Grand Challenges Conference Application

Computer Science Research Challenges – A NASA Problem-Driven View

Daniel E. Cooke, Professor and Chair, Computer Science Department, Texas Tech University.
Butler Hine, Manager, Intelligent Systems Program, NASA Ames Research Center.

NASA has recently increased its focus on computer science research. "When people think of space, they think of rocket plumes and the Space Shuttle, but the future of space is information technology. We must develop a virtual presence, in space, on planets, in aircraft and spacecraft." [Dan Goldin, May 1996] In the shadows of the largest wind tunnel in the world, the number one priority at the NASA Ames Research Center is computer science research.

NASA's increased focus has arisen out of its realization that the major technological roadblocks impeding the nation's ability to perform earth and space science, and to achieve its exploration goals, center on computation and Artificial Intelligence. To address many of these research challenges, the Intelligent Systems (IS) Program was recently established at NASA. The IS program office first determined the gaps in NASA capabilities taking into account current and future science and exploration goals. Based upon this "gap analysis" the program was designed to advance research efforts to begin addressing NASA's future needs. In the following, we summarize two categories of significant challenges facing NASA. Some of the significant goals related to these challenge areas are italicized below.

Near Term. Data Understanding and the Exploitation of Parallel Computing Capabilities. Currently, NASA is receiving two terabytes of data per day from earth observing satellites alone. NASA is able to acquire and store vast amounts of data, but the amount of data is duly stressing our ability to analyze the data. One can view these vast data sets as empirical data. Scientists typically endeavor to reduce empirical observations to concise theories, which explain the observations. Revolutionary approaches that provide theory-based access to these data sets are included among NASA's goals.

A major challenge is to develop systems capable of reducing large data sets to small algorithmic units – identifying causal relationships in the data sets. The resulting algorithms could be viewed as concise statements of the data – providing more manageable representations of the data that should lead to better understanding – and perhaps might be capable of reproducing the data sets. Efforts at the Santa Fe Institute are investigating the relationship between theories and the amount of data the theories explain. They are analyzing these relationships through an application of Kolmogorov's Complexity measure, called Algorithmic Information Content (AIC). Given a particular message string, the programs that will print the string and then halt are identified. The length of the shortest program is called the AIC of the string. One can envision a ratio where the size (i.e., number of characters) of the shortest program serves as the numerator and the size of the message (i.e., the number of characters in the program's output) serves as the denominator. For example, a program that computes millions of the digits of *pi* will result in a fraction close to zero. A message that is not the product of a formula or algorithm will simply be a print statement in which the entire message is a literal. In such cases, the ratio approaches one. One approach to data understanding might attempt to discover ways in which data can be analyzed in order to identify the shortest program that can produce the data. These algorithmic units, could then serve as the "theories" explaining the data, and could possibly result in data compression.

To perform large-scale data analysis often requires high performance computing. The NASA Ames Research Center has one of the largest supercomputing facilities in the world. Efforts are underway to purchase and install a 2,048 node Silicon Graphics Origin Supercomputer. Currently a 1,024 SGI processor system is operated. Adding more processors alone is not enough. People must ultimately be able to exploit the compute power the hardware provides. There are at least two challenges to our ability to exploit these systems. First, there is typically, a point of diminishing returns when new processors are added: the overhead in adding nodes outweighs the performance gains. The multilevel parallelism (MLP) package, developed at Ames, has succeeded in managing the overhead one normally experiences when adding more nodes to a problem solution, resulting in near linear speed-ups for key supercomputing applications. NASA Ames Research Center has conducted empirical studies, proceeding from 256 to 512 to 1,024 processor systems. Recently, SGI estimated that speed-ups would continue to improve in systems up to 30,000 processors. Therefore, efforts to extend our ability to exploit more processors in terms of speed-up concerns are being addressed at places like Ames.

An additional concern, however, continues to serve as a challenge. *Automated approaches to identify <u>what</u> parallelisms are inherent in a problem solution, in addition to current efforts that identify <u>how</u> to carry out the parallel operations, are needed.* Successes in this area will make parallel systems more accessible and will further advance our ability to exploit parallel architectures.

Far Term. Telepresence of Humans – Real and Virtual Presence of Humans on Distant Planets. Libration points are distributed throughout our solar system. These are points where gravitational forces of the surrounding planets cancel each other out. There is, for example, an earth-moon libration point and a sun-Mars libration point, connected by a stream of libration points, sometimes called Martin Lo paths. These "streams" are gravity-free, low energy transfer paths that dynamically run through space. Currently NASA is investigating the possibility of deploying way stations at the earth-moon and the sun-Mars libration points, and having a low-cost shuttle that will transfer people and supplies between them. This approach to Mars exploration significantly lowers the risk and the cost of exploration, especially when compared to the current estimates for a direct earth-Mars transfer of people and supplies to the Martian surface. If significant challenges to automated reasoning and human-centered computing are addressed, astronauts in the sun-Mars libration point may be able to have a deeper and richer exploration experience than they would on the surface. A revolutionary combination of human-robotic reasoning capable of achieving the levels of human reasoning in a relatively seamless fashion is required.

The most primitive form of reasoning is called **Stimulus-Response** reasoning. In this form of reasoning, an organism is capable of processing an external stimulus and determining an appropriate response. Some S-R reasoning is so primitive, that it is viewed as reflex rather than reasoning. When a person touches a hot stove, he/she will instinctively pull away with little or no consideration of the situation. Note, however, that data analysis to determine the causal relationship between the heat source and pain, as well as significant control and feedback are indeed taking place whether or not we are aware of these actions. A more sophisticated form of reasoning is called **Stimulus-Stimulus**. S-S reasoning occurs when an organism receives a stimulus and, in turn, produces a stimulus for another organism or for some tool or machine. The most sophisticated form of reasoning is called **Offline Reasoning**. One simple view of this form of reasoning is to envision humans as having a primitive brain that performs the S-R and S-S reasoning. This primitive brain deals with external stimuli and is incapable of originating thoughts that are not triggered by outside occurrences. Now envision a more sophisticated brain that spends its time monitoring the behavior of the primitive brain – reflecting upon and analyzing situations. Observation and sophisticated analysis leading to discovery and invention is O-L reasoning and embodies the creative ability of humans. *Challenges to automated reasoning include, the need to provide robotic devices with total and complete reflexive capabilities and to push the envelope to provide richer stimulus-stimulus reasoning.* These capabilities are required so that much of the human ability to skillfully move and use tools can be safely relegated to robots deployed to the surface. Many of our practiced skills with tools and with our own bodies become so reflexive, that performing these skilled activities does not always reach our awareness. If we can achieve these abilities in robots on the surface of Mars, then fine-grained teleoperated control of the devices will not be necessary. This is important to the feasibility of having astronauts explore Mars from a sun-Mars libration space station. Currently, it is difficult to teleoperate complex robotic systems with a communications time delay of more than a few seconds, based upon the current state-of-the-art in predictive control. There is a 7.2-light second round-trip delay from the sun-Mars libration point and the Martian surface. Therefore the robots on the surface will need to handle reflexive fine-grained control on their own. Humans, however, will be able to provide the higher levels of reasoning in an effective manner from the space station. (Humans on earth will not be able to perform the same functions due to time delays ranging from 6 to 45 minutes roundtrip, at the speed of light between earth and Mars.)

Humans will provide the high level reasoning functions. The challenges to a libration point teleoperation and virtual teleportation to the surface are many-fold. *The goal is for humans at the libration point to have a full sensory experience of being on the surface*. The machines on the surface will have to perform flawless S-R and S-S functions, and be equipped with sensors that will provide tactile, visual, auditory, and olfactory signals that, when sent to the space-station will be used to directly stimulate these senses in the astronauts. We are envisioning direct stimulation – e.g., we are not targeting simple streaming of high definition video. Instead, we anticipate having sensors and downstream equipment that will directly stimulate the retina. Therefore, cognitively and perceptually the astronauts will *be* on the surface of Mars. Their bodies, however, will be in the protected environment of the space station. *The information flow among the levels of reasoning will exist as a challenge.* Furthermore, the manner in which complexity is organized within the levels of reasoning as distributed among machines and humans presents a challenge. Machines will likely perform their reasoning in a manner different from the manner in which humans perform the reasoning. Therefore, the translations and “cognitive permeability” among the distributed levels of reasoning must be fully understood and seamless communication among the levels must be achieved. Finally, since language is the way humans organize and communicate their understanding of complex phenomena, language research challenges abound.

Biographical Sketch--Daniel E. Cooke, Ph.D.

Daniel Cooke serves as Professor and Chair of the Computer Science Department at Texas Tech University. Dr. Cooke has recently completed an 18-month assignment as the Manager of NASA’s Intelligent Systems Program. The program is a 5-year $320M national research initiative in computer science aimed at NASA relevant problems. Cooke led the activities to establish the technical content of the program, took it from formulation to implementation, and established the program office, which he headed at NASA Ames Research Center in Mountain View, California.

Previously, Dr. Cooke served as an Assistant, Associate, and then Professor and Chair of the Computer Science Department at the University of Texas at El Paso (UTEP).

Since 1990, Daniel Cooke has published more than 75 technical papers in the areas of computer language design and software engineering. He has served as PI or Co-PI on research grants totaling more than $8 million, edited many journal special issues, edited a book on Computer Aided Software Engineering, and served as chair or vice-chair for 15 international conferences or workshops. He currently serves as the Formal Methods Area Editor of the *International Journal of Software Engineering and Knowledge Engineering* and as the Chair of the IEEE Computer Society Technical Committee on Computer Languages.

Cooke has been an American Electronics Association Fellow, a MacIntosh-Murchison Faculty Fellow, and held the MacIntosh-Murchison Chair in Engineering at UTEP. In 1996 he was the recipient of the University of Texas at El Paso's Distinguished Achievement in Research Award.

Recent Papers

Daniel E. Cooke and Per Andersen, “Automatic Parallel Control Structures in SequenceL,” *Software Practice and Experience*, Volume 30, Issue 14, (November 2000), 1541-1570.

Daniel Cooke, "An Introduction to SEQUENCEL: A Language to Experiment with Nonscalar Constructs," *Software Practice and Experience,* Vol. 26(11), (November, 1996) 1205-1246.

Daniel E. Cooke, "SequenceL Provides a Different Way to View Programming," *Computer Languages* 24 (1998) 1-32*.*

Daniel E. Cooke and David Dampier, "A Realtime Version of SequenceL," *Annals of Software Engineering,* Volume 7 (1999), pp. 157-176*.*

Daniel E. Cooke, Joseph Urban, and Scott Hamilton “UNIX and Beyond: An Interview with Ken Thompson,” *COMPUTER* (May, 1999) Vol. 32 No. 5, pp. 58-64*.*

Daniel E. Cooke and Scott Hamilton, “New Directions for NASA Ames Research Center,” *COMPUTER* (January 2000) Vol. 33 No. 1, pp.63-71.

Daniel E. Cooke and Luqi, "Logic Programming and Software Maintenance," *Annals of Mathematics and Artificial Intelligence,* 21 (1997) 221-229.

Biographical Sketch--Butler Hine, Ph.D.

Dr. Hine is currently the Manager of the Intelligent Systems Program, a NASA effort to develop intelligent spacecraft and vehicles, and technology to enable highly capable teams of humans and automation to solve some of NASA’s most pressing problems. Prior to this, Dr. Hine was President and CEO of a Silicon Valley software company that developed advanced visualization tools for managing large corporate networks. His previous NASA experience includes directing the Intelligent Mechanisms Laboratory at NASA Ames Research Center, which pioneered the use of telepresence and virtual reality to control remote science exploration systems.