techniques for intelligent optimization and search landscape feature elicitation.

(#136) Date: June 6, 2002: David Kil, Rockwell Scientific Company
"Easy-to-use, integrated data mining"

Data mining is considered as an esoteric, domain-dependent art. In this talk, I will explore several key issues in designing an easy-to-use, integrated, and scalable data-mining tool that can be implemented in parallel computers with coarse granularity. The first issue deals with metadata creation by marrying signal processing to data mining in the context of preprocessing, which is the most important step in data analysis. Next, I will discuss automatic database exploration to find entity relations and discover meaningful relationships autonomously. Third, I will investigate the components of an intelligent data-mining expert engine that can guide novice users through a myriad of seemingly complex steps in data mining. Finally, I will demonstrate how these salient concepts can be applied to high-throughput gene-chip image analysis, mine countermeasure, and diagnostic applications. In summary, we will together explore what it will take to change the perception of data mining from an abstract art to a well-understood science.
Bio Sketch:
David Kil received his BS in EE and Chemistry at the University of Illinois at Urbana-Champaign (Bronze Tablet), MSEE from the Polytechnic University of New York, and MBA from Arizona State University. His primary research interest is in integrated data analysis that combines salient concepts from signal processing, image understanding, data mining, and fusion to extract the maximum amount of useful information from raw data. He has managed over ten 6.1 and 6.2 research programs, dealing with information extraction and knowledge discovery. He is currently leading a data-mining effort at Rockwell Scientific and consults for two biotech companies (Arcturus and Gene Networks) in high-throughput gene-chip image and tissue analyses. He has published over 25 papers and one research monograph titled Pattern Recognition and Prediction with Applications to Signal Characterization by Springer-Verlag.

(#135) Date: June 5, 2002: Asok Srivastava, Ph.D, Blue Martini Analytic Services
"Discovering Regimes in Time Series"

Many real-world time series are multistationary, meaning that the dynamics of a data generating process switches its mode of behavior. These modes of behavior are sometimes referred to as "regimes," and could manifest as a shift in the mean, variance, or some other statistic. Two key problems that arise in the analysis and prediction of such systems are: · To identify the time at which a system undergoes a mode change, and · To model the underlying dynamics of the system after the mode change. In an attempt to address these issues, we will discuss a model that builds on two model classes: mixture models and thermodynamic clustering-models. Using a maximum-entropy framework, we will derive association probabilities, a quantity which represents the probability of associating an input-output pair to a local model. Unlike standard mixture models, these association probabilities are parameterized by a scale parameter that allows us to sweep through multiple time series segmentations. The different segmentations that arise can shed light onto the two problems described above. Applications of such technology arise in a diverse set of fields, including engineering and aeronautical systems, economics, and physiology. We will describe the performance of this model on synthetic and real-world time series.
Bio-sketch
Ashok N. Srivastava, Ph.D. is currently Senior Director of Blue Martini Analytic Services and has fourteen years of experience in research and development in the fields of data analysis and

machine learning, data mining, signal processing, and applied physics. Some recent research activities include the development of patent pending algorithms for understanding nearest neighbor predictions, methods of clickstream analysis and visualization, creation of new methods to profile populations based on naive Bayes techniques, and the deployment of research results in a variety of industries and government.

(#134) Date: May 31, 2002: Dr. Richard Rohwer, HNC Software
"Information-theoretic clustering of probability distributions for semantic representation"

This presentation will begin with a brief overview of statistical pattern recognition research at HNC, stressing my particular interests, and then focus in on a specific body of work: a system for clustering probability distributions. This work is interesting for its motivation and for its technical components. The motivation involves the use of probability distributions to embody a very general concept of semantics. The algorithm simultaneously aggregates 2 categories of values into bins with minimal loss of statistically significant mutual information between the two. The number of bins is self-adjusting, trading statistical significance against information preservation. The method is formulated in the multinomial-Dirichlet framework, using the Evidence approximation to adjust the hyperparameters, after verifying the validity of the approximation in this setting. An adaptive simulated annealing procedure is used for training.
Bio Sketch:
Richard Rohwer graduated with a BS in Physics from Stanford in 1978 and a PhD in Physics from the University of Texas at Austin in 1985, with a thesis on quantum cosmology. Following a 1-year PostDoc at the University of Newcastle upon Tyne in England, Dr. Rohwer moved into Neural Networks, working at the Centre for Speech Technology Research of Edinburgh University until 1991. He joined the faculty of Aston University in Birmingham, where he was instrumental in setting up their then new Neural Computing Research Group. In 1996, he moved to Prediction Company in Santa Fe, NM, and on to HNC Software in San Diego in 1997, where he is now a Principal Scientist in the Advanced Technology Solutions Department. Dr. Rohwer currently leads grant-funded research involving diverse applications of statistical pattern recognition, including topics in natural language, machine vision, and bioinformatics.

(#133) Date: May 30, 2002: Dr. Robert Filman, RIACS - NASA Ames Research Center
"Polychotomic Encoding: A Better Quasi-Optimal Bit-Vector Encoding of Tree Hierarchies"

Polychotomic Encoding is an algorithm for producing bit vector encodings of trees. Polychotomic Encoding is an extension of the Dichotomic Encoding algorithm of Raynaud and Thierry. Polychotomic and Dichotomic Encodings are both examples of hierarchical encoding algorithms, where each node in the tree is given a "gene"---a subset of the integers {1,...,n}. The encoding of each node is then the union of that node's gene with the genes of its ancestors. Reachability in the tree can then be determined by subset testing on the encodings. Dichotomic Encoding restructures the given tree into a binary tree, and then assigns two bit, incompatible (chotomic) genes to each of the two children of a node. Polychotomic Encoding substitutes a multibit encoding for the children of a node when the restructuring operation of Dichotomic Encoding would produce a new heaviest child (child requiring the most bits to represent a tree of

its children) for that node. We prove that Polychotomic Encoding never produces an encoding using more bits than Dichotomic Encoding. Experimentally, Polychotomic Encoding produces a space savings of up to 15% on examples of naturally occurring hierarchies, and 25% on trees in the randomly generated test set.
Bio Sketch:
Robert Filman is a Senior Scientist at the Research Institute for Advanced Computer Science (RIACS) at NASA Ames Research Center, working on frameworks for developing distributed applications. Prior to coming to NASA in May 1999, Dr. Filman worked in the research groups of Lockheed Martin, IntelliCorp and Hewlett-Packard, and on the faculty of the Computer Science Department at Indiana University, Bloomington. He is Associate Editor-in-Chief of IEEE Internet Computing and is on the editorial boards of the Journal of Software Maintenance and Evolution and the International Journal of Artificial Intelligence Tools. Dr. Filman received his B. S. (Mathematics), and M.S. and Ph. D. (Computer Science) from Stanford University.

(#132) Date: May 16, 2002: Nick Kingsbury, Reader in Signal Processing, University of Cambridge, UK
"Image Processing with Complex Wavelets"

This talk will describe the Dual Tree Complex Wavelet Transform (DT CWT), which is a form of discrete wavelet transform, which generates complex coefficients by using a dual tree of wavelet filters to obtain their real and imaginary parts. This introduces limited redundancy (2^m:1 for m-dimensional signals) and allows the transform to provide approximate shift invariance and directionally selective filters (properties lacking in the traditional wavelet transform) while preserving the usual properties of perfect reconstruction and computational efficiency with good well-balanced frequency responses. We will then describe briefly why the DT CWT is particularly suitable for images and other multi-dimensional signals, and discuss some applications of the transform that take advantage of its unique properties. In particular we will consider its application to denoising, deconvolution, texture analysis, segmentation, content-based retrieval and watermarking.
Bio Sketch:
Nick Kingsbury received the honors degree in 1970 and the Ph.D. degree in 1974, both in electrical engineering, from the University of Cambridge. He is a member of the IEEE. From 1973 to 1983 he was a Design Engineer and subsequently a Group Leader with Marconi Space and Defence Systems, Portsmouth, England, specializing in digital signal processing and coding, as applied to speech coders, spread spectrum satcomms, and advanced radio systems. Since 1983 he has been a Lecturer in Communications Systems and Image Processing at the University of Cambridge and a Fellow of Trinity College, Cambridge. He was appointed to a Readership in Signal Processing in 2000. His current research interests include image compression, error-robust source coding techniques, and image analysis and enhancement techniques, particularly those based on wavelet decompositions. He is especially interested in the application of complex wavelets to images and 3-D datasets.
E-mail: ngk@eng.cam.ac.uk
Web home page: http://www.eng.cam.ac.uk/~ngk

(#131) Date: May 10, 2002: Dr. William Macready, VP Science at BiosGroup

"Optimization Over Sequence Spaces"

Landscapes are functions defined over graphs and have proven themselves to be a useful tool for the design and analysis of local search algorithms for combinatorial optimization. Recent advances in kernel methods from machine learning offer the promise of learning landscapes thereby allowing learning to greatly improve optimization. I will discuss beginning steps in this direction for optimization problems defined over sequence spaces. Optimization over sequence spaces includes many important problems from physics (e.g. spin glasses) and biology (e.g. protein folding)

Bio Sketch:

Dr. William Macready is currently VP Science at BiosGroup, a consulting company applying insights from complexity science, simulation, machine learning, and optimization to solve difficult business problems. William's interests center on the probabilistic inference for the design of efficient optimization and machine learning algorithms. Before joining BiosGroup William worked at the Santa Fe Institute for complex systems as a postdoctoral fellow and at IBM research as a scientist. He has published on the theory of landscapes, molecular evolution, adaptive organizations, economics, optimization, machine learning, and methods of quantifying complexity.

(#130) Date: May 6, 2002: Alexander Rudnicky, School of Computer Science, Carnegie Mellon University

"Dialog systems that learn"

Descriptions of dialog systems commonly focus on core components such as dialog management and understanding. Yet to develop a working system capable of non-trivial operation, a great deal of effort is also spent on defining and implementing the interface between the dialog system and the domain (or "back-end"). Often this aspect comes to dominate the total effort devoted to implementation. We have begun to explore how corpus-based learning techniques can be brought to bear on this problem. This talk will define the problem and will describe our current work on the discovery of operative concepts from transcripts of goal-directed human-human conversation.

(#129) Date: May 2, 2002: Dr. Manny Rayner, Netdecisions Technology Centre, Cambridge, England

"Applying Explanation Based Learning to Speech Recognition"

Accurate recognition is essential for speech-enabled command and control tasks. The required accuracy is made possible by using a suitable grammar for constraining the space of recognition hypotheses. For complex applications, construction of this grammar represents a major investment of effort. There is consequently a strong motivation to develop general grammars that can easily be reused between applications. Experience shows however that such general grammars tend to have serious problems with both efficiency and scalability. In this talk, we will show how it is possible to use Explanation Based Learning (EBL) to specialize a general grammar coded in a high-level, logic-based formalism against a corpus of domain-specific examples. A series of experiments, carried out on the RIALIST group spoken interface to a simulated version of the Personal Satellite Assistant, suggest that specialized grammars have

significantly better run-time properties and strikingly better scalability than the general grammars from which they were derived. In particular, the relationship between the number of rules in a grammar and the time required to compile it into an executable form appears to be roughly quadratic for an EBL specialized grammar, compared to roughly exponential for a general grammar. (Joint work with Beth Ann Hockey and John Dowding, RIACS)
Bio Sketch:
Manny Rayner is Senior Architect for Voice System Design at the Netdecisions Cambridge Technology Center, where his work focuses on development of tools for implementation of commercial voice-enabled systems. Prior to joining Netdecisions, he headed the RIACS Spoken Language group (RIALIST). He also worked for eight years at SRI International, where he was project lead on the Spoken Language Translator, one of the world's first large-scale automatic speech translation projects. He has over fifty publications, including a book on the Spoken Language Translator which was published last year by Cambridge University Press. He has a B.A. in Mathematics from the University of Cambridge and a Ph.D. in Computer Science from Stockholm University.

(#128) Date: May 2, 2002: Dr. Yuval Davidor, Schema
"The Use of Innovative Evolutionary Computation to Solve Intractable Problems"

Schema is one of the first and few commercial companies in which its business and core technology was founded on evolutionary computation (EC). Founded in 1994, the Schema group's main purpose was to solve intractable problems in science, industry and commerce. During this time the group addressed various projects such as static and dynamic missile balancing, warehouse shelf space management, container stowage optimization, and NMR field optimization. In 1997 Schema expanded into the wireless industry and introduced its Falcom platform for spectrum management. To date, about 30% of North American cellular subscribers obtain wireless services that have been optimized for capacity and quality of service by the Falcom platform. Based on Schema's optimization technologies that employ network modeling and simulation capabilities, the platform eliminates the tedious trial and error processes required for manually configuring a network for optimal performance. Schema's solutions are deployed and benchmarked by leading wireless operators worldwide, including Verizon Wireless, Cingular Wireless, U.S. Cellular, and BellSouth International. The presentation will briefly cover past EC applications that the group has developed and will mainly focus on the current Falcom product for optimal frequency allocation and spectrum management.
Bio Sketch:
Dr. Yuval Davidor is the founder of Schema. Previously he was a scientist in the department of computer science at the Weizmann Institute and was a member of the School for Advanced Studies of the Hebrew University. Davidor has authored and co-authored books, academic publications, and several patents in the field of evolutionary computation and optimization. Since 1991, Davidor has devoted his career to the practical application of optimization technologies to solve intractable problems ranging from defense applications to logistics and telecom. Davidor has a bachelor's degree in Engineering from Tel-Aviv University and a doctorate in artificial intelligence from Imperial College, the University of London.

(#127) Date: April 23, 2002: James Allen

"Towards A Generic Spoken Dialogue System, University of Rochester"

While there is great interest and activity in building spoken dialogue systems today, most applications involved very limited domains that require no significant reasoning. Our goal is to design and build systems that approach human performance in conversational interaction in domains that require significant reasoning. We limit our study to "Practical dialogues": dialogues in which the conversants are cooperatively pursuing specific goals or tasks. These include planning (e.g., designing a kitchen), information retrieval (e.g., finding out the weather in New York), customer service (e.g., booking an airline flight), advice-giving (e.g., helping assemble some modular furniture) or crisis management (e.g., a 911 center assistant). In fact, our belief is that the class of practical dialogues includes most anything about which people might want to interact with a computer. While each of these different genres of tasks require significantly different reasoning components and have different structures, we believe that we can develop an generic model of practical dialogue systems that enables us to build domain-independent components that can relatively easily be adapted to different domains. I will describe our work so far and illustrate with examples from some systems we have built over the past five years.
Bio Sketch:
James Allen received his Ph.D. in 1979 from the University of Toronto. He currently holds the position of Professor (87-present) and holds the John H. Dessauer Chair at the University of Rochester (92-present); formerly he was Department Chair (87-90); Associate Professor (84-87), and Assistant Professor (79-84). He served as Editor-in-Chief, Computational Linguistics (83-93) and was a Presidential Young Investigator (84-89). Dr. Allen is the author of Natural Language Understanding, and Reasoning About Plans, and co-editor of Readings in Planning. Dr. Allen is a Fellow of the AAAI. James Allen's research interests lie at the intersection of language and reasoning, and span a range of issues including natural language understanding, dialogue systems, knowledge representation, common-sense reasoning and planning. In the last five years, he has been focusing on designing and building end-to-end spoken dialogue systems that require and exploit common-sense reasoning to collaborate with the user in problem solving. The Rochester Intelligent Planning System (TRIPS) is a planning assistant that can converse in spoken natural language with a person to create, discuss and evaluate various plans involving freight shipments by train. No prior training on how to interact with the system is required.

(#126) Date: April 22, 2002: Supratik Mukhopadhyay, University of Pennsylvania
"Model Checking, Program Analysis and Constraint Databases"

Bugs in unverified (software) systems can cause disasters ranging from rebooting a PC to the failure of a space mission. Model checking is an automatic technique for verifying systems in which a desired behavior of a system is verified over a given system (the model) through exhaustive enumeration of all states reachable by the system and the behaviors that traverse through them. Program analysis refers to the technique(s) of automatically ascertaining information about a program without actually running the program. Constraint databases tightly integrate database and constraint solving methods thereby bridging the gap between efficient, declarative database programming and efficient constraint solving. We establish connections between the seemingly different fields of model checking for infinite state systems, program analysis and constraint databases. This connection allows us to derive uniformly solution large

number of problems in verification of software. In particular, we derive uniformly, symbolic and (in most cases) local algorithms for interprocedural dataflow analysis, points-to analysis, aliasing analysis, automatic checking of array bound violation and other memory errors in C programs as well as automatic verification of safety and liveness properties of embedded software. I will also show how to seamlessly integrate deductive reasoning and abstract interpretation techniques within our methodology. The combined "framework" is used to derive "lightweight" tools for automatically verifying (rather falsifying = finding bugs in) software. I will share my experiences in developing and using a tool based on the described methodology for automatically discovering bugs in C programs.
Bio Sketch:
Supratik Mukhopadhyay received his PhD in Computer Science from the Max Planck Institute for Computer Science, Saarbruecken, Germany, in May 2001. He started as a postdoctoral researcher at the Department of Computer and Information Sciences at the University of Pennsylvania from June 2001. His research interests include program analysis, embedded systems, software engineering and distributed systems.
(#125) Date: April 18, 2002: Michael Sims, Code IC/Center for Mars Exploration, ARC
"Long Day's Drive: Advanced IT enabled long distance rover traverse of Mars"

NASA is expected to soon call for proposals for Mars missions for launched in 2007. A team including Ames, CMU and Ball Aerospace will propose a long distance rover to explore the polar layered deposits in the northern region of Mars. In this talk I will describe that proposed mission, called Long Day's Drive, and will describe how clever design decisions allow mobility of 100 km or more. Central to this approach is relatively modest capabilities for rover 'self safing' autonomy. We need to have great confidence in the reliability of these self safing systems. Additionally, for Long Day's Drive some autonomy elements are more important and some less important than for the currently planned Mars Smart Rover which is to launch in 2009. Due to Long Day's Drive's long traverse over the polar layered deposits it will be possible to gather information that is otherwise obtainable by very deep drilling. We gather that information with considerably less complexity than drilling systems require. In this talk I will discuss the mission simplifications that have led to more autonomy with less effort and I will discuss those areas where advanced autonomy can significantly enhance the science return.

(#124) Date: April 12, 2002: Alexander Verbraeck, Delft University of Technology
"Testing logistic control systems through advanced use of simulation"

Faculty of Technology, Policy and Management Systems Engineering Group
P.O. Box 5015, 2600GA Delft, THE NETHERLANDS
e-mail: a.verbraeck@tbm.tudelft.nl
http://www.tbm.tudelft.nl/webstaf/alexandv
Control systems for logistic and transport systems are extremely complex. Currently control systems are usually fully tested for the first time at the shop floor after commissioning. This usually results in a lot of costly failures that occur at the start-up stages of control systems, after all investments in the physical equipment have taken place. In our research, simulation plays an extended role that simulation can play in testing of fully automated logistic systems and their control systems before commissioning. We follow a three-step approach in testing both logistic

and control systems, where a component based model of the control system and physical system plays an important role. A simulated control system is used to control simulated, emulated, and real prototypes of logistic resources. Until now, we tested our approach with different simulation COTS packages such as Simple++ / eM-Plant, AutoMod, and Arena. For one project for real-time control of free-ranging vehicles the control system has been implemented in all three simulation packages to control logistic resources at a vehicle TestSite. The TestSite is a special laboratory for testing new technologies in logistic automation, where we have 10 scale model vehicles of 2 metres long, 3 full truck size vehicles of 5-6 metres long, and loading and unloading docks. The simulation based control system enabled us to test the control strategies under different circumstances.
Bio Sketch:
Alexander Verbraeck has an MSc in mathematics (cum laude, 1987) and a PhD in computer science (1991) from Delft University of Technology in the Netherlands. He worked as assistant professor in information systems until 1995, when he was appointed associate professor in the systems engineering group of the faculty of Technology, Policy and Management (TPM) of TU Delft. Current research focuses on complex distributed systems such as supply chains, real-time control and emulation of equipment using simulation, and on the development of generic, object oriented simulation libraries. Alexander has been the chair of the European Board of the Society for Computer Simulation International for five years and he is a member of ACM, IEEE-CS, INFORMS, AIS, and EuroSim. His research results have been published and presented at many international conferences. In 2001, Alexander Verbraeck was the department chair of the Information, Communication and Systems department and he served on the Faculty Board. From January 2002, Alexander also has an appointment as part-time research professor in the R.H. Smith School of Business of the University of Maryland in the Logistics, Business, and Public Policy Department.

(#123) Date: April 11, 2002: Alessandro Cimatti, IRST (Institute for Scientific and Technological Research) in Trento, Italy
"Formal Methods at IRST: Symbolic Model Checking and its Applications"

Goal of this talk is to overview the research lines and the technology transfer projects of the Formal Methods group at IRST, Trento, Italy. These activities are based on the use of symbolic model checking, a formal verification technique that allows for the analysis of digital systems represented as Finite State Machines. In the talk, I will first describe NuSMV, a software platform for symbolic model checking developed at IRST as an OpenSource project. NuSMV combines the traditional use of Binary Decision Diagrams and the recently introduced Bounded Model Checking technology, based on propositional satisfiability. Then, I will show how symbolic model checking techniques are being extended and applied to the problems of Planning in Nondeterministic Domains, Safety Analysis, and Multi-agent reasoning.
Bio Sketch:
Alessandro Cimatti is a senior research at IRST (Institute for Scientific and Technological Research) in Trento, Italy, where he is the head of the Formal Methods group. He participated in and led several industrial technology transfer projects aiming at the effective integration of Formal Methods within the development process of safety-critical systems and commercial

controllers. His research interests include Symbolic Model Checking and its applications to Planning in Nondeterministic Domains and to Multi-Agent reasoning.

(#122) Date: April 4, 2002: Paul Gazis, San Jose State University Foundation
"Artificial Intelligence Techniques for Large-Scale Surveys of Space Science Data"

Many problems in space physics require large-scale surveys of extensive data sets to identify and classify qualitative features such as shocks, discontinuities, energetic particle enhancements, or specific types of spectra. Such surveys can be difficult to accomplish using conventional programming techniques and the manpower requirements associated with direct physical examination of the relevant data sets can be prohibitive. Artificial Intelligence (AI) techniques represent a potential solution to this problem. We have systematically applied and evaluated mature AI approaches, such as expert systems and several representative neural network architectures, to evaluate their suitability to perform large-scale surveys of solar wind plasma and interplanetary magnetic field (IMF) data. Several techniques were found to be particularly promising. Our results suggest that it should be possible to use AI techniques to identify and classify a broad range of different types of phenomena in large spectral data sets and extended time series that would be difficult or impossible to examine using any other means.
Bio Sketch:
Dr. Gazis is a research associate at the San Jose State University Foundation. His thesis, from the Massachusetts Institute of Technology, was on solar wind evolution. Dr. Gazis worked at MIT Lincoln Laboratory to apply AI techniques to analyze images from infra-red laser radars and at the NASA Ames Space Science Division using conventional techniques to analyze solar wind data from the Pioneer and Voyager spacecraft. His interests involve the application of AI and statistical methods to the analysis of large space science data sets.

(#121) Date: April 3, 2002: Dr. Atul G. Kelkar, Iowa State University
"Neural Networks for Modeling and Control of Dynamic Systems"

This talk is aimed at presenting some analytical, simulation, and experimental results on control of linear and nonlinear dynamic systems using neural network-based model predictive control. In particular, a class of model predictive controllers, namely, Generalized Predictive Controller (GPC) is used in this work. A brief background of neural networks with particular emphasis on their generalization capability that is central to their use in modeling and control of dynamic systems is given. Results related to the approximation accuracy of neural network and corresponding error bounds are discussed. Basic formulation of GPC control law for linear as well as nonlinear system is given. For linear unconstrained systems, conditions on GPC control parameters for ensuring closed-loop stability are given. A finite horizon stable GPC control strategy based on infinite horizon LQ performance function is proposed. For linear plants with uncertainties robustness conditions are derived for GPC control law. For control of uncertain linear systems and nonlinear systems, a neural network-based GPC architecture is presented. For uncertain linear systems, neural network is used in on-line learning configuration to account for plant uncertainties. For nonlinear systems affine in control, the nonlinear prediction equations are obtained for neural network-based predictor and closed-form of GPC control law is derived. The conditions for weight update that will ensure convergence of weights to optimal weights are given. Several experimental results are given to demonstrate the effectiveness of neural network-based GPC control schemes. Finally, open research issues in the area of NGPC are discussed and future research plan is presented.

(#120) Mar. 28, 2002: Ewen Denney, University of Edinburgh
"Towards Component Retrieval in Constructive Logic"

The gazing technique was proposed in 1992 as a solution to the problem of lemma use in automated theorem proving. The idea is to consider a hierarchy of abstraction spaces, prove a goal by rewriting in an abstract space, then gradually refine this plan in concrete spaces, patching where necessary. In this talk, I discuss work in progress which aims to extend and apply this technique, in a constructive higher-order setting, to the problem of component retrieval. This suggests an interesting framework for combining deductive-based retrieval and adaptation. This is being carried out in the context of a project which formalizes part of the Java Card virtual machine, and we will describe some aspects of this.
Bio Sketch:
Dr Ewen Denney is a research fellow in the Division of Informatics at the University of Edinburgh. His thesis, from Edinburgh's Laboratory for the Foundations of Computing Science, was on a theoretical framework for formal program development. He has since worked on formal methods and theorem proving projects, in France and Hong Kong. His interests are in semantics, program synthesis, theorem proving, and their tool support and mathematical foundations.

(#119) Mar. 21, 2002: Dr. Walt Brooks, RIACS-ARC
"Earth Science Technology Office (ESTO) Status and Plans"

ESTO manages the development of advanced technologies and applications that are needed for cost-effective Earth Science missions. ESTO plays a major role in shaping ESE research and application programs of the future. The office is organized into three major areas including Information Systems. The objectives of the Advanced Information Systems Program area is to identify, develop and (where appropriate) demonstrate advanced information system technologies. These demonstrations should: reduce the risk, cost, size, and development time of Earth Science Enterprise (ESE) space-based and ground-based information systems, increase the accessibility and utility of Earth science data, and enable new Earth observation measurements and information products. This talk will address the overall approach, objectives, upcoming initiatives and solicitations relevant to Ames Information System Research and summarize the results of a recent workshop designed to identify technology drivers.
Bio Sketch:
Dr. Walt Brooks is assigned to the level II NASA HQ Earth Science Technology Office at GSFC. His work focuses on Information Technology in support of future Earth Science Enterprise Missions. Previous work has included management of the Ames Supercomputer Facilities(NAS) and Program /Project management of several IR Space Astronomy missions.

(#118) Mar. 21, 2002: Dr. David McAllester
"The Role of Simplicity in Learning Theory"

Science values simplicity. All other things being equal, a simple explanation is preferable to a complex one. Bayesians assign higher prior probability to simple theories. But is it really true that a simple theory is a-priori more likely than a complex one? It turns out that one can justify a preference for simplicity independent of Bayesian assumptions. The justification involves only the law of large number and the observation that the number of simple theories is limited. This talk will present this justification and go on to describe more general "laws of large numbers" that justify more sophisticated methods of evaluating the accuracy of predictive rules. Applications to statistical natural language modeling will be discussed.
Bio Sketch:
David McAllester received his Ph.D. from the Massachusetts Institute of Technology AI Laboratory in 1987. He was on the faculty of the Cornell Computer Science Department for the 87/88 academic year and served on the faculty of the MIT AI Laboratory from 1988 to 1995. Since 1995 he has been a principal member of the technical staff at AT&T Laboratories-Research. David McAllester's research interests include knowledge representation, automated reasoning (automated theorem proving and formal verification), static analysis of computer programs, computer game playing (computer chess algorithms), constraint satisfaction algorithms, Bayesian networks, reinforcement learning, general PAC bounds in the theory of machine learning, and natural language modeling.

(#117) Mar. 7, 2002: Dr. David Bell, Xerox PARC
"Collaborative Systems for the Distributed Management of Knowledge in Engineering Contexts"

There are few organizing structures more ubiquitous in modern organizations than standardized procedures and processes. For example, in field service engineering contexts, standard procedures are used in an attempt to optimize machine diagnosis and repair with respect to factors such as parts and labor costs; and in research & development (R&D) contexts, standard processes are used in an attempt to optimize coordinated work with respect to factors such as technology and market risk, product quality and cost, and program schedule. The Internet is enabling new forms of collaborative systems that address these same organizational goals, but do so through more distributed means. Eureka and Sparrow Web™ are two collaborative systems developed at Xerox PARC that take advantage of the social nature of the Internet to support distributed management of knowledge in engineering contexts. Eureka has been used to support the work practices of field service technicians, and Sparrow Web™ has been used to support the work practices of R&D scientist and engineers. Both systems were developed using a human-centered and participatory design methodology that involved iteration between understanding actual work practices and co-designing and implementing collaborative systems. This talk will contrast the traditionally centralized approaches for managing knowledge using standardized procedures and processes, with the more distributed approaches enabled through Eureka and Sparrow Web.
Bio Sketch:
David Bell has been a member of the research staff at Xerox, where he was a core member of the research teams that invented the initial Eureka system and the latest version of the Sparrow Web system. During his ten years at PARC he worked in the Scientific & Engineering Reasoning Area of the Systems & Practices Laboratory, and conducted research in the areas of expert and case-based systems, knowledge and information management, and participatory human-centered

design. David recently led a national research program on enterprise learning and knowledge sharing in the National Science Foundation sponsored Center for Innovation in Product Development at MIT. David received his Ph.D. from Cornell University, with a dissertation on "Product Development Process Dynamics."

(#116) Mar. 7, 2002: Maarten Sierhuis, RIACS/NASA-Ames
"Brahms: An overview of the tool and current research"

In this talk I present an overview of the current state of the Brahms multiagent modeling and simulation environment, and will discuss how we are using Brahms in our NASA research. Since my last Brahms presentation at ARC, the Brahms virtual machine has been fully integrated with the Java language, which now allows us to use Brahms not only as a simulation environment, but also as an intelligent agent development environment. I will discuss these new capabilities and will also discuss our progress with integrating Brahms with the KAoS agent framework from IHMC/UWF and Boeing Corp., allowing Brahms to run agents distributed in a networked environment. I will also present the first release of our new Brahms Interactive Development Environment, as well as our Brahms Virtual Reality Environment effort with DigitalSpace, Inc., integrating Brahms with DigitalSpace's OWorld and web-based virtual 3D worlds using AdobeR AtmosphereT.
Bio Sketch:
Maarten is a Senior Research Scientist at RIACS/USRA, NASA Ames Research Center, Moffett Field, CA. He leads the Brahms Project in the Work Systems Design & Evaluation group (headed by Dr. William J. Clancey). Before coming to RIACS in 1998, Maarten was a member of technical staff, first in the Expert Systems Laboratory and later in the Work Systems Design group, at NYNEX Science & Technology. Maarten has a Ph.D. in Social Science and Informatics from the University of Amsterdam in The Netherlands, and an Engineering degree in Informatics from the Polytechnic University of The Hague, The Netherlands. His research interests lie in understanding the essence of human-centered computing, a new and upcoming multi-disciplinary field that could provide a leap in how information systems are designed and implemented. His view on this new field is influenced by his current research on modeling and simulating work practices in human organizations, as well as his past industry experience in developing knowledge-based systems attempting to make human organizations more effective and efficient.

(#115) Feb. 27, 2002: Chad Carson, Digital Integrity
"Blobworld: Region-based image retrieval"

I will describe "Blobworld," a content-based image retrieval system created as part of the NSF/DARPA/NASA-funded UC Berkeley Digital Library Project. This research lives at the intersection of statistical machine learning, image analysis, and information retrieval. Stock photo database users generally want to find objects in images, but most previous systems retrieved images based only on low-level features such as global color and texture histograms. The Blobworld system uses a new approach to image retrieval that approaches object-level queries. Blobworld is based on finding coherent image regions which roughly correspond to objects. Each image is segmented into regions by fitting a mixture of Gaussians model to the pixel distribution in a joint color-texture-position feature space. Each region ("blob") is then

associated with color, texture, and shape descriptors. Querying is based on the user selecting one or two regions of interest and specifying the importance of each feature type for each region. The query system is online at http://elib.cs.berkeley.edu/photos/blobworld/. Experiments indicate that queries for distinctive objects have much higher precision using Blobworld than using global image features. Blobworld querying is also more intuitive than global-feature querying because it allows the user to interact with the internal representation of the image; this helps the user formulate effective queries and understand their results. This is joint work with Serge Belongie, Jitendra Malik, Megan Thomas, Joe Hellerstein, Ray Larson, Ginger Ogle, and Joyce Gross.
Bio Sketch:
Chad Carson received a Ph.D. in Electrical Engineering and Computer Sciences from the University of California at Berkeley in 1999. His Ph.D. research, part of the Berkeley Digital Library Project, included work in image and information retrieval, computer vision, and statistical machine learning. After Berkeley, Dr. Carson joined Digital Integrity, a company that provided enterprise software and Internet services for full-content search in terabyte-scale text collections. He served as an engineering manager and product manager, leading the team developing end-user applications built on Digital Integrity's core technology. Dr. Carson received bachelor's degrees in Electrical Engineering and History from Rice University in 1994.

(#114) Feb. 21, 2002: Stephanie Seneff, Spoken Language Systems Group at the Laboratory for Computer Science, MIT
"Dialogue Design Strategies in Spoken Conversational Systems"

The Spoken Language Systems group at MIT's Laboratory for Computer Science has been developing mixed-initiative dialogue systems for spoken access to information for over a decade. This talk will focus on two of our most recently developed systems, the Mercury flight reservation system and the Orion system for task delegation. The latter system is able to call a user back at a prescribed time with pertinent information such as a traffic report or a flight arrival status. The emphasis of the talk will be on aspects concerned with dialogue design, including control, user feedback, confirmation, and error recovery. The overall system framework, which makes use of the Galaxy Communicator architecture, will also be discussed. We hope to be able to give a live demonstration of one or both of the systems.

(#113) Feb. 21, 2002: Dr. Vipul Kashyap, Technical Director of the ELBook project, Stanford Medical Informatics
"Enabling the Semantic Web: The role of metadata, semantics and domain specific ontologies"

We propose a vision of the Semantic Web, where domain specific ontologies and vocabularies will be the core infrastructure, as well as the basis of new metaphor for people and applications to interact with the web. We first illustrate with examples of how ontological concepts can be mapped to heterogeneous data, such as structured databases and textual databases. We then discuss the critical problem of inter-ontology interoperation, the core technology required to make the vision above a reality. Interoperation across ontologies requires queries to be re-written using inter-ontology relationships leading to translations that may not be semantics preserving. We present a novel approach for estimating loss of information due to the change in semantics. Measures for loss of information are defined based on intentional information; as well as on well

established metrics like "precision" and "recall" based on extensional information. These are then used to select results from multiple translations across multiple ontologies. In this talk, we establish the critical role of metadata and domain specific ontologies in developing the semantic web infrastructure.
Bio Sketch:
Vipul Kashyap works in the areas of Information and Knowledge Integration and Management, E-commerce and Semantic Web technologies. He has been active in the Semantic Web research community, and has organized panels and workshops on related topics. He has worked at R&D Labs of MCC and Telcordia Technologies (formerly known as Bellcore) on issues related to Information Integration and Agent based infrastructures and is the recipient of a Ph.D. from Rutgers University. His research interests are: interoperation across multiple domain ontologies on the Semantic Web and issues of loss of information as a consequence of ontology mismatches. Currently, he is investigating the feasibility of using sociological approaches for creating and evolving knowledge on the Semantic Web. Vipul has recently published a book on Information Brokering, has participated in panels, has been a member of conference program committees. He has published around 40 research articles and papers at various conferences and prestigious journals.

(#112) Feb. 20, 2002: Dr. Dan Berrios, Technical Director of the ELBook project, Stanford Medical Informatics
"Semi-Automated Indexing Using Domain-specific Semantics for High Precision Information Retrieval"

The use of full text resources like textbooks is frequently neither straightforward nor expedient. Faced with an urgent information need, a reader often must rely on manual inspection of a table of contents or alphabetized keyword index to guide her search. Indexes that would allow scientists to retrieve the information they need from text sources more rapidly and with greater precision must contain more knowledge than merely the location of the beginning of textbook sections or the numbers of pages on which one or two concepts are discussed. Entries in these indexes must mirror the questions that drive readers to use the text source to seek knowledge. Furthermore, these indexes must point the reader to more specific locations in the text. We have developed a system, ISAID (Internet-based Semi-automated Indexing of Documents), to generate electronic indexes for HTML documents that are more detailed and more useful to readers. ISAID is part of ELBook, an integrated system for high-precision information retrieval from full-text resources. Users of ISAID see indexes proposed by the system, based on natural language processing of documents using domain-specific semantics from the Unified Medical Language System. In this seminar, I will discuss the design and implementation of ISAID and ELBook, including a controlled evaluation of ISAID's methods, in which users were timed and the indexes they generated compared.
Biosketch:
Dan Berrios is Technical Director of the ELBook project at Stanford Medical Informatics. His research interests include collaborative information management, digital libraries and publishing, and natural language processing and web ontologies for information indexing and retrieval. He received a Bachelors degree in mathematics and biochemistry from Brown University in 1985, an M.D. from the University of California, San Francisco and Masters of Public Health in

epidemiology and biostatistics from the University of California, Berkeley in 1990, and a Ph.D. in biomedical informatics from Stanford University in 2001.

(#111) Feb. 8, 2002: Serdar Uckun, Director of Advanced Technology, Blue Pumpkin Software
"AWARE: Interpreting and Presenting Weather Data for Aviation Decision Making"

A significant percentage of general aviation accidents and fatalities are attributable to weather. Ironically, most weather-related aviation accidents are inherently preventable by simply not launching a flight into potentially hazardous conditions. Although a variety of tools are available for disseminating raw weather data to pilots, accurately interpreting weather data in the context of a mission profile remains an art mastered only after thousands of hours of flying experience. In 1998, NASA launched a program named AWIN (Aviation Weather INformation) to develop technologies that may help reduce weather-related aviation fatalities. One of the AWIN programs, AWARE is a four-year effort jointly funded by Rockwell and NASA LaRC to interpret and present aviation weather data in order to assist pilots with aviation decision making tasks. In this talk, I will discuss the progress made in the AWARE program and expand on some of the research issues common to data-rich, knowledge-poor decision support problems.
Bio Sketch:
Serdar Uckun is the Director of Advanced Technology at Blue Pumpkin Software, an enterprise software company focusing on scheduling and optimization applications for workforce management. Prior to joining Blue Pumpkin in 2000, he was Assistant Director and Manager of the Intelligent Systems Department at Rockwell Science Center, Palo Alto, CA. He served as the Program Manager for AWARE between from its inception in 1998 until early 2000. He has an M.D. from Ege University, Izmir, Turkey, an M.S. in Biomedical Engineering from Bogazici University, Istanbul, Turkey, and a Ph.D. in Biomedical Engineering from Vanderbilt University, Nashville, TN. Between 1992 and 1994, he received postdoctoral training in Computer Science at the Knowledge Systems Laboratory at Stanford University. His research interests include decision making under uncertainty, situational awareness, and scheduling.

(#110) Feb. 7, 2002: Dr. Norman Lamarra of JPL
"MarsNet Middleware Services for Remote Exploration"

Recently, JPL's Center for Space Mission Information and Software Systems sponsored a study focused on issues and architectures for future space-based networks, focusing primarily on MarsNet. It studied mission and science challenges, and sketched an approach to addressing some of these challenges via modern IT, (processors, operating systems, software architectures and middleware services). A major recommendation of that study was that middleware services be developed to support future interplanetary networks (particularly for in-situ spacecraft), and that these could be deployed incrementally over several missions to the mutual benefit of all such missions. Early work has begun with a prototype messaging middleware constructed with three "layers": application, middleware, and space protocol. This is intended to demonstrate separability that relieves applications from dealing with vagaries of the space communication. The prototype middleware is derived from a messaging API developed by JPL for the U.S. Marines, in which dispersed assets are required to communicate information over heterogeneous (and intermittent) networks, including low-bandwidth radios. We refer to all layers between

application and data link layer as "middleware", and preach the "service" approach to developing, deploying, operating, and evolving such middleware. Such middleware is in wide commercial deployment; for example, CORBA provides a set of distributed object services (naming, security, etc.) built upon (pluggable) lower-layer transport services (e.g., TCP/IP). Java also provides similar capabilities; for example, JINI and JXTA. With this background, this "messaging" prototype represents the necessary first step toward building a set of "middleware services" that are space-deployable and are designed ab initio to be combined into progressively higher-level capabilities. The next most important step proposed is a data-management prototype, leveraging Object-Oriented Data Technology infrastructure developed by JPL's Enterprise Data Management team. It is envisioned that evolution of such services would occur gradually over several years (and multiple missions), but we believe that resultant ubiquitous and rich interplanetary network infrastructure could be enabling for future space exploration.
Bio Sketch:
Norm Lamarra is a Principal software systems engineer in JPL's Engineering & Communications Infrastructure section. He received the Ph.D. degree from UCLA in System Science in 1982. Prior degrees earned were M.Sc. in Radar Technology (1974) and B.Sc. in Mathematical Physics (1973), both from the University of Birmingham, UK. During the 25 years he has been in the U.S., Dr. Lamarra has worked in the fields of radar system analysis, real-time signal processing, adaptive antenna arrays, simulation and modeling for engineering and physiological systems, and software systems. He joined JPL in 1994.

(#109) Feb. 4, 2002: Simon Buckingham Shum, Knowledge Media Institute, Open University, UK
"ScholOnto: Towards a Tool for Distributed Scientific Discourse"

Digital libraries (DLs) are gradually becoming standard resources for researchers, but while useful, these also flood us with even more information than we already have to deal with. There remains a yawning gap in the scientist's digital toolkit: tools to track ideas and results in a field, and tools to express and analyze one's understanding of their significance. This is not surprising in that we're talking about meaning rather than datasets or published information. What is the significance of this idea in relation to others? According to whom? How does the expert community perceive this theory, model, language, empirical result? Where did this idea come from? What kind of evidence supports it, and challenges it? Are there different camps on this issue? Scientific research in contested/poorly understood domains thus requires 'sensemaking' tools. Freeform annotation and discussions are the dominant solutions at present, but these generally have little structure and low status in scientific publishing, and are consequently perceived as very informal media. The Scholarly Ontologies (ScholOnto) project [http://kmi.open.ac.uk/projects/scholonto], funded by the UK EPSRC, is exploring an alternative scenario. We are developing an ontology-based 'Claims Server' to support scholarly interpretation and discourse, investigating the practicality of publishing not only documents, but associated conceptual structures in a collective knowledge base. The system enables researchers to make claims: to describe and debate, in a network-centric way, their view of a document's contributions and relationship to the literature. It thus provides an interpretational layer above raw DLs (books; papers; datasets; software tools...). This contrasts with most DL/semantic web applications that require consensus on the structure of a domain, and an agreed metadata scheme

that tries to iron out inconsistency, ambiguity and incompleteness. ScholOnto is all about supporting principled disagreement, conflicting perspectives, and the resulting ambiguities and inconsistencies, because they are the very stuff of research, and the objects of explicit inquiry. In this presentation I'll describe where we've got to (1 year into a 3 year project), and demo the current version. I'll also illustrate how we're using another sensemaking tool, Compendium, to support ScholOnto's design (Compendium is also in use at ARC, by Maarten Sierhuis). Your insights and suggestions are very welcome.
Bio Sketch
Simon Buckingham Shum is a Senior Lecturer at the Open University's Knowledge Media Institute, a 70 strong R&D lab focused on the interaction between technology and knowledge. His background is in cognitive psychology, ergonomics and human-computer interaction. His central interest is in the human dimensions to knowledge media, with specific application to scholarly publishing and knowledge management. http://kmi.open.ac.uk/sbs

(#108) Jan. 15, 2002: Cathy Marshall, Microsoft Corporation
"Reading in the Digital Library: A tale of e-books in action"

How will we read digital library materials? The answer to this question is far from straightforward; rather it has proven to be a matter that has polarized today's readers, writers, publishers, and librarians. Will we print digital documents as we need them, or will we be lured to new computer-based reading technologies like e-books? And if we are to read on a computer, what then? What might make it worthwhile to move away from a medium as convenient and malleable as paper? In this talk, I will take a work practice and technology in action-based look at reading, annotating, and collaborating over digital library materials to answer these questions.
Bio Sketch:
Cathy Marshall is an Architect at Microsoft Corporation and an active participant in the international Hypertext, Digital Library, and WWW research communities. Her research lies in the disciplinary interstices of computer science, social science, and the arts. She was a long-time member of the research staff at Xerox PARC and an affiliate of the Center for the Study of Digital Libraries at Texas A&M University. Since 1983, Cathy has been working on computer support for intellectual work from the multiple perspectives of designer, theorist, (feral) self-styled ethnographer, and writer. Her current work investigates the use of mobile pen-based computers to support reading, research, collaboration, and annotation. Her homepage is http://www.csdl.tamu.edu/~marshall

(#107) Jan. 10, 2002: Dr. Gil Syswerda, i2 Corporation
"Genetic Algorithms and Schedule Optimization"

In this talk, I will introduce genetic algorithms, and provide some computer demonstrations to show how they work on simple function optimization problems. I will then show how this technique can be applied to combinatorial optimization problems and to scheduling problems, which can be treated as constrained combinatorial optimization problems. Lastly, I will discuss what is required to take an optimization algorithm as the one described and create a successful commercial scheduling product.
Bio Sketch

Gil Syswerda received his Bachelors and Masters degrees in Computer Science from the University of Michigan in 1987, where he studied artificial intelligence in general and genetic algorithms in particular under John Holland, their inventor. In 1988, he joined the intelligent systems group at BBN Systems and Technologies, where he worked on scheduling and machine vision systems. In 1993, he co-founded Optimax Systems to produce commercial production scheduling systems. Optimax was purchased by i2 Technologies in 1997. Mr. Syswerda is currently a strategic technology and business advisor to i2.

(#106) Jan. 9, 2002: Joseph E. Flaherty, Scientific Computation Research Center, Rensselaer Polytechnic Institute
"Adaptive and Parallel Discontinuous Galerkin Methods for Hyperbolic Systems"

The discontinuous Galerkin method (DGM) provides an appealing approach to address problems having discontinuities, such as those that arise in hyperbolic conservation laws. Originally developed for neutron transport problems, the DGM has been used to solve both ordinary and partial differential equations. The DGM may be regarded as a way of extending finite volume methods to arbitrarily high orders of accuracy. The solution space is a piecewise continuous (polynomial) function relative to a structured or unstructured mesh. As such, it can sharply capture solution discontinuities relative to the computational mesh. It maintains local conservation on an elemental basis. Regardless of order, the DGM has a simple communication pattern to elements with a common face that makes it useful for parallel computation. It can handle problems in complex geometries to high order. And, it is useful with adaptivity since inter element continuity is neither required for h-refinement (mesh refinement and coarsening) nor p-refinement (method order variation). We describe several aspect of the method including basis construction, data structures, flux evaluation, solution limiting, local time stepping, and a posteriori error estimation. We further describe a framework for controlling parallel adaptive computation. The parallel data management system can handle high-order techniques and maintain a dynamic load balance in homogeneous and heterogeneous computing environments. Results of serial and parallel computations are presented for unsteady compressible flow problems involving instabilities and other complex two- and three-dimensional phenomena.
Bio Sketch:
Joseph E. Flaherty is an Amos Eaton Professor in the Department of Computer Science and Mathematical Sciences at Rensselaer Polytechnic Institute of Troy, NY. He studied Aeronautical Engineering and received his Ph.D. in applied mechanics at the Polytechnic Institute of Brooklyn. He is currently Dean of the school of science at Rensselaer. Dr. Flaherty has researched and published extensively in the areas of computational science and numerical analysis such as adaptive finite analysis, adaptive local refinement, adaptive discontinuous Galerkin technique, grid generation and adaptive algorithms, error estimation and diffusion modeling. He also developed the degree programs at Rensselaer in Computational Science and Engineering and in Numerical Analysis.
Contact: flaherje@cs.rpi.edu

(#105) Dec. 13, 2001: Dr. Naveen Ashish, IBM Silicon Valley Laboratory
"Information Mediation: Integrating Information from Multiple Online Information Sources"

NOTE: This is an encore presentation, previously given on Oct. 23, 2001. Due to terrorist attacks, the former presentation was held off base, therefore this is an opportunity for those who missed that one to be able to attend. This talk centers around "Information Mediators", which are software systems that provided integrated query access to multiple distributed information sources such as databases or Web sources. I will begin with an introduction to such systems, and provide an overview of several interesting research issues in building such systems. I will then discuss in detail an approach to optimizing the performance of these systems by locally storing data at the mediator side. Optimizing performance is the problem that I had addressed for my doctoral dissertation. I will conclude with talking about various applications of information mediators, future work in building mediators systems and finally, an overview of commercial ventures in this area.
Bio Sketch:
Naveen Ashish holds a PhD in Computer Science from University of Southern California, Los Angeles. He is currently pursuing research in Life Sciences Development at the IBM Silicon Valley Laboratory, San Jose. He has been an assistant Professor in the Department of Computer Science at University of Georgia, Athens. His interests lie mainly in the area of information management, particularly information agents and information gathering and integration on the World Wide Web. His research background includes work and contributions in the areas of information mediators, integration of data from heterogeneous data sources and extraction and integration of data from semi-structured Web sources. He has worked in areas such as query planning and optimization for information mediators, information modeling and also semi-automatic wrapper generation for extracting data from Web sources. Note: Please see the homepage of the Information Agents Research Group, for an overview.

(#104) Dec. 3, 2001: Dr. Bob Young, SciComp Inc
"Using Knowledge to Save Work"

Knowledge-based systems can save work both for end users, such as numerical modelers, and for the developers of the systems that help the users. For example, user interface components can be generated from the same knowledge bases that drive the performance of a system. One instance of how this can work comes from the domain of numerical simulation, which plays a very important role in many industries. SciComp was founded with the belief that we could help numerical modelers to produce high quality programs with much less effort on their part. SciComp's SciFinance product is a knowledge-based system that transforms concise problem descriptions, given in financial and mathematical terms, into running C-code simulations. The system has knowledge of programming, mathematics and finance. We took advantage of this environment to generate a number of different kinds of hyper-linked documents that are important components of the users' interface to SciFinance. They include reference documents, "design history" documents and viewers that offer different views of the same data, aimed at different kinds of users: mathematically knowledgeable user vs. system developer. Most of the documents were automatically generated, and hence a major labor saving, only possible because of the knowledge-based environment.
Bio Sketch:

My main interest is Human Computer Interaction: using information technology to make complex tasks easier for the people who do them. In pursuit of this goal, I have worked on problems of interest to geologists, petrophysical engineers, mechanical designers, test equipment designers and most recently, modelers with a need for numerical simulation. These problems have required large amounts of diverse knowledge for their solution. SciComp Inc: (Austin, TX)Director of Technology, A founding employee of SciComp. SciComp products generate simulation programs from a very high-level specification language. The programs solve partial differential equations and stochastic differential equations. The technology is applicable in many fields, but has been initially targeted at finance. Among a variety of activities, I designed and implemented the Electronic Information System (EIS) for our product system. Schlumberger: Aided ATE designers in developing simple ways to organize the information about a large set of software tools to help new employees get up to speed quickly without extensive human mentoring. Led a group that developed prototypes of CAD engineering tools that significantly improved both the speed and quality of customers' mechanical design process by augmenting traditional geometric representations with constraints and engineering relationships. The work led to the successful transfer and commercial re-implementation of the patented core ideas. Developed and prototyped knowledge-based tools and supported engineering groups using the tools and commercializing prototypes. PhD Department of Computer Sciences, University of Texas at Austin Explorations in Story Understanding Using Contextually Accessed Semantic Networks. A Knowledge-Based-System for text understanding
Contact Info: Bob Young
SciComp Inc
5806 Mesa, Suite 250
Austin, TX 78731
512-451-1050 Ext.204 512-451-1622 (FAX)

(#103) Nov. 29, 2001: Dr. Alexander Repenning, President and CEO of AgentSheets® Inc. Dr. Robert B. Owen is Vice-President, Technical Development, for AgentSheets Inc. "The End-User Programmable Web"

While information available via computer networks is no longer scarce, it is completely useless to many individuals because it is not provided in an appropriate format. Language, modality, and format are a few of the causes underlying representational mismatches between information producers and information consumers. Few information producers provide multiple representations of information that accommodate users with specific requirements. Analysts cite the cost of producing and maintaining multiple representations as well as the inability to predict all the possible uses of the information produced as major reasons for this common practice. In addition, information producers lack the knowledge and technology they need to deliver information to all users in a form that can be utilized. Unfortunately, with the current Web browser model, information consumers have little representational control over the HTML and XML content created by information producers. A number of XML-based efforts in research (e.g., W3's SOAP) and commerce (e.g. Microsoft's .NET) explore more flexible schemes to computationally map information spaces from information producers to information consumers. However, these efforts are focused on software developers, not computer end-users. We explore the use of Every-Citizen Interfaces that make diverse use of Web-based information by giving sufficient control to individual information consumers through end-user programming. The goal is to enable them, without the need of professional programming skills, to computationally process information and to interact with information through different modalities, including spoken dialogue interaction in their native language with animated computer agents. AgentSheets is an authoring tool to build multi-agent simulations and games that can access and transform Web-based information. A number of applications will be presented including distributed eco world simulations built by elementary school children, voice controlled mountain bike advisor agents, and space shuttle payload simulations built by NASA scientists. More info: http://www.agentsheets.com/showcase.html

Bio Sketch:

Repenning:

Dr. Alexander Repenning is the CEO and President of AgentSheets Inc., a Computer Science professor and member of the Center for Lifelong Learning & Design at the University of Colorado in Boulder. Repenning's mission is to build cognitive tools, that he calls Thought Amplifiers, helping people to explore and communicate complex ideas interactively. He has worked in research and development at Apple Computer, Asea Brown Boveri, Hewlett Packard, Martin Marietta, and Xerox PARC. Repenning is the creator of the award-winning AgentSheets simulation and game-authoring tool. His research interests include end-user programmable agents, computers in education, and artificial intelligence. Repenning has served as advisor to the National Academy of Sciences, the National Science Foundation and the European Commission. Repenning received his Ph.D. in computer science from the University of Colorado in 1993. He can be reached at alexander@agentsheets.com

Owen:

Dr. Robert B. Owen is Vice-President, Technical Development, for AgentSheets Inc. As one of the Founders of AgentSheets Inc., he serves on the AgentSheets Board of Directors in addition to his Vice-President T.D. duties. Dr. Owen received his Ph.D. in physics from Virginia Polytechnic Institute and State University and his Ph.D. in anthropology from the University of Colorado, with an emphasis on computer science methods. He served as Research Scientist at NASA Marshall Space Flight Center from 1966 until 1987. Dr. Owen spent over a decade designing, building, and flying aircraft laser interferometer systems for NASA. He has 25 years of experience in holographic optical systems, and was among the first to use holography and interferometry in low-gravity environments. While a NASA employee, he designed and constructed interferometers that he operated on over 8000 low-gravity parabolas in the NASA KC-135 low-gravity simulation aircraft. He has been directly responsible for optically characterizing the microgravity behavior of a wide range of materials science and biological samples. He developed an in-line digital holographic system for analysis of marine particulates, and a digital holographic interferometer for Space Station materials science experiments. Dr. Owen was Senior consultant for optics on the CAST Spacelab experiment. He developed a neural network-based adaptive optics system that automatically maintained alignment in microgravity and on the KC-135 aircraft. He has converted many laboratory systems into practical optical instruments. Dr. Owen uses neural networks and other AI (artificial intelligence) methods to develop smart instruments. He has worked in optical computing and adaptive optics. Dr. Owen is author or co-author on 65 scientific publications, holds 4 patents, and is the recipient of scores of honors. He was a director of the Colorado R&D Consortium, and is a graduate-trained government contract manager. He presently serves on the Board of Directors of the

Colorado Photonics Industry Association. He is a member of the Optical Society of America and SPIE-The International Society for Optical Engineering.

Dr. Alexander Repenning, President and CEO
AgentSheets® Inc.
6560 Gunpark Drive
Boulder, CO 80301

(#102) Nov. 27, 2001: Dr. William Hefley, Q-Labs and SEI
"Avoiding the Terrible, Horrible, No-Good Very Bad Day:
Ensuring at Least Three C's in Human-Centered Computing"

When things go wrong in everyday life, we sometimes think of the children's book and the "terrible, horrible, no-good very bad day." When things go wrong in mission operations of advanced or autonomous space missions, it is possible that the same conclusions about bad days could be drawn. In ensuring that human-centered computing for mission operations supports these unforeseen circumstances, the deployed computing environment must support three key factors: context, communication, and collaboration. This talk will address the importance of these factors and identify some of the challenging research problems related to these factors.
Bio Sketch:
Dr. Bill Hefley is currently a Principal Consultant with Q-Labs and a resident affiliate of the Software Engineering Institute (SEI). He has over twenty-five years experience in the information technology field. His interests in software process improvement include organizational aspects of managing and developing human resources and the integration of usability/HCI concerns. His dissertation addressed interpersonal dynamics within software teams. Dr. Hefley has been on the faculty at Carnegie Mellon University and taught for IBM's Executive Consulting Institute and the Software Engineering Institute (SEI). At the SEI, he was involved in technology transition activities and led the team that created the People Capability Maturity Model®. In prior industry roles in both government and commercial sectors, he led systems development and user interface design projects for space and C3I applications, was the project manager for the STS-39 payload operations part-task trainer, and contributed to financial and manufacturing systems for firms in the heavy manufacturing and semiconductor industries. Dr. Hefley has consulted in the United States, Denmark, Australia, Japan, and the Netherlands.

(#101) Nov. 1, 2001: Lynn Harper, Lead of Integrative Studies, Astrobiology and Space Research Directorate
"Space Station and NASA's New Biology: Think of it as Evolution in Action"

Reduced to its essentials, the justification for a space mission is to transport people and equipment to extraterrestrial locations for the purposes of discovery and development. Affordability of space missions is what constrains the scope, magnitude, distance, duration and desirability of missions. For these reasons, NASA has a standing goal to reduce the cost of a payload pound by an order of magnitude. However, affordability can also be approached from a different direction: amplify the value of a payload pound by many orders of magnitude. Three convergent revolutions - in biotech, infotech, and microminiaturization - can now meet in the very long duration engineered extraterrestrial ecology of the International Space Station. Together, these capabilities provide the means to explore the only life we know during its first generations beyond the planet of origin with unprecedented depth, sensitivity, scope, and resolution. Space provides life's second data point: habitation in an utterly alien environment. Over earth's 4.5 billion year evolution, many environmental changes shaped terran life, but changes in gravity were not among them. New space biology studies are likely to reveal features of terran life that literally cannot be seen on Earth. Early - and very limited -- application of these new technologies in space has already yielded results that were medically important and commercially interesting. The importance of the scientific opportunity is akin to using contemporary biotechnology tools to document the evolution of the first sea creatures to the land. Early studies conducted between 1998 and now show that high fidelity studies on cells and small organisms can be conducted in space and the inflight data preserved sufficiently well to allow postflight applications of contemporary biotechnology's most sophisticated and comprehensive analytical techniques. This type of approach enables orders of magnitude more data return from space biosciences than has ever been possible -- certainly far more than the original designers ever envisioned. Important investigations - even pioneering studies - can be conducted in small volumes and under highly constrained conditions. Participation opportunities and data amplification can continue via data mining and analyses of sample archives. This defines the floor of new millennium space biosciences. Concepts for these studies can be accomplished will be presented. Discussions on how to further improve the value of these investigations will be encouraged.

(#100) Oct. 29, 2001: Jim Martin from Rockwell Science Center
"Necessity is the Mother of Invention"

I will present a brief review of some of the technical projects with which I have been involved in the past, including experimental particle physics, robotics, technology access, and artificial intelligence. The common theme has been: "How do we turn scientific and technical capabilities into something that can be used, or sold as a product?". Critical factors in determining the best approach are understanding the domain of use and getting the full participation of one's partners/customers. Keeping their support requires producing usable and useful deliverables on an on-going basis. Flexibility is essential, especially since requirements evolve and mature during a project. My take-away lesson from the last two decades of R&D bears a great similarity to the old aphorism: Need is very much a primary motivator of innovation.

(#99) Oct. 23, 2001: Dr. Naveen Ashish, IBM Silicon Valley Laboratory
"Information Mediation: Integrating Information from Multiple Online Information Sources"

This talk centers around "Information Mediators", which are software systems that provided integrated query access to multiple distributed information sources such as databases or Web sources. I will begin with an introduction to such systems, and provide an overview of several interesting research issues in building such systems. I will then discuss in detail an approach to optimizing the performance of these systems by locally storing data at the mediator side. Optimizing performance is the problem that I had addressed for my doctoral dissertation. I will conclude with talking about various applications of information mediators, future work in building mediators systems and finally, an overview of commercial ventures in this area.

(#98) Oct. 23, 2001: Rachel Sussman, Tanenhaus student at University of Rochester and Stanford University
"Do actions really speak louder than words?: Eye-movements during passive listening tasks"

There is an ever-growing body of work linking eye-movements to language processing in action based tasks. When performing a task within a circumscribed visual world, subjects' eye-movements to objects in the scene reflect their moment-by-moment understanding of auditorily presented instructions. Relatively fewer studies have looked at how language drives eye-movements in more passive listening domains. That is, in the absence of any pressing, real-world task to perform, are eye-movements still dependent on and time-locked to the syntactic processing of spoken language? This talk will give a brief overview of work that examines these questions, as well as present results from a new study that examines eye-movements during interrogatives. The study presented subjects with a spoken narrative and an accompanying display of characters and objects mentioned in the story. After the narrative, subjects were required to verbally answer a "comprehension question" about the story. All of the entities relevant to the question were present in the visual display. Patterns of eye-movements were found to vary depending on the type of question asked (yes/no vs. a question involving a WH-word) as well as the preferred transitivity of the verb in the question. Since subjects were not required to interact with the display in any way, this result provides strong support that eye-movements to visual displays are not merely an upshot of motor planning, but are largely driven by the syntax of linguistic input.

(#97) Oct. 18, 2001: Pentti Kanerva, Swedish Institute of Computer Science (SICS)
"Computing with Large Random Patterns: Where AI and Neural Nets May Meet"

In traditional computing we calculate with numbers (numeric computing) and we manipulate pointers in computer memory (symbolic computing). These methods have been used in Artificial Intelligence (AI), and impressive results have been achieved in specialized areas. However, we have not been able to build systems that learn in ways resembling human learning, which is the basis of human intelligence. Artificial neural nets are more recent and have had some success with humanlike learning. At SICS we have studied the mathematical foundations of these newer methods that compute with large random patterns, rather than with numbers and pointers. Artificial neural nets are simplified models of real neural circuits. Large patterns are used because the brain's circuits are large--they have thousands to millions of neurons--and because many neurons are active at any one time. The patterns are random because no two brains are wired exactly alike and yet can learn to act alike, for example they can learn to agree about the meanings of words and phrases. This means that each brain has its own internal code based on its peculiar connections. Neural activity patterns are modeled by points of a high-dimensional space, or multicomponent (high-d) vectors where each dimension represents an element of a pattern or a neuron of a circuit. Thus the dimensionality of the vectors is in the thousands to millions. Spaces with such high dimensionality have surprising properties. The study of pattern computing then involves the study of the geometric and algebraic properties of high-dimensional spaces and of

computing based on them. Roughly speaking, geometric properties refer to distances between points of the space--i.e., to judgments about similarity--and algebraic properties refer to the making of new meaningful patterns from existing ones. Until recently neurocomputing has relied almost entirely on the geometric properties of the space, which have proven sufficient for tasks such as recognizing handwritten and spoken words. We have focused on the algebraic properties, to provide computational operations for higher mental functions such as learning by analogy and human use of language. In technical terms, the algebraic operations are needed for the encoding and decoding of compositional structure and for analogical mapping of concepts and structures. In this talk I will demonstrate these ideas with the simplest of vectors, namely, with long random bit strings, and indicate what other kinds of vectors and operations have been used.
Bio Sketch
Pentti Kanerva has been intrigued by the prospect of understanding brains in computer terms since first learning about computers long ago in Finland. His thesis was the basis of the book Sparse Distributed Memory (MIT Press, 1988). In 1985-92 he was a principal investigator of an SDM research group at RIACS and since then at SICS. Recent work has been on representing compositional structure, and making semantic vectors for words and larger texts, with large random patterns.

(#96) Oct. 11, 2001: Dr. Kurt C. Wallnau, Carnegie Mellon's Software Engineering Institute
"Packaging Prediction: Prediction-Enabled Component Technologies"

The real innovations of software component technology lie in the areas of 'packaging' and 'deployment'. By focusing on packaging and deployment, it becomes possible to consider how component technology can be used to accelerate experimentation, integration, and transition of prediction technologies. By prediction technology I mean those technologies that permit engineers to analyze and predict the emergent attributes of an assembly of components (security, performance, safety, reliability, etc.) from the known and trusted properties of its constituent components. In this presentation I will explain why progress is possible now, how logical and empirical approaches to certification and prediction can be reconciled, the nature of component interface, the role of certification and trust, and other related ideas. I will then describe a proof of feasibility prediction enabled component technology (PECT) that demonstrates some of the key ideas of predictable assembly. I will close the presentation with a discussion of open issues, and activities that we and others are pursuing to establish a community of collaborators in predictable assembly from certifiable components.
Bio Sketch:
Kurt C. Wallnau is a senior member of the technical staff at Carnegie Mellon's Software Engineering Institute. Mr. Wallnau currently leads an SEI project in Predictable Assembly from Certifiable Components (PACC). Prior to his work on PACC, Mr. Wallnau led the SEI COTS-Based Systems project. His recent book, Building Systems from Commercial Components (Addison-Wesley) describes the exploratory design processes required by use of commercial components today. Mr. Wallnau's interest in PACC lies in the hope that we can substantially improve the engineering rigor of component-based design processes.

(#95) Oct. 9, 2001: Mirelle Ducasse, INSA/IRISA, Rennes, France
"A generic approach to monitor program executions (joint work with E. Jahier)"

Monitoring requires to gather data about executions. The monitoring functionalities currently available are built on top of ad hoc instrumentations. Most of them are implemented at low-level; in any case they require an in-depth knowledge of the system to instrument. The best people to implement these instrumentations are generally the implementers of the compiler. They, however, cannot decide which data to gather. Indeed, hundreds of variants can be useful and only end-users know what they want. In this article, we propose a primitive which enables users to easily specify what to monitor. It is built on top of the tracer of the Mercury compiler. We illustrate how to use this primitive on two different kinds of monitoring. Firstly, we implement monitors that collect various kinds of statistics; each of them is well-known, the novelty is that users can get exactly the variants they need. Secondly, we define two notions of test coverage for logic programs and show how to measure coverage rates with our primitive. To our knowledge no definition of test coverage exist for logic programming so far. Each example is only a few lines of Mercury. Measurements show that the performance of the primitive on the above examples is acceptable for an execution of several millions of trace events. Our primitive, although simple, lays the foundation for a generic and powerful monitoring environment.

(#94) Oct. 4, 2001: Dr. Kanna Rajan, NASA Ames Research Center & The MER P&S prototyping team
"The MER Planning/Scheduling Prototype Effort"

Since February 01, a joint ARC/JPL team has been developing a software prototype that uses advanced Planning and Scheduling techniques to assist in scheduling of high level science goals for the MER mission. The overall objective of this effort is to ensure that scientists are able to rapidly generate a composite picture of a Sol's activities in an optimal manner consistent with inter-activity constraints and resource bounds. The system operates in a "mixed-initiative" mode that allows user preferences and intent to be fully captured in the planning process while alleviating the burden on the human user of routine tasks such as enforcing flight and mission rules. This effort combines two systems with a long heritage: JPL's APGEN Mission Planning system; and the Automated Planner/Scheduler flown on-board DS1 as part of the Remote Agent Experiment. It leverages the best of both progenitors by allowing the extensive user interaction facilities in APGEN to be wedded to the constraint maintenance features of the automated planner, providing a tool for rapid scenario analysis and high level science plan generation. Users will see the familiar APGEN interface, but will have additional optional features to automate routine aspects of the science planning process. In this talk and demonstration of the prototype software, we will show the flexibility the system offers scientists in adding powerful search capabilities to incrementally generate plans, which can then be edited, rearranged or merged with other plans. The talk will focus on the background and the process involved in getting to this stage (including the developmental methodology), and the potential impact this has on MER. The software demonstration will highlight the various features that were incorporated in the system after feedback from a number of MER personnel and project scientists. The team that developed the prototype includes John Bresina/ARC, Will Edgington/ARC, Ari Jonsson/ARC, Adans Ko/JPL, Pierre Maldague/JPL and Paul Morris/ARC.

## *III.B. RIACS-Supported Workshops*

As part of its mission of fostering ties with the academic community in IT, RIACS provides financial, administrative, and technical support for selected workshops involving RIACS scientists. The following workshops were supported during this reporting year:

**AIPS 02**
RIACS collaborated with NASA Ames to provide support for the Sixth International Conference on AI Planning & Scheduling (AIPS 02), held in Toulouse, France, on 23-27 April 2002. In particular, RIACS and NASA Ames provided student scholarships for conference attendees. The International Conference on AI Planning & Scheduling (AIPS) is a well established scientific meeting for the presentation of new results in the growing field of planning and scheduling. More information on this conference may be found at http://www.laas.fr/aips/aips2.html.

**Workshop on Collectives and the Design of Complex Systems**
Organized jointly by RIACS and NASA Ames, the Workshop on Collectives and the Design of Complex Systems took place at NASA Ames Research Center on 6-9 August 2002. With the advent of extremely cheap computing, we are moving to a world filled with distributed systems of computationally sophisticated components. Very often there are performance criteria by which each of these components can rank its own behavior, as well criteria by which we rank the dynamic behavior of the complex system as a whole. Examples of such overall performance criteria are total throughput in a data network, total scientific information gathered by a constellation of deployable instruments, and, in the natural world, GDP growth in a human economy, or percentage of available free energy exploited by an ecosystem.

No current scientific discipline provides a thorough understanding of the relation between the structure of such "collectives" and how well they meet their overall performance criteria. This workshop aims to be the first step at overcoming this problem by laying the foundation of the study of collectives. More information on this workshop may be found at http://ic.arc.nasa.gov/~kagan/cdcs2002/cdcs2002.html

**IDU Review and Workshop**
RIACS supported a review and workshop held by the Intelligent Data Understanding component of the NASA Intelligent Systems project. The review and workshop were held in Monterey, CA, on 4-6 September 2002. The Intelligent Data Understanding component's goal is to develop and integrate sophisticated techniques that assist in the data/information extraction and knowledge discovery process in onboard space missions and on the ground. More information on the project may be found at http://www.cict.nasa.gov/Public/is.php.

RIACS has also been working with various NASA and university groups in preparation for workshops to be held in the coming year.

**Workshop on Biology/Information Science/Nanotechnology Fusion, and NASA Missions**
RIACS, USRA, and NASA Ames jointly organized an Invitational Workshop on Biology/Information Science/Nanotechnology Fusion, and NASA Missions, held at NASA Ames Research Center on 7-9 October 2002. NASA missions to explore the universe, to query the origin and evolution of life, and to observe Earth generate challenges and hard problems that test limits in science and engineering. Several of these challenges and their potential solutions lie in the fields of Biology, Information Science, and Nanotechnology (BIN). These disciplines either directly or in combination are widely regarded as a source of breakthrough solutions for NASA missions.

One motivation for fusion among fields of science and engineering has been the opportunities afforded by the methods and insights from one field to solve hard problems or challenges in a second field. Exploitation of these opportunities has led to a natural merging of disciplines and the formation of new disciplines. This workshop discussed science challenges and hard problems generated by NASA missions with the goal of matching these challenges to anticipated innovative solutions from cross-disciplinary research in the fields of BIN. Further, workshop discussions motivated and supported planning for a broader fusion among the three fields, to the benefit of NASA. More information on this workshop may be found at http://binfusion.arc.nasa.gov/

Other workshops being supported and planned for the coming year include:

- Workshop on Foundations for Intelligent Physical Agents, to be held October 14-18, 2002. Collaboration with University of West Florida.
- NASA Planning and Scheduling Workshop, to be held October 27-29, 2002. Collaboration with Automation and Robotics area, Computational Sciences Division.
- Information Power Grid Workshop, to be held December 11-12, 2002. Collaboration with NASA Advanced Supercoming Division.

## *III.C. RIACS Technical Reports*

RIACS encourages its scientists and staff to make technical papers and other reports available through RIACS technical reports. RIACS technical reports are available online at http://www.riacs.edu/trs through our use of the EPrints software system. The following reports were posted on the system during FY02.

**01.25 Maximally Informative Statistics for Localization and Mapping**
*Author(s): Matthew C. Deans, Robotics Institute, Carnegie Mellon University*
*Number of Pages: 14*
*Publication: Submitted to ICRA 2002*

**Abstract:**
This paper presents an algorithm for localization and mapping for a mobile robot using monocular vision and odometry as its means of sensing. The approach uses the Variable State Dimension filtering (VSDF) framework to combine aspects of Extended Kalman filtering and nonlinear batch optimization. This paper describes two primary improvements to the VSDF. The first is to use an interpolation scheme based on Gaussian quadrature to linearize measurements rather than relying on analytic Jacobians. The second is to replace the inverse covariance matrix in the VSDF with its Cholesky factor to improve the computational complexity. Results of applying the filter to the problem of localization and mapping with omnidirectional vision are presented.

**01.26 (Almost) Featureless Stereo -- Calibration and Dense 3D Reconstruction Using Whole Image Operations**
*Author(s): V. N. Smelyanskiy, NASA Ames*
*R.D. Morris, D.A. Maluf, P. Cheeseman, RIACS*
*Number of pages: 8*

**Abstract:**
The conventional approach to shape from stereo is via feature extraction and correspondences. This results in estimates of the camera parameters and a typically sparse estimate of the surface. Given a set of calibrated images, a dense surface reconstruction is possible by minimizing the error between the observed image and the image rendered from the estimated surface with respect to the surface model parameters. Given an uncalibrated image and an estimated surface, the camera parameters can be estimated by minimizing the error between the observed and rendered images as a function of the camera parameters. We use a very small set of matched features to provide camera parameter estimates for the initial dense surface estimate. We then re-estimate the camera parameters as described above, and then re-estimate the surface. This process is iterated. Whilst it can not be proven to converge, we have found that around three iterations results in excellent surface and camera parameter estimates.

### 01.27 Robust Automatic Feature Detection and Matching Between Multiple Images

*Author(s):* *Maurice Ringer, Cambridge University Engineering Department*
*Robin D. Morris, RIACS*
*Number of pages: 25*

**Abstract:**
This report deals with the problem of identifying common points (features) observed in multiple images. It describes various currently popular solutions to this problem and a complete system for automatically performing detection and matching of these features based on these current techniques. The system is shown well on many image sequences, although, as the proposed system is a concatenation of many existing techniques, it also exhibits failure under similar conditions to these techniques, namely when images are taken from extremely different angles.

### 01.28 Designing Adaptive Low-Dissipative High Order Schemes for Long-time Integrations

*Author(s):* *H.C. Yee, NASA Ames Research Center*
*B. Sjogreen, Department of Numerical Analysis and Computer Sciences, KTH*
*Number of Pages: 60*
*Publication:* *A Chapter for a book entitled ``Turbulent Flow Computation," (Eds. D. Drikakis \& B. Geurts), Kluwer Academic Publisher, 2001*

**Abstract:**
A general framework for the design of adaptive low-dissipative high order schemes is presented. It encompasses a rather complete treatment of the numerical approach based on four integrated design criteria:

(1) For stability considerations, condition the governing equations before the application of the appropriate numerical scheme whenever it is possible.

(2) For consistency, compatible schemes that possess stability properties, including physical and numerical boundary condition treatments, similar to those of the discrete analogue of the continuum are preferred.

(3) For the minimization of numerical dissipation contamination, efficient and adaptive numerical dissipation control to further improve nonlinear stability and accuracy should be used.

(4) For practical considerations, the numerical approach should be efficient and applicable to general geometries, and an efficient and reliable dynamic grid adaptation should be used if necessary.

These design criteria are, in general, very useful to a wide spectrum of flow simulations. However, the demand on the overall numerical approach for nonlinear stability and accuracy is much more stringent for long-time integration of complex multi-scale viscous shock/shear/turbulence/acoustics interactions and numerical combustion. Robust classical numerical methods for less complex flow physics are not suitable or practical for such applications. The present approach is designed expressly to address such flow problems, especially unsteady flows. The minimization of employing very fine grids to overcome the production of spurious numerical solutions and/or instability due to under-resolved grids is also sought. The incremental studies to illustrate the performance of the approach are summarized.

Extensive testing and full implementation of the approach is forthcoming. The results shown so far are very encouraging.

### 01.29 Efficient translation of LTL formulae into Büchi automata

*Author(s):* *Dimitra Giannakopoulou, RIACS*
*Flavio Lerda, School of Computer Science, Carnegie Mellon University*
*Number of pages: 17*

**Abstract**

Model checking is a fully automated technique for checking that a system satisfies a set of required properties. With explicit-state model checkers, properties are typically defined in linear-time temporal logic (LTL), and are translated into Büchi automata in order to be checked. This report presents how we have combined and improved existing techniques to obtain an efficient LTL to Büchi automata translator. In particular, we optimize the core of existing tableau-based approaches to generate significantly smaller automata. Our approach has been implemented and is being released as part of the Java PathFinder software (JPF), an explicit state model checker under development at the NASA Ames Research Center.

### 02.01 Source-Code Instrumentation and Quantification of Events

*Author(s):* *Robert E. Filman, RIACS*
*Klaus Havelund, Kestrel Technology*
*Number of pages: 5*
*Presentation:* *Workshop on Foundations Of Aspect-Oriented Languages (FOAL), Aspect-Oriented Software Development 2002*

**Abstract:**

Aspect-Oriented Programming is making quantified programmatic assertions over programs that otherwise are not annotated to receive these assertions. Varieties of AOP systems are characterized by which quantified assertions they allow, what they permit in the actions of the assertions (including how the actions interact with the base code), and what mechanisms they use to achieve the overall effect. Here, we argue that all quantification is over dynamic events, and describe our preliminary work in developing a system that maps dynamic events to transformations over source code. We discuss possible applications of this system, particularly with respect to debugging concurrent systems.

### 02.02 Polychotomic Encoding: A Better Quasi-Optimal Bit-Vector Encoding of Tree Hierarchies

*Author(s):* *Robert E. Filman, RIACS*
*Number of pages: 18*
*Conference:* *16th European Conference on Object-Oriented Programming*

**Abstract:**

Polychotomic Encoding is an algorithm for producing bit vector encodings of trees. Polychotomic Encoding is an extension of the Dichotomic Encoding algorithm of Raynaud and Thierry. Polychotomic and Dichotomic Encodings are both examples of hierarchical encoding

algorithms, where each node in the tree is given a gene---a subset of {1,...,n}. The encoding of each node is then the union of that node's gene with the genes of its ancestors. Reachability in the tree can then be determined by subset testing on the encodings.
Dichotomic Encoding restructures the given tree into a binary tree, and then assigns two bit, incompatible (chotomic) ``genes" to each of the two children of a node. Polychotomic Encoding substitutes a multibit encoding for the children of a node when the restructuring operation of Dichotomic Encoding would produce a new heaviest child (child requiring the most bits to represent a tree of its children) for that node. The paper includes a proof that Polychotomic Encoding never produces an encoding using more bits than Dichotomic Encoding. Experimentally, Polychotomic Encoding produces a space savings of up to 15% on examples of naturally occurring hierarchies, and 25% on trees in the randomly generated test set.

### 02.03 Synthesizing Certified Code

*Author(s):* *Michael Whalen, Univ. Minnesota*
*Johann Schumann, RIACS*
*Bernd Fischer, RIACS*
*Number of pages: 12*

**Abstract:**
Code certification is a lightweight approach to demonstrate software quality on a formal level. Its basic idea is to require code producers to provide formal \emph{proofs} that their code satisfies certain quality properties. These proofs serve as \emph{certificates} which can be checked independently. Since code certification uses the same underlying technology as program verification, it also requires many detailed annotations (e.g., loop invariants) to make the proofs possible. However, manually adding these annotations to the code is time-consuming and error-prone.
We address this problem by combining code certification with automatic program synthesis. We propose an approach to generate simultaneously, from a high-level specification, code and \emph{all} annotations required to certify the generated code. Here, we describe a certification extension of {\sc Au\-to\-Bayes}, a synthesis tool which automatically generates complex data analysis programs from compact specifications. \AB\ contains sufficient high-level domain knowledge to generate detailed annotations. This allows us to use a general-purpose verification condition generator to produce a set of proof obligations in first-order logic. The obligations are then discharged using the automated theorem prover E-\SETHEO. We demonstrate our approach by certifying operator safety and memory safety for a generated iterative data classification program without manual annotation of the code.

### 02.04 Injectors and Annotations

*Author:* *Robert E. Filman, RIACS*

*Number of pages: 6*

*Conference:* *Workshop on Concrete Communication Abstractions Of The Next 701 Distributed Object Systems, ECOOP'2002, Málaga, Spain, June, 2002.*

**Abstract:**

We argue for argue for extending the conventional distributed objects model in two important ways: (1) Having the ability to insert injectors (filters, wrappers) into the communication path between components, and (2) Having the ability to annotate communications with additional information, and to propagate these annotations through an application.

### 02.05 Realizing Aspects by Transforming for Events

*Author(s):* *Robert E. Filman, RIACS*
*Klaus Havelund, Kestrel Technology*

*Number of pages: 7*

*Conference:* *Workshop on Declarative Meta Programming at Automated Software Engineering 2002, Edinburgh, Scotland, September 2002*

**Abstract:**

We explore the extent to which concerns can be separated in programs by program transformation with respect to the events required by these concerns. We describe our early work on developing a system to perform event-driven transformation and discuss possible applications of this approach.

### 02.06 A Bibliography of Aspect-Oriented Programming

*Author(s):* *Robert E. Filman, RIACS*

*Number of pages: 29*

**Abstract:**

A bibliography of the literature related to Aspect-Oriented Programming.

### 02.07 Technology Requirements for Information Management

*Author(s):* *Sara Graves, UAH*
*Craig A. Knoblock, Univ. of Southern California*
*Larry Lannom, CNRI*

*Number of pages: 32*

**Abstract:**

This report provides the results of a panel study conducted into the technology requirements for information management in support of application domains of particular government interest, including digital libraries, mission operations, and scientific research.

The panel concluded that it was desirable to have a coordinated program of R&D that pursues a science of information management focused on an environment typified by applications of

government interest – highly distributed with very large amounts of data and a high degree of heterogeneity of sources, data, and users.

# IV. RIACS Staff

The primary mechanism used at RIACS is the engagement of RIACS scientists in NASA projects. RIACS staff scientists are recruited from the broad university research community to provide a nucleus of activity to both collaborate with NASA scientists and to provide an "attraction point" for bringing in visitors from academia. These scientists typically work in clusters on NASA projects – collaborating closely with NASA scientists and onsite contractors, providing senior technical expertise and coordinating visitors from academia.

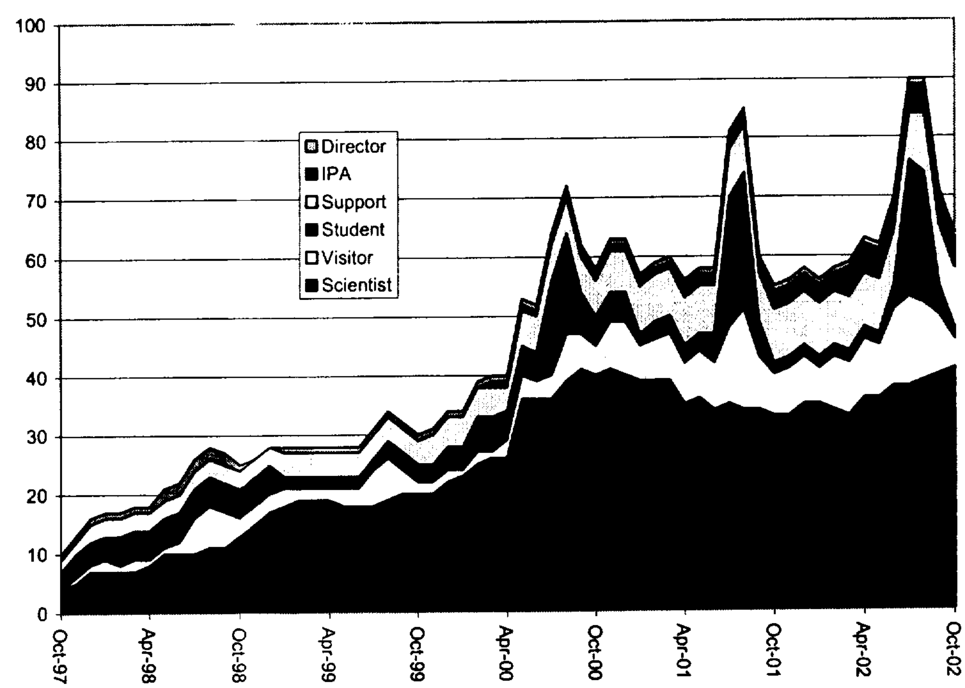


**Figure 4: RIACS Staff**

Figure 4 shows the growth in RIACS staff during the current cooperative agreement. Over the period shown, the number of scientific staff onsite has grown from three to a current 41 permanent scientific staff members. In addition, five scientists (Daniel Clancy, James Hieronymus, Butler Hine, David Maluf, and Kanna Rajan) are now on loan to NASA under the IPA program. Including panelists for the Information Management Technology Requirements Study and participants in the Summer Student Research Program, 52 visiting scientists and 44 visiting students spent time at RIACS during the year.

The following sections lists both the permanent and visiting staff at RIACS over the reporting period, with the time they have been at RIACS shown in parentheses.

## *IV.A. Management Staff*

**Barry Leiner**, Director – Ph.D. Electrical Engineering, Stanford University, 1973. (06/28/99 – present)

**Serdar Uckun**, Deputy Director – Ph.D, Biomedical Engineering, Vanderbilt University, 1992. (3/22/02 – present)

**Kathleen M. Connell**, Associate Director for Bio/IT Fusion – BA, Human and Organizational Development, UC Berkeley. (4/1/01 – present)

**Johann Schumann**, Associate Director for Software Engineering – Ph.D, 1991, Computer Science, Technical University of Munich, Germany. (03/03/00 – present)

**Richard Washington**, Associate Director for Autonomous Systems – Ph.D., 1994, Computer Science, Stanford University. (04/03/00 – present)

**Diana Martinez**, Administrator – BS, 2002, Business Administration, University of Phoenix (08/18/97 - present)

## *IV.B. Administrative and Support Staff*

**Beatrice Burnett**, Administrative Assistant (11/3/97 - present)

**Sue Christman**, Administrative Assistant (4/06/00 - present)

**Bernadette Pagan**, Administrative Assistant (9/17/01 – present))

**Peggy Leising**, Project Facilitator (04/01/00 – present)

**Ryan Nelson**, Systems Administrator (subcontractor) (6/17/00 – present)

**Roscoe 'Jack' Jackson**, Administrative Assistant (10/10/00 – present)

**Charlene Daley**, Administrative Assistant (12/26/00 – present)

**Karen Brennan**, Administrative Assistant Intern, De Anza College (7/02/01 – 6/30/02)

**Cindy Quach**, Administrative Assistant Intern, De Anza College (7/1/02 – present)

**Manuel Nanez**, Administrative Assistant Intern, De Anza College, 7/2/02 – present)

## *IV.C. Scientific Staff*

**Janice Aikins**, Ph.D., 1980, Computer Science, Stanford University, Technology strategy and technology transition of Ames information technology projects to programs in NASA centers. (1/16/02 – 5/7/02)

**Gregory Aist**, Ph.D., 2000, Language and Information Technology, Carnegie Mellon University, Spoken dialogue systems and intelligent tutoring systems. (11/26/01 – present))

**Linda Andrews**, BS, 1990, San Jose State University, Physics graduate work, University of California, Davis, Information systems technology and program management in support of Life Science Enterprise. (6/20/01 present

**Naveen Ashish**, Ph.D., 2000, Computer Science, University of Southern California, Information mediation and information agents for data extraction and fusion from multiple sources (3/25/02 – present)

**Anupa Bajwa**, Ph.D., 1995, Aerospace Engineering, Pennsylvania State University, Model-based reasoning for autonomous systems diagnosis. (04/03/00 – present)

**Esfandiar Bandari**, Ph.D, 1995, Computer Science, University of British Columbia, Computational Vision, Signal Processing, 3-D Reconstruction and Medical Imaging (09/2/98 – present)

**David Bell**, Ph.D., 1993, Engineering Design, Cornell University, Collaborative Systems, engineering for complex systems. (5/1/02 – present)

**Daniel Berrios**, Ph.D., 2001, Media & Information Science, Stanford University, Information sharing (5/1/02 – present)

**Walter F. Brooks**, Ph.D., 1977, Physics, Stevens Institute of Technology, Advanced air traffic control, advanced information systems technology in support of Earth Science Enterprise. (04/04/00 – present)

**Peter J. Cheeseman**, Ph.D., 1979, Monash University, Artificial Intelligence, computational complexity, bayesian inference, computer vision, plasma physics (09/1/97-present)

**Daniel J. Clancy**, Ph.D., 1997, Computer Science, University of Texas, Austin. Model-based reasoning, artificial intelligence, diagnosis, health management (01/24/00 – present)

**Kathleen M. Connell**, BA, Human and Organizational Development, UC Berkeley, Strategic planning for the leveraging of advanced information technology in astrobiology and space science research. (4/1/01 – present)

**James Crawford**, Ph.D., 1990, Computer Science, University of Texas, Austin, Knowledge representation, satisfiability, constraint satisfaction, planning and scheduling. (11/05/01 – present))

**Richard Dearden,** Ph.D., 2000, Computer Science, University of British Columbia. Artificial intelligence, reasoning, and decision making under uncertainty (07/19/00 – present)

**John Dowding**, MSE, 1988, Computer Science, University of Pennsylvania, Spoken dialogue systems, speech recognition, language processing (9/18/00 – present)

**Lisa Faithorn**, Ph.D, 1990, Anthropology, University of Pennsylvania, Technology enhanced team research communication and calibration, human dimensions and technology options (9/18/00 – present)

**Robert E. Filman,** Ph.D., 1979, Computer Science, Stanford University. Frameworks for distributed computing. (12/01/99 – present)

**Bernd Fischer**, Ph.D., 2001, U Passau, Germany, Computer Science, Automated Software Engineering, Machine Learning (11/01/98 – present)

**Dave Gehrt**, JD Law, 1972, University of Washington, UNIX system administration, security, and network based tools (1/84 - 7/85, 2/1/88 - present).

**Dimitra Giannakopoulou**, Ph.D, 1999, Distributed Software Engineering, Imperial College, London, Distributed computing, design, analysis and implementation (8/28/00 – present)

**Zann Gill**, M.Arch, 1979, Harvard University, Collaborative problem-solving environments (CPSE), concept formation (the creative process). (03/07/00 –2/20/02)

**Howard Goldstein**, 1963, MS, Chemical Engineering, University of Arizona, Research on thermal protection systems that utilize intelligent systems technology and smart materials (7/1/00 – present)

**Julian Gomez**, Ph.D., 1985, Computer and Information Science, Ohio State University, Binding 3-D computer graphics with Information Technology (5/6/02 – present)

**Mohana M. Gurram**, BA, 2000, Computer Science/Engineering, University College of Engineering, Kakatiya University, Data mediation and synthesis project. (6/3/02 – 8/30/02)

**James Hieronymus**, Ph.D., 1971, Applied Physics, Cornell University, Spoken Dialogue System, automated speech recognition, acoustic phonetics. (11/1/00 – present)

**Butler Hine**, Ph.D., 1988, Astronomy, University of Texas, Austin. Autonomous Systems for Space and Planetary Exploration. (11/01/99 - present) (Currently IPA to NASA Ames)

**Beth Ann Hockey**, Ph.D., 1998, Linguistics, University of Pennsylvania, (01/01/99 – present)

**Marjory J. Johnson**, Ph.D., 1970, Mathematics, University of Iowa, High-performance networking for both space and ground applications (1/9/84 -present).

**Ari K. Jonsson,** Ph.D., 1997, Computer Science, Stanford University, New Techniques in constraint satisfaction and Automated Planning. (10/1/97 - present)

**Chang-Sung Kim**, Ph.D., 2001, Computational Fluid Dynamics, Seoul National University, Research on hemodynamics of human circulatory system. (1/17/02 – present)

**Frank Kuehnel**, Ph.D, 2000, Physics, Michigan State University, Computervision in 3D super-resolution and artificial intelligence. (08/01/00 - present)

**John Loch**, Ph.D., 1998, Computer Science, University of Colorado at Boulder (12/28/98 – 3/8/02)

**William McReady,** Ph.D., 1991, Condensed Matter Theory, University of Toronto, Machine Learning, Knowledge Discovery. (8/26/02 – present)

**Ronald Mak**, M.S., 1975, Computer Science, Stanford University, Information structures (3/18/02 – present)

**David Maluf**, Ph.D., McGill University, Canada, 1995, Post Doctoral Fellow, Stanford University, knowledge-based systems, database management systems, image databases, computer vision, and man machine interfaces. (01/03/00 – present)

**Robin D. Morris**, Ph.D., 1995, Engineering, Trinity College, Cambridge, U.K., Bayesian Inference and Image Analysis, Mathematical Modeling of Natural Imagery and Associated Computational Methods. (3/1/98 – present)

**Lewis Peach – USRA Chief Engineer** (09/04/99 – present)

**Charles Pecheur**, Ph.D., 1996, Electrical Engineering, Computer Science, University of Liege, (11/06/98 – present)

**Kanna Rajan**, MA, 1989, Computer Science, New York University, Automated reasoning, planning and scheduling. Model-based reasoning, knowledge representation and space craft autonomy. (04/03/00 – present)

**Grigore Rosu**, Ph.D, 2000, Computer Science, University of California, San Diego. Formal methods in system specification and verification, applied logics in computer science. (9/1/00 – present)

**Johann Schumann**, Ph.D, 1991, Computer Science, Technical University of Munich, Germany (Neural networks, automated theorem, parallel symbolic systems (03/03/00 – present)

**Karl Schweighofer**, Ph.D, 1995, Chemistry, UCSC. Research in structural genomics and bioinformatics (8/1/01 - present)

**Maarten Sierhuis**, Ph D, 2001, University of Amsterdam, Intelligent multi-agent simulation agent-oriented programming languages (4/1/98 – present)

**Ashok Srivastava**, Ph.D., 1996, Electrical Engineering, University of Colorado, Machine Learning, Knowledge Discovery. (8/19/02 – present)

**Doron Tal**, Ph.D 1997, Computational Neuroscience, Boston University, Computer vision, Bayesian inference, machine learning, robotics, artificial intelligence. (3/14/02 – present)

**Jean Jerry Toung,** MS, 2000, Electrical Engineering and Computer Science, Universite de Nantes, France, High performance WAN network monitoring and technology deployment for the NGI/NREN project. (07/15/00 – present)

**Udo von Toussaint**, Ph.D., 2000, Physics, University of Bayreuth, Germany, Bayesian inference on ill-posed inverse problems. (1/2/01 – 1/18/02)

**Ian A. Twombly**, Ph.D., 1997, Biophysics, John Hopkins University, (9/1/98 – present)

**Willem Visser**, Ph.D., 1998, Computer Science, University of Manchester (10/19/98 – present)

**Richard Washington**, Ph.D., 1994, Computer Science, Stanford University. Artificial Intelligence, Robotics, Autonomous Systems, Plan Execution. (04/03/00 – present)

## *IV.D. Visiting Scientists and Consultants*

Note: Dates shown for visiting scientists are not necessarily continuous.

**Howard Barringer**, Ph.D., 1978, Computer Science, University of Manchester, Formal specification and verification methods for software and hardware.(4/3/02 – 6/7/02)

**Marsha Berger**, Ph.D., 1982, New York University, Courant Institute, Computer Science, Automated mesh generation and flow computations using Cartesian embedded boundary approach (6/12/02 – 8/22/02)

**Pavel Bochev**, Ph.D., 1994, Mathematics, Virginia Tech, Sandia National Laboratory, Numerical methods for PDE's; optimal control of Navier-Stokes (5/19/02 – 5/23/02)

**Julia Brodsky**, MS, 1996, Space Physics, St. Petersberg Technical University, International Space Station training expert. (10/01/01 – 9/30/02)

**Charles Brownstein**, Ph.D., 1975, Lehigh University, Information Management Requirements Panel Study (11/28/01 – 11/30/01)

**Alessandro Cimatti**, Ph.D., IRST, Trento, Italy, Symbolic Model Checking and Applications (April 2002)

**Charles Dey**, University of Illinois, Computational analysis and visualization development; internet/intranet database design, analysis, development (1/1/02 – present)

**Suhrit Dey**, Ph.D., Aerospace Engineering, Mississippi State University; Professor, Eastern Illinois University, Computational Immunology with applications to prevention/cure of breast cancer. (12/17/01 – 12/30/01)

**Michael Duke**, Ph.D., 1963, Geochemistry, California Institute of Technology, Moon and Mars Geoscience; exploration and development of space resources (5/1/01 – 4/30/02)

**Mohammed Adel El Sayed El-Beltagy**, Ph.D., 2000, Mechanical Engineering, University of Southampton, Machine Learning (6/6/02 – 6/7/02)

**Michael Folk**, Ph.D., 1972, Computer Science, Syracuse University, Participate in the Information Management Requirements Panel Study (11/28/01 – 12/01/01)

**Scott Fouse**, MS, 1983, Electrical Engineer, University of Southern California, Advanced information technologies (11/28/01 – 11/30/01)

**James C. French**, Ph.D., 1982, Computer Science, University of Virginia, Participate in the Information Management Requirements Panel Study (11/28/01 – 12/01/01)

**Tarek El-Ghazawi**, Ph.D., George Mason University, Distributed Shared-memory Programming with the Unified Parallel C Information Power Grid, (09/21/00 – present)

**Yolanda Gil**, Ph.D., 1992, Carnegie Mellon University, Information Management Requirements Panel Study (11/28/01 – 11/30/01)

**Genevieve Gorrell**, MS., 1999, Computer Speech and Language Processing, Cambridge University, Spoken dialogue systems (5/27/02 – 8/16/02)

**Sara Graves**, Ph.D., Computer Science, University of Alabama at Huntsville, Distributed information systems; data mining, (9/1/01 – 9/30/02)

**Daniel Hammerstrom**, Ph.D., 1977, EE University of Illinois. Oregon Graduate Institute. Massively parallel VLSI architectures for intelligent systems. (03/01/00 – 9/30/02)

**William Hefley**, Ph.D., 1998, Information Technology, Carnegie Mellon University, Organization and Technology Change, (11/23/01 – 11/27/01)

**Andre Jalobeanu**, Ph.D., 2001, Universite de Nice Sophia Antipolis, France, 3D modeling, super-resolution, inverse problems in imaging (1/02/02 – present)

**Roby Joehanes**, MS, 2002, Computer Science, Kansas State University, Artificial Intelligence; Robotics (7/1/02 – 7/19/02)

**Richard G. Johnson**, Ph.D. - Physics, Indiana University, 1956, Global environmental problems and issues (11/1/92 - present).

**Vipul Kashyap**, Ph.D., 1997, Computer Science, Rutgers University, Information integration and management, knowledge management, metadata, ontologies and the semantic web. (2//20/02 – 2/22/02)

**Nicholas Kingsbury**, Ph.D., 1975, Electrical Engineering, University of Cambridge, Image processing and image coding with emphasis on wavelet transforms and statistical models for signals. (5/13/02 - 5/25/02)

**Craig Knoblock**, Ph.D., 1991, Computer Science, Carnegie Mellon University, Co-Chair of the Information Management Requirements Panel Study (9/01/01 – 9/30/02)

**Carl Lagoze**, MSE, 1987, Software Engineering, Wang Institute, Participate in the Information Management Panel Study (11/28/01 – 12/01/01)

**Norman Lamarra**, Ph.D., 1982, Computer Science, UCLA; CalTech/JPL, Software architecture for distributed systems. (2/7/02 –2/8/02)

**Lawrence Lannom**, MS, 1976, Library and Information Science, University of Illinois, Co-Chair of the Information Management Panel Study (9/01/01 – 9/10/02)

**Ronald Larsen**, Ph.D., 1981, Computer Science, University of Maryland, Information Management Requirements Panel Study (11/28/01 – 11/30/01)

**Supratik Mukhopadyay**, Ph.D., 2001, Computer Science Max Planck Institute. University of Pennsylvania, Program Analysis, distributed systems, formal methodology. (4/22/02 – 4/23/02)

**Stanley Peters**, B.S., 1963, Mathematics, MIT, Spoken dialogue systems and semantics. (12/20/01 – present)

**Martha Pollack**, Ph.D., 1986, Computer and Information Science, University of Pennsylvania, AIPS conference presenter – Artificial Intelligence, Plan Management, Reasoning in Dynamic Environments. (4/22/02 – 4/28/02)

**Richard Rohwer**, Ph.D., Physics, University of Texas, Austin, Machine Learning (5/30/02 – 5/31/02)

**Steven W. Running**, Ph.D., 1979, Forest Ecophysiology, Colorado Statue University, Global ecology using remote sensing data to study ecological systems to solve environmental problems. (11/28/01 – 11/29/01),

**Gerardo Schneider**, MS., 1996, Computer Science, Verimag, Verification of hybrid systems. (3/28/02 – 3/28/02)

**Simon Buckingham-Shum**, Ph.D., 1991, Psychology, and HCI, University of York, UK. Knowledge media and sensemaking hypertext and discourse. (2/4/02 – 2/5/02)

**Bjorn Sjogreen**, Ph.D., 1988, Numerical Analysis, Uppsala University. Royal Institute of Technology, Sweden. Development of efficient and accurate numerical methods for the computation of compressible fluid flows. (5/14/02 – 8/14/02)

**Zhendong Su**, Ph.D., 2002, Computer Science, U. C. Berkeley, Static Analysis for Software Systems. (9/3/02 – 1/3/03)

**Katia Sycara**, Ph.D., 1987, Computer Science, Georgia Technology Institute, Participate in the Information Management Requirements Panel Study (11/28/01 – 12/01/01)

**Sheridan M. Tatsuno**, MS, 1977, Planning and Public Policy, Harvard's Kennedy School of Government, Dreamscape Global, Development of an organizational outline for an innovative NASA-industry collaborative problem-solving environment (CPSG) for adoption by IASC. (7/112/01 – 9/01/01)

**Sankaran Venkateswaran**, Ph.D. -Sr. Research Associate, University of Tennessee Space Institute, preconditioning methods and use of neural networks for application to semi-conductor materials (5/18/98 – 9/30/02)

**Robert Young**, Ph.D., 1982, University of Texas, Director of Technology, SciComp, Inc. Simulation programs from high-level specification language. (12/01/01 - 12/04/01)

**David Zingg**, Ph.D., University of Toronto, Aerospace Engineering, Navier-Stokes equations (07/01/00 – 09/30/00) (7/1/01 - 9/28/01)

## *IV.E. Visiting Students*

**Allessandro Acquisti,** MS, Information Management, University of California, Berkeley, Simulating work practices using Brahms agent-oriented language. (1/16/01 – present)

**Greg Barish,** MS., 1998, Computer Science, University of Southern California. AIPS conference attendee (4/22/02 – 4/28/02)

**Dan Bohus**, BS, 2000, Computer Science, Carnegie Mellon University, spoken dialogue systems, conversation agents. (6/3/02 – 6/16/02)

**Nate Blaylock**, B.A., University. of Rochester, mentored by John Dowding, Spoken Dialogue Systems Among Collaborating Intelligent Agents (6/10/02 – present)

**Mary Ellen Campana**. BS, 2000, Computer Science, University of Rochester, Using eye movements to improve spoken dialogue system performance. (6/3/02 – 8/27/02)

**Jamieson Cobleigh,** MS, 2001, Computer Science, University of Massachusetts, Software engineering and verification (6/10/02 – 8/16/02)

**Judah ben De Paula,** BS, 2000, Computer Science, Georgia Tech, Knowledge base inference engines (6/3/02 – 9/30/02)

**Karen Brennan,** (7/02/01 – 6/30/02)

**Emanuel Grant**, MSE, 2002, Computer Science, Colorado State University, (6/03/02 – 8/16/02)

**Alex David Groce**, BS, 1999, Computer Science, Carnegie Mellon University, Formal verification of Java programs. (6/02/02 – 8/16/02)

**Laleh Haghshenass**, B.A., Computer Science, Workspace Analysis and Brahms Modeling. (7/16/02 – present)

**Laurie Satsue Hiyakumoto**, M.S., Construction Engineering, Carnegie Mellon University, AIPS Conference attendee (4/22/02 – 4/28/02)

**Max Horstmann**, BA, 2001, Computer Science University of Karlsruhe, Application of MDP models on Mars Rovers. (6/3/02 – 8/31/02)

**Frank Hutter**, BA, 2001, Computer Science, Darmstadt University of Technology, Situational awareness for Mars Rovers. (6/10/02 – 8/16/02)

**Charis Kaskiris**, MS, 1996, International Trade Policy and Analysis, University of Michigan, Brahms modeling of a planetary astronaut/robot EVA. (5/28/01 – Present)

**Sarfraz Khurshid**, MS, 1998, Computer Science, MIT, Symbolic analysis of Java programs for test case generation. (6/24/02 – 8/30/02)

**Flavio Lerda**, MS, 2000, Computer Engineering, Politecnico di Torino, Italy, Formal verification of Java/C/C++ programs (6/3/02 – 8/16/02)

**Laurentiu Leustean**, MS, 1997, Computer Science, University of Bucharest, Optimality control and kalman filters. (6/10/02 - 8/16/02)

**Colleen McCarthy**, M.S., Computer Science, University of Pittsburgh, AIPS attendee (4/22/02 – 4/28/02)

**Katharine M. Mullen**, Computer Science, Bard College, Development of the AutoBayes automated software synthesis system. (6/24/02 – 8/30/02)

**Tae-young Oh**, Device modeling and simulation, Stanford University. (1/1/98 – present)

**Bart Peintner**, M.S., Computer Science, University of Michigan, AIPS attendee (4/22/02 / 4/28/02)

**Ny Aina Razermera** Mamy, MS, 2000, Computer Science, University of Oregon, Software verification principles and tools. (6/25/02 – 9/2/02)

**Oksana Tkachuk**, MS, Computer Science, Kansas State University, SSRP02, Verification of software (5/20/02 – 7/26/02)

**Kristen Brent Venable**, MS, 2001, Mathematics, Padova, Italy, Temporal constraint and satisfaction problems. (7/1/02 – 9/6/02)

**Rene E. Vidal**, MS, 2000, Electrical Engineering, UC Berkeley, Bayesian Computer Vision, 6/4/02 – present)

**Chieh-Chih Wang**, MS, 1996, Robotics and Computer Science, Carnegie Mellon University, Computer vision localization and mapping (6/3/02 – 8/10/02)

**Terry Zimmerman**, M.S., 2002, Computer Science, Arizona State University, AIPS Conference attendee (4/22/02 – 4/27/02)

## *IV.F. Science Council*

**Dr. David Bailey,** Chair, Lawrence Berkeley Labs

**Dr. Nabil Adam,** Rutgers University

**Dr. Daniel G. Bobrow,** Xerox Palo Alto Research Center

**Dr. Jeffrey M. Bradshaw,** University of West Florida

**Dr. Robert (Jack) Hansen,** University of West Florida

**Dr. Abdollah Homaifar,** North Carolina A&T State University

**Dr. Philip Kuekes,** Hewlett Packard Laboratories

**Dr. Mitchell P. Marcus,** University of Pennsylvania

**Dr. Alain Rappaport,** Carnegie Mellon University

RIACS
Mail Stop T041-5
NASA Ames Research Center
Moffett Field, CA 94035