Notes on the Technical Sessions of the Third HPCD Workshop
July 18-19, 1996
Rutgers University
New Brunswick, NJ

Brian D. Davison
Khaled Shehata Rasheed

September 1996

1 Welcoming Address – Joseph Seneca, University VP for Academic Affairs

1.1 Summary of Presentation

Speaking to close to 50 workshop attendees, Joseph Seneca kicked off the beginning of the workshop with a very brief address. In it, he extended a University welcome to everyone. He noted the importance of using a multidisciplinary paradigm (through-out the University) and the importance in general of partnerships between university and industry, allowing for risk sharing.

2 Introduction to the HPCD Project and to the Workshop – Saul Amarel

2.1 Summary of Presentation

Dr. Amarel laid out the goals of the project, where the project stands, a quick summary of what has been accomplished and what remains to be done in the future.

The HPCD project is a four year, large scale, multi-disciplinary, multi-institutional project supported by DARPA. The broad objective of HPCD is to build on top of advances in

- High performance computing
- Artificial intelligence
- Modeling/Simulation technology
- Software for design automation

to develop a new generation of engineering automation technology that can bring about dramatic gains in industrial productivity.

The focus of the project is on complex engineering design problems. These tend to be computationally intensive, hard to decompose, have many (possibly conflicting) design goals and constraints, making it hard to reason from goals to possible solutions. Examples of the problems addressed in HPCD include:

- Design of inlets and nozzles for jet engines
- Conceptual design of aerospace vehicles
- Design of innovative surface ships
- Design of photolithographic processes for electronic chip fabrication
- Design of microprocessors tailored to specific classes of problems

The project thrust is to:

- Develop methodologies/systems that enable substantial improvement in the solution of such complex engineering design problems with the help of computers.
- Explore increased automation of processes for solving complex engineering design problems.
- Develop prototype systems and experience in specific realistic domains/tasks-testbeds.
- Shorten the distance between science and engineering with the help of HPC.

The design problems that we are exploring often need computer power at the level of HPC. However, raw computer power alone will not be always sufficient for handling the computational complexity of the design tasks that we are considering. This leads us to look for:

- Ways of decomposing domain theories into a multiplicity of models at various levels of abstraction and approximation.
- Ways of using models appropriately in design decisions.

2.2 Questions and Comments

When he was asked whether design speedup included time for manufacturing, Dr. Amarel answered that we focused on design, and not manufacturing but that some areas considered manufacturing constraints such as microprocessor design. Also, the work in microlithography focuses on design of manufacturing processes.

Bob Lucas noted that the statement that "Speedups in design time of at least one order of magnitude were achieved in several design domains/tasks" was particularly strong, and asked what caused the order of magnitude speedup? Dr. Amarel responded saying that in most cases it wasn't the computers, it was the algorithms. He gave the example of Steve Orzag's work in masks – parallelization produced a 100 times factor speedup. Bob Lucas wanted more, though, asking how much was a result of algorithms and how much was hardware, etc. Dr. Amarel's answer was that algorithms and methodology were responsible in some domains such as aircraft and nozzle design, but parallelism was responsible in the masks domain.

3 Design of Propulsion Systems and Components – Doyle Knight, Andrew Gelsey et al

3.1 Summary of Presentation

The objectives of this part of the project are:

- To improve both the efficiency and quality of high speed inlet design through the integration of computational fluid dynamics and artificial intelligence into an automated optimal design methodology
- To develop and broaden methodologies for automated optimal design which extend to other design domains.

In year three of the HPCD project, work on the redesign of the NASA P8 Hypersonic inlet was completed. The design of an axi-symmetric Supersonic mixed-compression inlet, as a prototype inlet for Project Cheapshot (Tomahawk missile replacement), was started and completed on year Three. The existence of two types of analysis codes motivated a multi-level design strategy in which the faster and less accurate analysis code was used for the major part of the search and the expensive high quality code was used for verification purposes. The search engines used included Random Probes (RP), CFSQP and the adapted Genetic Algorithm (GA) developed at Rutgers. As expected, CFSQP and the GA did much better than RP and the GA did a little better than CFSQP. These results were published in AIAA conference papers and submitted to Journal of Aircraft. A 22% improvement in performance over conventional methods was achieved for the same or less calendar and human time.

The ongoing research includes the design of a 3-D High Speed Civil Transport (HSCT) Inlet and the design of a 3-D Supersonic Missile Inlet.

3.2 Questions and Comments

Bob Lucas commented that a parallel version of GASP would be useful.

4 Design of Air Vehicles and related methodologies – Andrew Gelsey et al

4.1 Summary of Presentation

Dr. Gelsey described various research tasks that were accomplished on the third year of the project. He presented the ideas mainly in the context of aerospace vehicle design, though some of those ideas were applied successfully in other domains (missile inlet design for instance).

The work on the Modeling/Simulation Associate (MSA) has continued from the beginning of the project. The main function of the MSA is to supervise the communication between the Design Associate (DA) and the simulators used in the design process. The need for the presence of the MSA stems from the fact that most existing simulators were developed for use by human designers rather than automated design codes such as the DA. In previous years, various methods of communication were explored such as:

- Methods for detecting model violations
- Methods for handling model violations within the simulators
- (Crude) communication between MSA and DA: failure driven random multistart

In the third year, more advanced communication strategies were explored and contrasted. The model constraints strategy (a strategy in which, upon failure of a simulation, the MSA returns to the DA a quantified description of the degree of violation of each modeling assumption) was shown to be very useful in the case of CFSQP optimization. The Model Penalties strategy (similar to the model constraints except that all the modeling assumption violations are consolidated into one penalty quantity) was adequate in the case of genetic algorithm based optimization.

The idea of multilevel design and modeling was explored in greater depth in the third year. The presence of two models with different levels of abstraction in the domain of supersonic aircraft design made it possible to reduce the cost of optimization by an order of magnitude through careful switching between models.

Genetic Algorithms were successful in reducing the cost of optimization in the supersonic aircraft design domain by an order of magnitude, using only one level of abstraction. The GA used was adapted for continuous design space search as part of the research. The multilevel design and modeling strategy with CFSQP gave better results than using a single level with GAs. The idea of combining multilevel modeling and GA optimization is very promising but experiments are yet to be completed.

4.2 Questions and Comments

Bob Lucas asked how you know when you've made enough model constraints for a simulator. Dr. Gelsey responded, saying that it is much like debugging a program – you converge to something that can be trusted (and you repeatedly go back to experts to verify new designs).

Paul Kantor wanted to know why Fuel Mass wasn't a design constraint (as opposed to a model constraint). Dr. Gelsey answered by saying that it was computed backwards and related to the assumption that fuel fits in a fuel tank.

Dr. Amarel agreed with the question, but Dr. Gelsey commented that it is all the same to the simulator; in general it depends on how the model is formulated.

Bob Lucas posed the question, "(When exploring a search space,) how do you determine which direction to go?" Dr. Gelsey's response was that since evaluation is inexpensive (less than a second) we try all directions and compute the gradient.

Eduardo Sontag asked "How do you know the 'global optimum'?" In answering, Dr. Gelsey said that it's really the apparent global optimum. The space has many stopping points. The only way is to start with many starting points. If lots of them get to some best point, it's probably a global optimum. Additionally, we want a robust optimum (i.e. one that is isolated so that small changes such as manufacturing variations move it out of optimality would be bad).

Al Despain asked, "What if you wanted to consider discrete design changes like 1, 2 or 4 engines?" Dr. Gelsey's answer was that one possibility is to use real numbers, like 2.3 and check neighbors, or use mixed programming. After stating that "you are doing abstraction by ignoring parameters," When Dr. Despain asked whether linear programming would work better, Dr. Gelsey responded that it would likely not be as good, since we had an idealized nozzle.

Dr. Sontag also asked what the population sizes used in the Genetic Algorithms were. Khaled Shehata Rasheed responded, saying they were 10 times the dimensionality of the space. Additionally, Dr. Sontag wanted to know how the gradients were estimated. Dr. Gelsey explained that they computed full gradients, and use rule-based gradients to handle nearby unevaluable points.

5 Design Associate (DA), Modeling-Simulation Associate (MSA), Visualization-aided Interfaces and Applications to Ship and Aircraft Design – Tom Ellman, Deborah Silver et al

5.1 Summary of Presentation

This part of the project is mainly concerned with increasing the robustness and efficiency of the automated design process. This is done through the development of new methodologies for interaction and collaboration among the different pieces of code used in the design/optimization process as well as the use of visualization and designer interaction to guide the search process. Some of the subareas of research include:

- An interactive visual environment for engineering design (The DA-MSA)
- Knowledge-based management of legacy codes
- Multi-level modeling in engineering design optimization

The objectives of the DA-MSA are to support rapid development of automated design systems in domain specific applications and to reduce design time while improving design quality and reliability. In year three, the DA-MSA was enhanced by adding the following components:

- A system for representing and reformulating design optimization strategies
- A system for maintaining a design strategy development record
- Methods of using multiple models and multiple search spaces in design optimization
- A framework for integrating visualization and program synthesis in a new generation of engineering design environments

Knowledge based management of legacy codes for automated design enables rapid construction of good evaluation models for automated design from legacy simulation programs. It also improves the robustness of design optimization by avoiding or handling failures in simulation codes. The idea was implemented and tested on yacht, nozzle, and airframe design domains. The effectiveness of failure handling schemata was demonstrated. The reusability of schemata was demonstrated across test domains.

5.2 Questions and Comments

Bob Lucas asked about the size of the databases and tables used in the development record. Dr. Ellman responded saying that we keep track of every evaluation of takeoff mass but we can control it. Dr. Silver said that it depends; you can dump everything but you don't normally want to. It could be gigabytes of data – the user steers it.

Bob Lucas also asked how a searcher decides when to abandon the space. Dr. Ellman answered that we are using sequential quadratic programming which does the search locally (i.e. determines when to explore the space). In this case, we performed a parameter study to generate the data visualized here.

When Paul Kantor asked whether the human was crucial to determine which constraint causes the failure, Dr. Ellman answered that it currently was, but doesn't have to be in the future.

During Deborah Silver's part of the talk, Paul Kantor said that there are several spaces involved – which do you visualize? Dr. Silver answered that it could be anything – it's a general technique. So far we have only used design parameters.

Bob Lucas asked whether the visualization is being performed post mortem, to which Dr. Silver answered yes, unless the code runs in real-time in which case it could be used for computational steering. Paul Kantor suggested waiting a few years so that it can be done in real-time. Dr. Silver responded that the problems we are interested in keep getting more difficult.

Another question asked was whether there are in fact systems that allow the user to do computational steering. Dr. Silver answered that there are systems that use supercomputers on small problems but these haven't been broadly successful since simulations are expensive.

Larry Reeker asked have you gotten any insights as to why people can visualize these things... using their human sensory reasoning? Lou Steinberg suggested that humans use sensory pattern recognition and domain knowledge to make it possible. To automate the role of the person doing the steering you should mimic this in addition to pattern recognition. Paul Kantor added that this system, as an interactive tool, needs to work well for lots of different kinds of people.

6 Design of the Voice Mimic – James Flanagan et al

6.1 Parametric representation of speech–the Voice Mimic Approach – James Flanagan

6.1.1 Summary of Presentation – James Flanagan

In this part of the project, our objective is to use the computational power of HPC and new understanding in acoustic signal processing to represent speech information in terms of compact, robust, natural parameters. We use an Adaptive Voice Mimic as our approach. We have produced a working prototype system for articulatory description of speech.

6.2 System realization – Gael Richard

6.2.1 Summary of Presentation

We are continuing to improve and extend the voice mimic system to handle fricative sounds. Despite the simplicity of the model used for fricative generation, this system has produced vowel/consonant/vowel utterances of very encouraging quality. This necessitated several significant improvements of the speech synthesizer, a design of a multi-stage open-loop analysis, and the definition of an efficient closed-loop control. The best results so far have been obtained using a rather detailed open-loop analysis and a highly constrained closed-loop analysis based on the initial solution provided by the open-loop analysis. Significant improvements have also been obtained using perceptual codebooks (automatically generated using a perceptual criterion based on the first two formant frequencies).

Four open-loop sound samples were played. For each sample, natural speech and synthetic speech compared favorably.

6.2.2 Questions and Comments

Al Despain asked, saying, "You are talking about vocal chords giving the fundamental frequency... How do you infer this frequency?" The response provided was that the fundamental frequency can be computed by fitting perfectly periodic synthetic signals to natural speech segments.

6.3 Speech synthesis from fluid flow – Daniel Sinder

6.3.1 Summary of Presentation

We are attempting to characterize speech generation as a fluid dynamic process, to develop reduced models for improved speech synthesis, and to utilize those models in the voice mimic system. Numerical simulations and flow visualizations were used. Vowels of encouraging quality were produced. Additionally, we have begun experimental and computational studies of stylized dental fricatives.

A video demonstration of flow animations was presented, showing the pressure contours and particle velocities in a uniform tube and in stylized shapes of a dental fricative.

6.4 Use of articulatory parameters for speaker recognition – Magali Goirand

6.4.1 Summary of Presentation

Preliminary experiments were performed using vocal tract length (VTL) estimation. VT length is useful for identification because it is speaker specific and independent of speakers emotions or health; dependent on the utterance and stable on vowels. The general scheme is to compare the processed signal to codebooks to estimate the VT length. Encouraging results for the VT length estimation were achieved in terms of compute time and stability of results.

A tape of an X-ray database of a vocal tract was shown while the subject was speaking.

6.4.2 Questions and Comments

A question was asked as to whether all the models were hard boundary models. The answer was that while some were soft boundary the flow and the Navier-Stokes modems are all hard.

6.5 Goals/Milestones/Contract performance – Jim Flanagan

6.5.1 Summary of Presentation

Dr. Flanagan summarized the achievements over the past year, including generation of consonants as well as vowels, and performance of articulatory-based speech recognition.

7 Design of Microlithographic Processes - Don Smith, Eytan Barouch, Steve Orszag et al

7.1 Summary of Presentation

Don Smith outlined the topics in Cluster I (Area I.1: Design for manufacture at sub-1/4 micron feature size and Area I.2: Design of Computers). He described the need for a rapid, flexible response to changes in technology or application requirements. In lithography, for example, brute force rescaling isn't possible, and entire redesign is too expensive. He showed that it was possible, however, to adaptively redesign components that were negatively affected by rescaling.

To perform adaptive redesign, a Circuit Design Associate has been implemented (with two main parts: a Geometric Design Associate and an Electrical Design Associate). These two subsystems communicate and determine which edges of the circuit to move to improve overall speed. It was successfully tested with an XOR circuit produced in MAGIC; a faster design was produced with no additional skew.

7.2 Questions and Comments

Al Despain asked whether edges were extended just one at a time, and Dr. Smith answered yes, and that the system moved metal edges over poly but didn't move contacts. Dr. Despain also suggested that additional operators would be useful.

Dr. Despain also asked whether Spice was too computationally expensive, to which Dr. Smith replied that it was, for what they wanted to do (i.e. scale beyond the XOR example), and Lou Steinberg mentioned that Spice is also harder to wrap.

8 Microprocessor Design - Lou Steinberg, Alvin Despain et al

8.1 Summary of Presentation

The objective of this research area is to study design across multiple levels of artifact representation. Digital circuit design already has many levels and much use of CAD systems, but little support for the overall design process. In particular, feedback from lower levels of representation is not automated at all. As a result, most existing tools are brittle, making new technologies difficult to adopt.

We are focusing on the coordination and control of tasks at different levels, incorporating information flow in both directions. We believe the key is to balance exploration with decision making. To that end, we've developed a method for rational control of multi-level stochastic design. At each step, the method determines which abstraction level to work on next and which of alternative designs to choose by using statistical and probabilistic techniques to determine the best choice. This approach has been applied to artificial "stub" design problems and to the problem of module-placement in circuit design.

Other work includes encapsulating IRSIM to simulate circuits at the switch level within an automated system; building a framework for doing instruction-scheduling to evaluate pipelined processors, and an instruction-level simulator.

8.2 Questions and Comments

In the context of the multi-level rational control strategy, Al Despain asked, given n levels, what do they represent, and how many levels is best? Dr. Steinberg answer was that this theory gives us a way to analyze this.

Saul Amarel asked whether the strategy was more than an issue of control, i.e. you need representations at a level and be able to evaluate it. To this, Dr. Steinberg agreed.

Dr. Despain also made the suggestion to try skipping the middle level of the testbed system (i.e. doing placement in a single level) and compare the results, and Dr. Steinberg agreed.

9 Tools and Algorithms for Parallel Processing, Apostolos Gerasoulis et al

9.1 Summary of Presentation

The obvious goal for this research is to utilize parallel processing to achieve significant speedups on HPCD problems. To do so, we need to develop algorithms for mapping problems to processors and for scheduling work on those processors to maximize throughput.

Recently we've developed new scheduling heuristics for dynamic and iterative programs with the properties that the task graph changes slowly through time and contains local regions of new computation. These heuristics allow for automatic re-mapping of the computation dynamically during runtime to increase overall performance.

We are interested in improving performance of an SAIC program that does ship body interactions called LAMP (Large-Amplitude Motion Program). We've analyzed the system for appropriate parallelization and are working with SAIC on a scalable version of LAMP that uses clustering techniques from PYRROS, and expect scalable performance for all LAMP codes. Additionally, we are applying our scheduling methods to Gaussian elimination within Helmholz equation solvers (used in the Microlithography area).

10 Wrap-up, Saul Amarel

10.1 Summary of Presentation

In his closing address, Saul Amarel summarized the activities that cut across the efforts in HPCD, pointing out the key technical issues, and summarizing the achievements that have been accomplished.

Our overall approach has been to develop prototype HPC-aided systems for complex engineering design. At the same time, we've had to develop methods, design principles, and tools and environments and apply these and other methods and tools to new design tasks.

Some areas that still need work; we need to show where we've had an impact, and we need to attempt to generalize – to show that what we've learned can apply to other domains.

There are three primary technical issues in our work:

1. Design Processes, including Design requirements, Solution evaluation, Search spaces, Control of search/optimization, and System frameworks;

2. Domain knowledge, including Model formulation, Development of simulation algorithms based on models Analysis of models, and Model reduction; and

3. High Performance Computing, including Parallelization and Software tools for programming in parallel and distributed environments.

From our experience over the last few years, we can make a few technical notes, such as that handling a design problem is an evolutionary process, and that we've seen that the design loop is the essential core of design processes.

The conclusions that I draw include the fact that we've made significant achievements in all areas and substantial technical progress – we've had a major impact in specific design domains/tasks. Additionally, as a research group we've gained cohesion, and have formed some clear ideas about future research.

To end our workshop, I'd like to summarize our achievements. We've had speedups in design time of at least one order of magnitude in several domains. Several simulation/analysis tools have been parallelized. Several complex searches/optimizations in design space are being guided by knowledge of regularities in the space. Complex multi-level search/optimization strategies are being developed. Solution evaluation with multi-level models is being explored in several areas. Issues of design representation and of setting problems and search spaces appear in each of the domains. Finally, problems of system integration, including interconnecting new software modules with legacy codes have been encountered in each area.