# REPORT DOCUMENTATION PAGE



AFRL-SR-BL-TR-00-
0764

| 1. AGENCY USE ONLY *(Leave blank)* | 2. REPORT DATE | 3. REPOR |
|---|---|---|
| | December, 1997 | |

**4. TITLE AND SUBTITLE**
1997 Summer Research Program (SRP), Graduate Student Research Program (GSRP), Final Reports, Volume 9, Rome Laboratory

**5. FUNDING NUMBERS**
F49620-93-C-0063

**6. AUTHOR(S)**
Gary Moore

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Research & Development Laboratories (RDL)
5800 Uplander Way
Culver City, CA 90230-6608

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Air Force Office of Scientific Research (AFOSR)
801 N. Randolph St.
Arlington, VA 22203-1977

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION AVAILABILITY STATEMENT**
Approved for Public Release

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

The United States Air Force Summer Research Program (USAF-SRP) is designed to introduce university, college, and technical institute faculty members, graduate students, and high school students to Air Force research. This is accomplished by the faculty members (Summer Faculty Research Program, (SFRP)), graduate students (Graduate Student Research Program (GSRP)), and high school students (High School Apprenticeship Program (HSAP)) being selected on a nationally advertised competitive basis during the summer intersession period to perform research at Air Force Research Laboratory (AFRL) Technical Directorates, Air Force Air Logistics Centers (ALC), and other AF Laboratories. This volume consists of a program overview, program management statistics, and the final technical reports from the GSRP participants at the Rome Laboratory.

**14. SUBJECT TERMS**
Air Force Research, Air Force, Engineering, Laboratories, Reports, Summer, Universities, Faculty, Graduate Student, High School Student

**15. NUMBER OF PAGES**

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | UL |

## GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reports. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to ***stay within the lines*** to meet ***optical scanning requirements.***

**Block 1.** Agency Use Only *(Leave blank)*.

**Block 2.** Report Date. Full publication date including day, month, and year, if available (e.g. 1 Jan 88). Must cite at least the year.

**Block 3.** Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

**Block 4.** Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

**Block 5.** Funding Numbers. To include contract and grant numbers; may include program element number(s), project number(s), task number(s), and work unit number(s). Use the following labels:

| | |
|---|---|
| **C** - Contract | **PR** - Project |
| **G** - Grant | **TA** - Task |
| **PE** - Program Element | **WU** - Work Unit Accession No. |

**Block 6.** Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

**Block 7.** Performing Organization Name(s) and Address(es). Self-explanatory.

**Block 8.** Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the report.

**Block 9.** Sponsoring/Monitoring Agency Name(s) and Address(es). Self-explanatory.

**Block 10.** Sponsoring/Monitoring Agency Report Number. *(If known)*

**Block 11.** Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with....; Trans. of....; To be published in.... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

**Block 12a.** Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

| | |
|---|---|
| **DOD** - | See DoDD 5230.24, "Distribution Statements on Technical Documents." |
| **DOE** - | See authorities. |
| **NASA** - | See Handbook NHB 2200.2. |
| **NTIS** - | Leave blank. |

**Block 12b.** Distribution Code.

| | |
|---|---|
| **DOD** - | Leave blank. |
| **DOE** - | Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports. |
| | Leave blank. |
| **NASA** - | Leave blank. |
| **NTIS** - | |

**Block 13.** Abstract. Include a brief *(Maximum 200 words)* factual summary of the most significant information contained in the report.

**Block 14.** Subject Terms. Keywords or phrases identifying major subjects in the report.

**Block 15.** Number of Pages. Enter the total number of pages.

**Block 16.** Price Code. Enter appropriate price code *(NTIS only)*.

**Blocks 17. - 19.** Security Classifications. Self- explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contains classified information, stamp classification on the top and bottom of the page.

**Block 20.** Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

UNITED STATES AIR FORCE

SUMMER RESEARCH PROGRAM -- 1997

GRADUATE STUDENT RESEARCH PROGRAM FINAL REPORTS

VOLUME 9

ROME LABORATORY

RESEARCH & DEVELOPMENT LABORATORIES

5800 Uplander Way

Culver City, CA 90230-6608

Program Director, RDL
Gary Moore

Program Manager, AFOSR
Major Linda Steel-Goodwin

Program Manager, RDL
Scott Licoscos

Program Administrator, RDL
Johnetta Thompson

Program Administrator, RDL
Rebecca Kelly-Clemmons

Submitted to:

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Bolling Air Force Base

Washington, D.C.

December 1997

## PREFACE

Reports in this volume are numbered consecutively beginning with number 1. Each report is paginated with the report number followed by consecutive page numbers, e.g., 1-1, 1-2, 1-3; 2-1, 2-2, 2-3.


| VOLUME | TITLE |
|---|---|
| 1 | Program Management Report |
| | ***Summer Faculty Research Program (SFRP) Reports*** |
| 2A & 2B | Armstrong Laboratory |
| 3A & 3B | Phillips Laboratory |
| 4A & 4B | Rome Laboratory |
| 5A , 5B & 5C | Wright Laboratory |
| 6 | Arnold Engineering Development Center, United States Air Force Academy and Air Logistics Centers |
| | ***Graduate Student Research Program (GSRP) Reports*** |
| 7A & 7B | Armstrong Laboratory |
| 8 | Phillips Laboratory |
| 9 | Rome Laboratory |
| 10A & 10B | Wright Laboratory |
| 11 | Arnold Engineering Development Center, United States Air Force Academy, Wilford Hall Medical Center and Air Logistics Centers |
| | ***High School Apprenticeship Program (HSAP) Reports*** |
| 12A & 12B | Armstrong Laboratory |
| 13 | Phillips Laboratory |
| 14 | Rome Laboratory |
| 15A & 15B | Wright Laboratory |
| 16 | Arnold Engineering Development Center |

GSRP FINAL REPORT TABLE OF CONTENTS i-xi

1. INTRODUCTION 1

2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM 2

3. RECRUITING AND SELECTION 3

4. SITE VISITS 4

5. HBCU/MI PARTICIPATION 4

6. SRP FUNDING SOURCES 5

7. COMPENSATION FOR PARTICIPATIONS 5

8. CONTENTS OF THE 1995 REPORT 6

APPENDICIES:

A. PROGRAM STATISTICAL SUMMARY A-1

B. SRP EVALUATION RESPONSES B-1

GSRP FINAL REPORTS

# SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Armstrong Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Benedict N Arrey | Univ of Texas at San Antonio , San Antonio , TX<br>Identification and Quantitation of N-menthyl-1-(3,4 Methylenedioxyhenyl)-2- Butanamine Together wiht | AL/AOT | 7- 1 |
| MS JoAnne E Bates | University of North Dakota , Grand Forks , ND<br>A Way to Condense the Time Consuming Procedure of Cognitive Task Analysis | AL/HRCT | 7- 2 |
| MR Brandon B Boke | Trinity University , San Antonio , TX<br>Effects of Brain Temperature on Fatugue in Rats due to maximal Exercise an Millimeter Microwave Radi | AL/OER | 7- 3 |
| MS Constance R Buford | Alabama A&M University , Normal , AL<br>Assessment of Coagulant Agents on The Reduction of Aqueous Film Forming Foam (AFFF) in Wastewater | AL/OEB | 7- 4 |
| MS Dawn D Burnett | Wright State University , Dayton , OH<br>The Effects of Individual Differences and Team Processed ofn Team Member Schema Similarity and Task | AL/CFHI | 7- 5 |
| MR Bradley E Collie | Arizona State University , Mesa , AZ<br>Perception of Velocity as a Function of the Oculomotor State of the Eyes | AL/HRA | 7- 6 |
| MR Gerald W DeWolfe | Univ of Texas at Austin , Austin , TX<br>Investigation and Validation of Submaximal Cycle Ergometry Protocols Used to Assess the Aerobic Capa | AL/PSP | 7- 7 |
| Kea U Duckenfield | The Virginia Institute of Marine Science , Gloucesster Point , VA<br>Direct Measurment of DNAPL/Water Contact Area in the Subsurace: One-And three-Dimensional Studies | AL/EQL | 7- 8 |
| MR Phillip T Dunwoody | University of Georgia , Athens , GA<br>The Effects of Task Structure on Cognitive Organizing Principles: Implicatins | AL/CFTO | 7- 9 |
| MR Daniel X Hammer | Univ of Texas at Austin , Austin , TX<br>Measurement of Dispersive Curves for Ocular Media by white-light Interferometry | AL/OEO | 7- 10 |
| MS Catherine R Harrison | Univ of Illinois Urbana/Champaign , Champaign , IL<br>Gender effects in Wayfinding Strategy: Implications for Teamand Individual Trainging | AL/HRCC | 7- 11 |

| Author | University/Institution<br>Report Title | Armstrong Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MS Laura J Hott | Wright State University , Dayton , OH<br>Examination of an Organizational Climate Measure and theRelationship with Grievances and Turnover | AL/HRG | 7- 12 |
| MS Vanessa D Le | Univ of Texas at Austin , Austin , TX<br>A Clearance Study of Nitrotyrosine From a Prostate Cancer Cell Line | AL/OER | 7- 13 |
| MS Kavita Mahajan | Trinity University , San Antonio , TX<br>the Effect of 2.06 GHz Microwave Irraditation on The Permeability of the Blood Brain Barrier | AL/OER | 7- 14 |
| MR Thomas R Mertz (Jr.) | Univ of Scranton , Scranton , PA<br>Protocol for Development of Amplicons for a Rapid & Efficient Method of Genotyping Hepatitis C Virus | AL/AOEL | 7- 15 |
| MR Michael J Miller | Texas A & M Univ-College Station , College Station , TX<br>An Psycholmetric Examination of the Multidimensional work ethic Profile among Air Force enlisted per | AL/HRCF | 7- 16 |
| MR Miguel A Moreno | Arizona State University , Tempe , AZ<br>The Effect of Size Disparity on Perception of Surface Slant in Steroscopic Moving Images | AL/HRA | 7- 17 |
| MR Brian W Moroney | University of Cincinnati , Cincinnati , OH<br>The Role of Multi-Modal Adaptive Interfaces in Providing Flight Path Guidance | AL/CFHI | 7- 18 |
| MR Randy J Mueller | University of Connecticut , Storrs , CT<br>Desorption and Biodegradation of Dinitrotoluenes in Aged soils | AL/EQL | 7- 19 |
| MR Mark A Murphy | Ohio University , Athens , OH<br>Implementation of Freflex/merlin Telcoperation | AL/CFBA | 7- 20 |
| MS Cynthia J Musante | North Carolina State U-Raleigh , Raleigh , NC<br>Well-Posedness for a Class of Nonlinear Distributed Parameter Models wiht Time Delay Arising in Adva | AL/OES | 7- 21 |
| MR David C Powell | The College of William and Mary , Gloucester , VA<br>Investigaatioon of the Iron-Beaaring Phases of the Columbus AFB Aquifer | AL/EQL | 7- 22 |

# SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Armstrong Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Christopher S Schreiner | Miami University , Oxford , OH<br>The Effect of Visual Similarity and Reference Frame Alignment on the Recognition of Military Aircra | AL/HRCT | 7- 23 |
| MR John N Sempeles | University of Florida , Gainesville , FL<br>OH Radical Reaction Rate Constant & Product Study of 2-Propoxyethanol | AL/EQL | 7- 24 |
| MS Julie A Stiles-Shipley | Bowling Green State University , Bowling Green , OH<br>The Effects of Observation and Training Schedule on The Acquisition of a complex Computer Based | AL/HRCT | 7- 25 |
| MR Robert S Tannen | University of Cincinnati , Cincinnati , OH<br>Integrating Multisensory Displays for an Adaptive Target Leading Interface | AL/CFHP | 7- 26 |
| MR Paul J Taverna | Tulane University , New Orleans , LA<br>A Preeliminary Examination of ECL Activity Geared Toward a CD+2 Sensor | AL/EQL | 7- 27 |
| MR James M Tickner | Univ of Scranton , Scranton , PA<br>Molecular typing of Cndida Parasilosis Via Amplified Fragment Lenghth Polymorphism and Repetitive S | AL/AOEL | 7- 28 |
| MS Deanne L Westerman | Case Western Reserve Univ , Cleveland , OH<br>A Test of the Misattributed-Activation Hypothesis of the Revelatin Effect in Memory | AL/HRCC | 7- 29 |

# SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Phillips Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Joshua C Bienfang | University of New Mexico , Albuquerque , NM<br>Frequency Stabilization of an Nd; Yag Laser | AFRL/DEL | 8- 1 |
| MR Marc L Breen | Tulane University , New Orleans , LA<br>A Study of Defects and Dark Current Mechanisms in Triple-Junction GaInP2/GaAs/Ge Photovoltaic Cells | PL/VTV | 8- 2 |
| MR Jerome T Chu | University of Florida , Gainesville , FL<br>The Characterization of High Performance Quantrum Well Infrared Photodetectors for Low Background O | PL/VTMR | 8- 3 |
| MR Theodore S Elicker | University of N. C.- Charlotte , Charlotte , NC<br>Simulation fan Modeling of Nanoelectronic Materials | PL/VTMR | 8- 4 |
| MR Jeffery M Ganley | University of New Mexico , Albuquerque , NM<br>A Preliminary Study of the Causes of Spring-IN in A Unidirectional Carbon Fiber/EPOXY Composite | PL/VTV | 8- 5 |
| Johnelle L Korioth | Univ of Texas at Dallas , Richardson , TX<br>A Prreliminary analysis of Stacked blumleins Used in Pulsed power Devices | PL/WSQ | 8- 6 |
| Kelly K Lau | Univ of Texas at Arlington , Arlington , TX<br>Experimental Validation of Theree-Dimensional Reconstruction of Inhomogeneity Images in turbid Media | PL/LIMI | 8- 7 |
| MS Ruthie D Lyle | Polytechnic University , Farmingdale , NY<br>A Quasi-Particle Analysis of Amplitude Scintillation | PL/GPS | 8- 8 |
| MR Shaun L Meredith | Massachusetts Inst of Technology , Cambridge , MA<br>Research on Plasma Diagnostics for Versatile Toroidal Facility: Gridded energy Analyzers | PL/GPS | 8- 9 |

# SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Phillips Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Eric J Paulson | Univ of Colorado at Boulder , Boulder , CO<br>A Study of Atomospheric Perturbations On a Suborbital Space Plane Skipping Trajectory | AFRL/PR | 8- 10 |
| MR Christopher W Peters | Univ of Michigan , Ann Arbor , MI<br>A New "Technique Used to Dertemine the Time Evolutin of The Frequency in Heterodyne Systems | PL/WSQN | 8- 11 |
| MR Michael J Rowlands | Massachusetts Inst of Technology , Cambridge , MA<br>Ducted Whistler waves and Emissions in the Laboratory and the Ionosphere | PL/GPS | 8- 12 |
| MS Lorena L Sanchez | University of New Mexico , Albuquerque , NM<br>A Preliminary Study of the Effects of Process Conditions on Curvature in Graphite/EPoxy Pultruded Ro | PL/VTV | 8- 13 |
| MR John H Schilling | Univ of Southern California , Los Angeles , CA<br>"A Study of Alternate Propellants for Pulsed Plasma Thrusters | PL/RKEE | 8- 14 |
| MR Kenneth F Stephens II | University of North Texas , Denton , TX<br>Investigation of an Explosively Formed Fus Using Mach2 | AFRL/DEH | 8- 15 |
| MS Jane A Vladimer | Boston University , Boston , MA<br>Low Latitude Ionospheric Tec Measured by Nasa Topex | PL/GPS | 8- 16 |
| MR Michael V Wood | Pennsylvania State University , University Park , PA<br>Characterization of Spatial Light Modulator For Application to real-time Hlography | PL/LIMS | 8- 17 |
| MR. Mark C Worthy | Univ of Alabama at Huntsville , Huntsville , AL<br>Library of the Natural Frequency Responses for Cylindrical & Rectangular Shaped Plastic Mines | PL/WSQW | 8- 18 |
| MR John Yoon | University of Florida , Gainesville , FL<br>Simulating Transient Temperature Distributions in Optically Pumped Multillayer Laser Structures | PL/LIDA | 8- 19 |

SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Rome Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Tony M Adami | Ohio University , Athens , OH | RL/C3 | 9- 1 |
| MR Richard S Andel | SUNY Binghamton , Binghamton , NY<br>Visual Target Tracking and Extraction from a Sequence of Images | RL/IRRE | 9- 2 |
| MR Patrick M Garrity | Central Michigan University , Mt. Pleasant , MI<br>An Examination of Java and CORBA Security | RL/CA-II | 9- 3 |
| MR Walter I Kaechele | Rensselaer Polytechnic Instit , Troy , NY<br>Operational Analysis of an Actively Mode-Locked Fiber Laser | RL/OCPA | 9- 5 |
| MR William J Kostis | Cornell University , Ithaca , NY | RL/OCSS | 9- 6 |
| MS Helen Lau | Syracuse University , Syracuse , NY<br>A Simulati9n Study on a Partitioning Procedure for Radar Signal processing Problems | RL/OCSS | 9- 7 |
| MR Myron R Mychal | Illinois Inst of Technology , Chicago , IL<br>Simulaton of a Robust Locally Optimum Receiver in Correlated Noise Using Autoregressive Modeling | RL/C3BB | 9- 8 |
| MS Maryanne C Nagy | SUNY OF Tech Utica , Utica , NY | RL/IWT | 9- 9 |
| DR Luke J Olszewski | Georgia Southern University , Statesboro , GA<br>Software Veification Guide Using PVS | RL/ERDD | 9- 12 |
| MR Charles M Palmer | George Washington University , Washington , DC<br>A Technique for locating and characterizing crystalline regions in simulated solids | RL/ERDR | 9- 10 |
| MR Dsunte L Wilson | Brown University , Providence , RI<br>System-Level Hardware/Software Partitioning of Hetergeneous Embedded Systems | RL/ERDD | 9- 11 |

# SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Wright Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Mark L Adams | Auburn University , Auburn , AL<br>A Study of Laser Induced Plasma Damage of Thin Metal Foils | WL/MNMF | 10- 1 |
| MR James T Belich | Bethel College , St. Paul , MN<br>Contribution of a scene Projecotr's Non-Uniformity to a test Article's output Image Non-Uniformity | WL/MNGI | 10- 2 |
| MR Jason W Bitting | Louisiana State University , Baton Rouge , LA<br>Visualization and Two-Color Digital PIV Measurements in Circular and Square Coaxial Nozzles | WL/POSC9 | 10- 3 |
| MR Lawrence L Brott | University of Cincinnati , Cincinnati , OH<br>Synthesis of a Novel Second Order Nonlinear Optical Polymer | WL/MLBP | 10- 4 |
| MS Danielle E Brown | Wright State University , Dayton , OH<br>An Experimental and Computational Analysis of the Unsteady Blade Row Potential Interaction in a Tran | WL/POTF | 10- 5 |
| MS Angela M Cannon | Pennsylvania State University , University Park , PA<br>the Synthesis of a Protected Carboxylic Acid Derivative for Attachment To C60 | WL/MLPJ | 10- 6 |
| MR Charles C Conklin | Florida State University , Tallahassee , FL<br>Vision Algorithms For Military Image Processing | WL/MNMF | 10- 7 |
| MR Mitchell G Dengler | University of Missouri - Rolla , Rolla , MO | WL/MLIM | 10- 8 |
| MR James D Drummond | University of Cincinnati , Cincinnati , OH<br>Invesstigation of Conductive Cladding Layers for Improved Polimg in Non-Linear Optical Polymer waveg | WL/MLPO | 10- 9 |
| MR Gary W Dulaney | Brown University , Providence , RI<br>Computer Simulation of Fire Suppression in Aircraft Engine Nacelles | WL/FIVS | 10- 10 |

SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Wright Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Robert L Parkhill | Oklahoma State University , Stillwater , OK<br>Organically modified Silicate Films as Corrosion Resistant Treatments for 2024-T3 Aluminum Alloy | WL/MLBT | 10- 23 |
| MS Annie R Pearce | Georgia Inst of Technology , Atlanta , GA<br>Cost-Based Risk Prediction and Identification of Project Cost Drives Using Artificial neural Network | WL/FIVC- | 10- 24 |
| MR Dax B Pitts | University of Cincinnati , Cincinnati , OH<br>A Study of Intra-Class Variability in ATR Systems | WL/AACR | 10- 25 |
| MR Jonathan M Protz | Massachusetts Inst of Technology , Cambridge , MA<br>An LPV Controller for a Tailess Fighter Aircraft Simulation | WL/FIGC | 10- 26 |
| MR Jason E Riggs | Clemson University , Clemson , SC | WL/MLPJ | 10- 27 |
| MR Thomas W Scott | University of Missouri - Rolla , Rolla , MO<br>Iso-Octane and N-Heptane Laminar Flame Numerical Study | WL/POPS | 10- 28 |
| MR Steven R Stanfill | University of Florida , Gainesville , FL<br>A study of HRR Super Resolution Analysis for Possible ATR Performance Enhancement | WL/AACR | 10- 29 |
| Adedokun W Sule-Koiki | Howard University , Washington , DC<br>Detection Techniques use in Forward-Lookeng Radar Signal A Literature Review | WL/AAMR | 10- 30 |
| MR. Robert M Taylor | Purdue University , West Lafayette , IN<br>Rapid Modeling for Aircraft Design Synthesis | WL/FIBD | 10- 31 |
| MS Laura E Williams | Georgia Inst of Technology , Atlanta , GA<br>Data Simulation Supporting Range Estimating forResearch and Sevelopment Alternatives | WL/FIVC- | 10- 32 |

SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Wright Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR Cornelious W Williams Jr. | University of Cincinnati , Cincinnati , OH<br>Allyl & Propargyl Resins | WL/MLBC | 10- 33 |
| MS Melissa R Wilson | University of Missouri - Rolla , Rolla , MO<br>A Study of The Particulate Emissions of A Well-Stirred Reactor | WL/POSC | 10- 34 |
| Sami Zendah | Wright State University , Dayton , OH<br>Develop an Explosive simulated Testing Apparatus for Impact Physics Research at Wright Laboratory | WL/FIV | 10- 35 |

| Author | University/Institution<br>Report Title | Wright Laboratory<br>Directorate | Vol-Page |
|---|---|---|---|
| MR David W Fanjoy | Purdue University , West Lafayette , IN<br>Demonstration of Genetic Algorithms for Wngineering Optimization Problems | WL/FIIB | 10- 11 |
| | Western Michigan University , Kalamazoo , MI<br>Comparison of self-assembled monolayers and chitosan as functional substrates for deposition fo ultr | WL/MLPJ | 10- 12 |
| MR Carl C Hoff | Wright State University , Dayton , OH<br>Similarity Measures for pattern Recognition | WL/AACA | 10- 13 |
| MR Adam R Hoffman | Wright State University , Dayton , OH<br>Evaluation and Integratin of Electrodynamic Simulation Packages for Madmel Program | WL/POOX | 10- 14 |
| MR. George W Jarriel, Jr. | Auburn University , Auburn , AL<br>Exploding Foil Initiator Flyer Velocity Measurement Using VISAR | WL/MNMF | 10- 15 |
| MR Brett A Jordan | Wright State University , Dayton , OH<br>Capacitor Based DC Backup Power Supply with Integrated Cahrging Circuit | WL/POOC | 10- 16 |
| MR Edward L Kiely | Ohio State University , Columbus , OH | WL/FIBD | 10- 17 |
| MS Janae N Lockett | University of Toledo , Toledo , OH<br>A Study of Electronics Design Environments in Terms of Computer aided Design A Psychological Persper | WL/AAST | 10- 18 |
| MS Stephanie Luetjering | University of Dayton , Dayton , OH<br>Fatigue Crack GrowthBehavior of Ti-22A1-23Nb | WL/MLLN | 10- 19 |
| MR Alfred L Malone | Auburn University , Auburn , AL<br>Electrical and Mathematical Characterization of th Semiconductor | WL/MNMF | 10- 20 |
| MR Herbert F Miles II | Tulane University , New Orleans , LA | WL/MLLN | 10- 21 |
| Dawn H Miller | Georgia Inst of Technology , Atlanta , GA | WL/FIVC | 10- 22 |

# SRP Final Report Table of Contents


| Author | University/Institution<br>Report Title | Arnold Engineering Development Center<br>Directorate | Vol-Page |
|---|---|---|---|
| MS Jessica L Thomas | Tennessee Univ Space Institute , Tullahoma , TN<br>Incorporating Condensation into Nastd | AEDC | 11- 1 |
| MR Derek E Lang | University of Washington , Seattle , WA<br>Hue Analysis Factors For Liquid Crystal Transient Heat Tranasfer Measurements | USAFA | 11 – 2 |
| MS Bridget V McGrath | Univ of Colorado at Colorado Springs , Colorado Spring , CO<br>A Setup for Photoassociation of cold, Trapped Cesium Atoms | USAFA | 11 – 3 |
| MS Donna M Lehman | Univ of Texas Health Science Center , San Antonio , TX<br>Relationship between Growth Hormone and Myelin Basic Protein Expression in Vivo | WHMC | 11 – 4 |

# 1. INTRODUCTION

The Summer Research Program (SRP), sponsored by the Air Force Office of Scientific Research (AFOSR), offers paid opportunities for university faculty, graduate students, and high school students to conduct research in U.S. Air Force research laboratories nationwide during the summer.

Introduced by AFOSR in 1978, this innovative program is based on the concept of teaming academic researchers with Air Force scientists in the same disciplines using laboratory facilities and equipment not often available at associates' institutions.

The Summer Faculty Research Program (SFRP) is open annually to approximately 150 faculty members with at least two years of teaching and/or research experience in accredited U.S. colleges, universities, or technical institutions. SFRP associates must be either U.S. citizens or permanent residents.

The Graduate Student Research Program (GSRP) is open annually to approximately 100 graduate students holding a bachelor's or a master's degree; GSRP associates must be U.S. citizens enrolled full time at an accredited institution.

The High School Apprentice Program (HSAP) annually selects about 125 high school students located within a twenty mile commuting distance of participating Air Force laboratories.

AFOSR also offers its research associates an opportunity, under the Summer Research Extension Program (SREP), to continue their AFOSR-sponsored research at their home institutions through the award of research grants. In 1994 the maximum amount of each grant was increased from $20,000 to $25,000, and the number of AFOSR-sponsored grants decreased from 75 to 60. A separate annual report is compiled on the SREP.

The numbers of projected summer research participants in each of the three categories and SREP "grants" are usually increased through direct sponsorship by participating laboratories.

AFOSR's SRP has well served its objectives of building critical links between Air Force research laboratories and the academic community, opening avenues of communications and forging new research relationships between Air Force and academic technical experts in areas of national interest, and strengthening the nation's efforts to sustain careers in science and engineering. The success of the SRP can be gauged from its growth from inception (see Table 1) and from the favorable responses the 1997 participants expressed in end-of-tour SRP evaluations (Appendix B).

AFOSR contracts for administration of the SRP by civilian contractors. The contract was first awarded to Research & Development Laboratories (RDL) in September 1990. After completion of the

1990 contract, RDL (in 1993) won the recompetition for the basic year and four 1-year options.

## 2. PARTICIPATION IN THE SUMMER RESEARCH PROGRAM

The SRP began with faculty associates in 1979; graduate students were added in 1982 and high school students in 1986. The following table shows the number of associates in the program each year.

| YEAR | SRP Participation, by Year | | | TOTAL |
|---|---|---|---|---|
| | SFRP | GSRP | HSAP | |
| 1979 | 70 | | | 70 |
| 1980 | 87 | | | 87 |
| 1981 | 87 | | | 87 |
| 1982 | 91 | 17 | | 108 |
| 1983 | 101 | 53 | | 154 |
| 1984 | 152 | 84 | | 236 |
| 1985 | 154 | 92 | | 246 |
| 1986 | 158 | 100 | 42 | 300 |
| 1987 | 159 | 101 | 73 | 333 |
| 1988 | 153 | 107 | 101 | 361 |
| 1989 | 168 | 102 | 103 | 373 |
| 1990 | 165 | 121 | 132 | 418 |
| 1991 | 170 | 142 | 132 | 444 |
| 1992 | 185 | 121 | 159 | 464 |
| 1993 | 187 | 117 | 136 | 440 |
| 1994 | 192 | 117 | 133 | 442 |
| 1995 | 190 | 115 | 137 | 442 |
| 1996 | 188 | 109 | 138 | 435 |
| 1997 | 148 | 98 | 140 | 427 |

Beginning in 1993, due to budget cuts, some of the laboratories weren't able to afford to fund as many associates as in previous years. Since then, the number of funded positions has remained fairly constant at a slightly lower level.

## 3. RECRUITING AND SELECTION

The SRP is conducted on a nationally advertised and competitive-selection basis. The advertising for faculty and graduate students consisted primarily of the mailing of 8,000 52-page SRP brochures to chairpersons of departments relevant to AFOSR research and to administrators of grants in accredited universities, colleges, and technical institutions. Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs) were included. Brochures also went to all participating USAF laboratories, the previous year's participants, and numerous individual requesters (over 1000 annually).

RDL placed advertisements in the following publications: *Black Issues in Higher Education, Winds of Change,* and *IEEE Spectrum.* Because no participants list either *Physics Today* or *Chemical & Engineering News* as being their source of learning about the program for the past several years, advertisements in these magazines were dropped, and the funds were used to cover increases in brochure printing costs.

High school applicants can participate only in laboratories located no more than 20 miles from their residence. Tailored brochures on the HSAP were sent to the head counselors of 180 high schools in the vicinity of participating laboratories, with instructions for publicizing the program in their schools. High school students selected to serve at Wright Laboratory's Armament Directorate (Eglin Air Force Base, Florida) serve eleven weeks as opposed to the eight weeks normally worked by high school students at all other participating laboratories.

Each SFRP or GSRP applicant is given a first, second, and third choice of laboratory. High school students who have more than one laboratory or directorate near their homes are also given first, second, and third choices.

Laboratories make their selections and prioritize their nominees. AFOSR then determines the number to be funded at each laboratory and approves laboratories' selections.

Subsequently, laboratories use their own funds to sponsor additional candidates. Some selectees do not accept the appointment, so alternate candidates are chosen. This multi-step selection procedure results in some candidates being notified of their acceptance after scheduled deadlines. The total applicants and participants for 1997 are shown in this table.

| 1997 Applicants and Participants | | | |
|---|---|---|---|
| PARTICIPANT CATEGORY | TOTAL APPLICANTS | SELECTEES | DECLINING SELECTEES |
| SFRP | 490 | 188 | 32 |
| (HBCU/MI) | ( 0 ) | ( 0 ) | ( 0 ) |
| GSRP | 202 | 98 | 9 |
| (HBCU/MI) | ( 0 ) | ( 0 ) | ( 0 ) |
| HSAP | 433 | 140 | 14 |
| TOTAL | 1125 | 426 | 55 |

## 4. SITE VISITS

During June and July of 1997, representatives of both AFOSR/NI and RDL visited each participating laboratory to provide briefings, answer questions, and resolve problems for both laboratory personnel and participants. The objective was to ensure that the SRP would be as constructive as possible for all participants. Both SRP participants and RDL representatives found these visits beneficial. At many of the laboratories, this was the only opportunity for all participants to meet at one time to share their experiences and exchange ideas.

## 5. HISTORICALLY BLACK COLLEGES AND UNIVERSITIES AND MINORITY INSTITUTIONS (HBCU/MIs)

Before 1993, an RDL program representative visited from seven to ten different HBCU/MIs annually to promote interest in the SRP among the faculty and graduate students. These efforts were marginally effective, yielding a doubling of HBCI/MI applicants. In an effort to achieve AFOSR's goal of 10% of all applicants and selectees being HBCU/MI qualified, the RDL team decided to try other avenues of approach to increase the number of qualified applicants. Through the combined efforts of the AFOSR Program Office at Bolling AFB and RDL, two very active minority groups were found, HACU (Hispanic American Colleges and Universities) and AISES (American Indian Science and Engineering Society). RDL is in communication with representatives of each of these organizations on a monthly basis to keep up with the their activities and special events. Both organizations have widely-distributed magazines/quarterlies in which RDL placed ads.

Since 1994 the number of both SFRP and GSRP HBCU/MI applicants and participants has increased ten-fold, from about two dozen SFRP applicants and a half dozen selectees to over 100 applicants and two dozen selectees, and a half-dozen GSRP applicants and two or three selectees to 18 applicants and 7 or 8 selectees. Since 1993, the SFRP had a two-fold applicant increase and a two-fold selectee increase. Since 1993, the GSRP had a three-fold applicant increase and a three to four-fold increase in selectees.

In addition to RDL's special recruiting efforts, AFOSR attempts each year to obtain additional funding or use leftover funding from cancellations the past year to fund HBCU/MI associates. This year, 5 HBCU/MI SFRPs declined after they were selected (and there was no one qualified to replace them with). The following table records HBCU/MI participation in this program.

| SRP HBCU/MI Participation, By Year | | | | |
|---|---|---|---|---|
| YEAR | SFRP | | GSRP | |
| | Applicants | Participants | Applicants | Participants |
| 1985 | 76 | 23 | 15 | 11 |
| 1986 | 70 | 18 | 20 | 10 |
| 1987 | 82 | 32 | 32 | 10 |
| 1988 | 53 | 17 | 23 | 14 |
| 1989 | 39 | 15 | 13 | 4 |
| 1990 | 43 | 14 | 17 | 3 |
| 1991 | 42 | 13 | 8 | 5 |
| 1992 | 70 | 13 | 9 | 5 |
| 1993 | 60 | 13 | 6 | 2 |
| 1994 | 90 | 16 | 11 | 6 |
| 1995 | 90 | 21 | 20 | 8 |
| 1996 | 119 | 27 | 18 | 7 |

## 6. SRP FUNDING SOURCES

Funding sources for the 1997 SRP were the AFOSR-provided slots for the basic contract and laboratory funds. Funding sources by category for the 1997 SRP selected participants are shown here.

| 1997 SRP FUNDING CATEGORY | SFRP | GSRP | HSAP |
|---|---|---|---|
| AFOSR Basic Allocation Funds | 141 | 89 | 123 |
| USAF Laboratory Funds | 48 | 9 | 17 |
| HBCU/MI By AFOSR (Using Procured Addn'l Funds) | 0 | 0 | N/A |
| TOTAL | 9 | 98 | 140 |

SFRP - 188 were selected, but thirty two canceled too late to be replaced.
GSRP - 98 were selected, but nine canceled too late to be replaced.
HSAP - 140 were selected, but fourteen canceled too late to be replaced.

## 7. COMPENSATION FOR PARTICIPANTS

Compensation for SRP participants, per five-day work week, is shown in this table.

1997 SRP Associate Compensation

| PARTICIPANT CATEGORY | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
|---|---|---|---|---|---|---|---|
| Faculty Members | $690 | $718 | $740 | $740 | $740 | $770 | $770 |
| Graduate Student (Master's Degree) | $425 | $442 | $455 | $455 | $455 | $470 | $470 |
| Graduate Student (Bachelor's Degree) | $365 | $380 | $391 | $391 | $391 | $400 | $400 |
| High School Student (First Year) | $200 | $200 | $200 | $200 | $200 | $200 | $200 |
| High School Student (Subsequent Years) | $240 | $240 | $240 | $240 | $240 | $240 | $240 |

The program also offered associates whose homes were more than 50 miles from the laboratory an expense allowance (seven days per week) of $50/day for faculty and $40/day for graduate students. Transportation to the laboratory at the beginning of their tour and back to their home destinations at the end was also reimbursed for these participants. Of the combined SFRP and GSRP associates, 65 % (194 out of 286) claimed travel reimbursements at an average round-trip cost of $776.

Faculty members were encouraged to visit their laboratories before their summer tour began. All costs of these orientation visits were reimbursed. Forty-three percent (85 out of 188) of faculty associates took orientation trips at an average cost of $388. By contrast, in 1993, 58 % of SFRP associates took

orientation visits at an average cost of $685: that was the highest percentage of associates opting to take an orientation trip since RDL has administered the SRP, and the highest average cost of an orientation trip. These 1993 numbers are included to show the fluctuation which can occur in these numbers for planning purposes.

Program participants submitted biweekly vouchers countersigned by their laboratory research focal point, and RDL issued paychecks so as to arrive in associates' hands two weeks later.

This is the second year of using direct deposit for the SFRP and GSRP associates. The process went much more smoothly with respect to obtaining required information from the associates, only 7% of the associates' information needed clarification in order for direct deposit to properly function as opposed to 10% from last year. The remaining associates received their stipend and expense payments via checks sent in the US mail.

HSAP program participants were considered actual RDL employees, and their respective state and federal income tax and Social Security were withheld from their paychecks. By the nature of their independent research, SFRP and GSRP program participants were considered to be consultants or independent contractors. As such, SFRP and GSRP associates were responsible for their own income taxes, Social Security, and insurance.

## 8. CONTENTS OF THE 1997 REPORT

The complete set of reports for the 1997 SRP includes this program management report (Volume 1) augmented by fifteen volumes of final research reports by the 1997 associates, as indicated below:

1997 SRP Final Report Volume Assignments

| LABORATORY | SFRP | GSRP | HSAP |
|---|---|---|---|
| Armstrong | 2 | 7 | 12 |
| Phillips | 3 | 8 | 13 |
| Rome | 4 | 9 | 14 |
| Wright | 5A, 5B | 10 | 15 |
| AEDC, ALCs, WHMC | 6 | 11 | 16 |

## APPENDIX A – PROGRAM STATISTICAL SUMMARY

A. Colleges/Universities Represented

Selected SFRP associates represented 169 different colleges, universities, and institutions, GSRP associates represented 95 different colleges, universities. and institutions.

B. States Represented

SFRP -Applicants came from 47 states plus Washington D.C. Selectees represent 44 states.

GSRP - Applicants came from 44 states. Selectees represent 32 states.

HSAP - Applicants came from thirteen states. Selectees represent nine states.

| Total Number of Participants | |
|---|---|
| SFRP | 189 |
| GSRP | 97 |
| HSAP | 140 |
| TOTAL | 426 |

| Degrees Represented | | | |
|---|---|---|---|
| | SFRP | GSRP | TOTAL |
| Doctoral | 184 | 0 | 184 |
| Master's | 2 | 41 | 43 |
| Bachelor's | 0 | 56 | 56 |
| TOTAL | 186 | 97 | 298 |

| SFRP Academic Titles | |
|---|---|
| Assistant Professor | 64 |
| Associate Professor | 70 |
| Professor | 40 |
| Instructor | 0 |
| Chairman | 1 |
| Visiting Professor | 1 |
| Visiting Assoc. Prof. | 1 |
| Research Associate | 9 |
| TOTAL | 186 |

| Source of Learning About the SRP | | |
|---|---|---|
| Category | Applicants | Selectees |
| Applied/participated in prior years | 28% | 34% |
| Colleague familiar with SRP | 19% | 16% |
| Brochure mailed to institution | 23% | 17% |
| Contact with Air Force laboratory | 17% | 23% |
| *IEEE Spectrum* | 2% | 1% |
| *BIIHE* | 1% | 1% |
| Other source | 10% | 8% |
| TOTAL | 100% | 100% |

## APPENDIX B – SRP EVALUATION RESPONSES

### 1. OVERVIEW

Evaluations were completed and returned to RDL by four groups at the completion of the SRP. The number of respondents in each group is shown below.

Table B-1. Total SRP Evaluations Received

| Evaluation Group | Responses |
|---|---|
| SFRP & GSRPs | 275 |
| HSAPs | 113 |
| USAF Laboratory Focal Points | 84 |
| USAF Laboratory HSAP Mentors | 6 |

All groups indicate unanimous enthusiasm for the SRP experience.

The summarized recommendations for program improvement from both associates and laboratory personnel are listed below:

A. Better preparation on the labs' part prior to associates' arrival (i.e., office space, computer assets, clearly defined scope of work).

B. Faculty Associates suggest higher stipends for SFRP associates.

C. Both HSAP Air Force laboratory mentors and associates would like the summer tour extended from the current 8 weeks to either 10 or 11 weeks; the groups state it takes 4-6 weeks just to get high school students up-to-speed on what's going on at laboratory. (Note: this same argument was used to raise the faculty and graduate student participation time a few years ago.)

## 2. 1997 USAF LABORATORY FOCAL POINT (LFP) EVALUATION RESPONSES

The summarized results listed below are from the 84 LFP evaluations received.

1. LFP evaluations received and associate preferences:

Table B-2. Air Force LFP Evaluation Responses (By Type)

| | | How Many Associates Would You Prefer To Get ? (% Response) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | SFRP | | | | GSRP (w/Univ Professor) | | | | GSRP (w/o Univ Professor) | | | |
| Lab | Evals Recv'd | 0 | 1 | 2 | 3+ | 0 | 1 | 2 | 3+ | 0 | 1 | 2 | 3+ |
| AEDC | 0 | - | - | - | - | - | - | - | - | - | - | - | - |
| WHMC | 0 | - | - | - | - | - | - | - | - | - | - | - | - |
| AL | 7 | 28 | 28 | 28 | 14 | 54 | 14 | 28 | 0 | 86 | 0 | 14 | 0 |
| USAFA | 1 | 0 | 100 | 0 | 0 | 100 | 0 | 0 | 0 | 0 | 100 | 0 | 0 |
| PL | 25 | 40 | 40 | 16 | 4 | 88 | 12 | 0 | 0 | 84 | 12 | 4 | 0 |
| RL | 5 | 60 | 40 | 0 | 0 | 80 | 10 | 0 | 0 | 100 | 0 | 0 | 0 |
| WL | 46 | 30 | 43 | 20 | 6 | 78 | 17 | 4 | 0 | 93 | 4 | 2 | 0 |
| Total | 84 | 32% | 50% | 13% | 5% | 80% | 11% | 6% | 0% | 73% | 23% | 4% | 0% |

**LFP Evaluation Summary.** The summarized responses, by laboratory, are listed on the following page. LFPs were asked to rate the following questions on a scale from 1 (below average) to 5 (above average).

2. LFPs involved in SRP associate application evaluation process:
   a. Time available for evaluation of applications:
   b. Adequacy of applications for selection process:
3. Value of orientation trips:
4. Length of research tour:
5. a. Benefits of associate's work to laboratory:
   b. Benefits of associate's work to Air Force:
6. a. Enhancement of research qualifications for LFP and staff:
   b. Enhancement of research qualifications for SFRP associate:
   c. Enhancement of research qualifications for GSRP associate:
7. a. Enhancement of knowledge for LFP and staff:
   b. Enhancement of knowledge for SFRP associate:
   c. Enhancement of knowledge for GSRP associate:
8. Value of Air Force and university links:
9. Potential for future collaboration:
10. a. Your working relationship with SFRP:
    b. Your working relationship with GSRP:
11. Expenditure of your time worthwhile:

(Continued on next page)

12. Quality of program literature for associate:
13. a. Quality of RDL's communications with you:
    b. Quality of RDL's communications with associates:
14. Overall assessment of SRP:

Table B-3. Laboratory Focal Point Reponses to above questions

| | *AEDC* | *AL* | *USAFA* | *PL* | *RL* | *WHMC* | *WL* |
|---|---|---|---|---|---|---|---|
| *# Evals Recv'd* | 0 | 7 | 1 | 14 | 5 | 0 | 46 |
| *Question #* | | | | | | | |
| 2 | - | 86 % | 0 % | 88 % | 80 % | - | 85 % |
| 2a | - | 4.3 | n/a | 3.8 | 4.0 | - | 3.6 |
| 2b | - | 4.0 | n/a | 3.9 | 4.5 | - | 4.1 |
| 3 | - | 4.5 | n/a | 4.3 | 4.3 | - | 3.7 |
| 4 | - | 4.1 | 4.0 | 4.1 | 4.2 | - | 3.9 |
| 5a | - | 4.3 | 5.0 | 4.3 | 4.6 | - | 4.4 |
| 5b | - | 4.5 | n/a | 4.2 | 4.6 | - | 4.3 |
| 6a | - | 4.5 | 5.0 | 4.0 | 4.4 | - | 4.3 |
| 6b | - | 4.3 | n/a | 4.1 | 5.0 | - | 4.4 |
| 6c | - | 3.7 | 5.0 | 3.5 | 5.0 | - | 4.3 |
| 7a | - | 4.7 | 5.0 | 4.0 | 4.4 | - | 4.3 |
| 7b | - | 4.3 | n/a | 4.2 | 5.0 | - | 4.4 |
| 7c | - | 4.0 | 5.0 | 3.9 | 5.0 | - | 4.3 |
| 8 | - | 4.6 | 4.0 | 4.5 | 4.6 | - | 4.3 |
| 9 | - | 4.9 | 5.0 | 4.4 | 4.8 | - | 4.2 |
| 10a | - | 5.0 | n/a | 4.6 | 4.6 | - | 4.6 |
| 10b | - | 4.7 | 5.0 | 3.9 | 5.0 | - | 4.4 |
| 11 | - | 4.6 | 5.0 | 4.4 | 4.8 | - | 4.4 |
| 12 | - | 4.0 | 4.0 | 4.0 | 4.2 | - | 3.8 |
| 13a | - | 3.2 | 4.0 | 3.5 | 3.8 | - | 3.4 |
| 13b | - | 3.4 | 4.0 | 3.6 | 4.5 | - | 3.6 |
| 14 | - | 4.4 | 5.0 | 4.4 | 4.8 | - | 4.4 |

## 3. 1997 SFRP & GSRP EVALUATION RESPONSES

The summarized results listed below are from the 257 SFRP/GSRP evaluations received.

Associates were asked to rate the following questions on a scale from 1 (below average) to 5 (above average) - by Air Force base results and over-all results of the 1997 evaluations are listed after the questions.

1. The match between the laboratories research and your field:
2. Your working relationship with your LFP:
3. Enhancement of your academic qualifications:
4. Enhancement of your research qualifications:
5. Lab readiness for you: LFP, task, plan:
6. Lab readiness for you: equipment, supplies, facilities:
7. Lab resources:
8. Lab research and administrative support:
9. Adequacy of brochure and associate handbook:
10. RDL communications with you:
11. Overall payment procedures:
12. Overall assessment of the SRP:
13. a. Would you apply again?
    b. Will you continue this or related research?
14. Was length of your tour satisfactory?
15. Percentage of associates who experienced difficulties in finding housing:
16. Where did you stay during your SRP tour?
    a. At Home:
    b. With Friend:
    c. On Local Economy:
    d. Base Quarters:
17. Value of orientation visit:
    a. Essential:
    b. Convenient:
    c. Not Worth Cost:
    d. Not Used:

SFRP and GSRP associate's responses are listed in tabular format on the following page.

Table B-4. 1997 SFRP & GSRP Associate Responses to SRP Evaluation

| | Arnold | Brooks | Edwards | Eglin | Griffis | Hanscom | Kelly | Kirtland | Lackland | Robins | Tyndall | WPAFB | average |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| # res | 6 | 48 | 6 | 14 | 31 | 19 | 3 | 32 | 1 | 2 | 10 | 85 | 257 |
| 1 | 4.8 | 4.4 | 4.6 | 4.7 | 4.4 | 4.9 | 4.6 | 4.6 | 5.0 | 5.0 | 4.0 | 4.7 | 4.6 |
| 2 | 5.0 | 4.6 | 4.1 | 4.9 | 4.7 | 4.7 | 5.0 | 4.7 | 5.0 | 5.0 | 4.6 | 4.8 | 4.7 |
| 3 | 4.5 | 4.4 | 4.0 | 4.6 | 4.3 | 4.2 | 4.3 | 4.4 | 5.0 | 5.0 | 4.5 | 4.3 | 4.4 |
| 4 | 4.3 | 4.5 | 3.8 | 4.6 | 4.4 | 4.4 | 4.3 | 4.6 | 5.0 | 4.0 | 4.4 | 4.5 | 4.5 |
| 5 | 4.5 | 4.3 | 3.3 | 4.8 | 4.4 | 4.5 | 4.3 | 4.2 | 5.0 | 5.0 | 3.9 | 4.4 | 4.4 |
| 6 | 4.3 | 4.3 | 3.7 | 4.7 | 4.4 | 4.5 | 4.0 | 3.8 | 5.0 | 5.0 | 3.8 | 4.2 | 4.2 |
| 7 | 4.5 | 4.4 | 4.2 | 4.8 | 4.5 | 4.3 | 4.3 | 4.1 | 5.0 | 5.0 | 4.3 | 4.3 | 4.4 |
| 8 | 4.5 | 4.6 | 3.0 | 4.9 | 4.4 | 4.3 | 4.3 | 4.5 | 5.0 | 5.0 | 4.7 | 4.5 | 4.5 |
| 9 | 4.7 | 4.5 | 4.7 | 4.5 | 4.3 | 4.5 | 4.7 | 4.3 | 5.0 | 5.0 | 4.1 | 4.5 | 4.5 |
| 10 | 4.2 | 4.4 | 4.7 | 4.4 | 4.1 | 4.1 | 4.0 | 4.2 | 5.0 | 4.5 | 3.6 | 4.4 | 4.3 |
| 11 | 3.8 | 4.1 | 4.5 | 4.0 | 3.9 | 4.1 | 4.0 | 4.0 | 3.0 | 4.0 | 3.7 | 4.0 | 4.0 |
| 12 | 5.7 | 4.7 | 4.3 | 4.9 | 4.5 | 4.9 | 4.7 | 4.6 | 5.0 | 4.5 | 4.6 | 4.5 | 4.6 |
| Numbers below are percentages | | | | | | | | | | | | | |
| 13a | 83 | 90 | 83 | 93 | 87 | 75 | 100 | 81 | 100 | 100 | 100 | 86 | 87 |
| 13b | 100 | 89 | 83 | 100 | 94 | 98 | 100 | 94 | 100 | 100 | 100 | 94 | 93 |
| 14 | 83 | 96 | 100 | 90 | 87 | 80 | 100 | 92 | 100 | 100 | 70 | 84 | 88 |
| 15 | 17 | 6 | 0 | 33 | 20 | 76 | 33 | 25 | 0 | 100 | 20 | 8 | 39 |
| 16a | - | 26 | 17 | 9 | 38 | 23 | 33 | 4 | - | - | - | 30 | |
| 16b | 100 | 33 | - | 40 | - | 8 | - | - | - | - | 36 | 2 | |
| 16c | - | 41 | 83 | 40 | 62 | 69 | 67 | 96 | 100 | 100 | 64 | 68 | |
| 16d | - | - | - | - | - | - | - | - | - | - | - | 0 | |
| 17a | - | 33 | 100 | 17 | 50 | 14 | 67 | 39 | - | 50 | 40 | 31 | 35 |
| 17b | - | 21 | - | 17 | 10 | 14 | - | 24 | - | 50 | 20 | 16 | 16 |
| 17c | - | - | - | - | 10 | 7 | - | - | - | - | - | 2 | 3 |
| 17d | 100 | 46 | - | 66 | 30 | 69 | 33 | 37 | 100 | - | 40 | 51 | 46 |

## 4. 1997 USAF LABORATORY HSAP MENTOR EVALUATION RESPONSES

Not enough evaluations received (5 total) from Mentors to do useful summary.

## 5. 1997 HSAP EVALUATION RESPONSES

The summarized results listed below are from the 113 HSAP evaluations received.

HSAP apprentices were asked to rate the following questions on a scale from 1 (below average) to 5 (above average)

1. Your influence on selection of topic/type of work.
2. Working relationship with mentor, other lab scientists.
3. Enhancement of your academic qualifications.
4. Technically challenging work.
5. Lab readiness for you: mentor, task, work plan, equipment.
6. Influence on your career.
7. Increased interest in math/science.
8. Lab research & administrative support.
9. Adequacy of RDL's Apprentice Handbook and administrative materials.
10. Responsiveness of RDL communications.
11. Overall payment procedures.
12. Overall assessment of SRP value to you.
13. Would you apply again next year? Yes (92 %)
14. Will you pursue future studies related to this research? Yes (68 %)
15. Was Tour length satisfactory? Yes (82 %)

| | Arnold | Brooks | Edwards | Eglin | Griffiss | Hanscom | Kirtland | Tyndall | WPAFB | Totals |
|---|---|---|---|---|---|---|---|---|---|---|
| # resp | 5 | 19 | 7 | 15 | 13 | 2 | 7 | 5 | 40 | 113 |
| 1 | 2.8 | 3.3 | 3.4 | 3.5 | 3.4 | 4.0 | 3.2 | 3.6 | 3.6 | 3.4 |
| 2 | 4.4 | 4.6 | 4.5 | 4.8 | 4.6 | 4.0 | 4.4 | 4.0 | 4.6 | 4.6 |
| 3 | 4.0 | 4.2 | 4.1 | 4.3 | 4.5 | 5.0 | 4.3 | 4.6 | 4.4 | 4.4 |
| 4 | 3.6 | 3.9 | 4.0 | 4.5 | 4.2 | 5.0 | 4.6 | 3.8 | 4.3 | 4.2 |
| 5 | 4.4 | 4.1 | 3.7 | 4.5 | 4.1 | 3.0 | 3.9 | 3.6 | 3.9 | 4.0 |
| 6 | 3.2 | 3.6 | 3.6 | 4.1 | 3.8 | 5.0 | 3.3 | 3.8 | 3.6 | 3.7 |
| 7 | 2.8 | 4.1 | 4.0 | 3.9 | 3.9 | 5.0 | 3.6 | 4.0 | 4.0 | 3.9 |
| 8 | 3.8 | 4.1 | 4.0 | 4.3 | 4.0 | 4.0 | 4.3 | 3.8 | 4.3 | 4.2 |
| 9 | 4.4 | 3.6 | 4.1 | 4.1 | 3.5 | 4.0 | 3.9 | 4.0 | 3.7 | 3.8 |
| 10 | 4.0 | 3.8 | 4.1 | 3.7 | 4.1 | 4.0 | 3.9 | 2.4 | 3.8 | 3.8 |
| 11 | 4.2 | 4.2 | 3.7 | 3.9 | 3.8 | 3.0 | 3.7 | 2.6 | 3.7 | 3.8 |
| 12 | 4.0 | 4.5 | 4.9 | 4.6 | 4.6 | 5.0 | 4.6 | 4.2 | 4.3 | 4.5 |
| Numbers below are percentages | | | | | | | | | | |
| 13 | 60% | 95% | 100% | 100% | 85% | 100% | 100% | 100% | 90% | 92% |
| 14 | 20% | 80% | 71% | 80% | 54% | 100% | 71% | 80% | 65% | 68% |
| 15 | 100% | 70% | 71% | 100% | 100% | 50% | 86% | 60% | 80% | 82% |

**Associate did not participate in the program.**

# EIKONES-AN OBJECT-ORIENTED LANGUAGE FORIMAGE ANALYSIS AND PROCESSING

**Richard Andel**
**Graduate Student**
**Department of Computer Science**

**Binghamton University (SUNY)**
**Center for Intelligent Systems**
**T.J. Watson Schoo**
**Binghamton, NY 13902l**

**Final Report for:**
**Graduate Student Research Program**
**Rome Laboratory**

**Sponsored by:**
**Air Force Office of Scientific Research**
**Bolling Air Force Base, Washington, DC**

**And**

**Rome Laboratory**

**August 1997**

# EIKONES - AN OBJECT-ORIENTED LANGUAGE FOR IMAGE ANALYSIS AND PROCESSING

**Nikolaos G. Bourbakis**
Professor
Dept. of Electrical Engineering & Dept. of Computer Science
and
**Richard Andel**
Graduate Student
Dept. of Computer Science
Binghamton University (SUNY),
Center for Intelligent Systems,

**Abstract**

This paper presents an object oriented language, called EIKONES, for efficient and flexible image processing. The EIKONES language provides to the user flexibility and friendliness for image processing which are not available in other image processing tools. The basic idea behind EIKONES is the consideration of the image processing algorithms as objects and the appropriate development of a formal grammar for its actual implementation. Results are provided by using the EIKONES language implemented by C++ for a PC environment. The implementation of EIKONES Language has a current size of 3K lines code.

# EIKONES - AN OBJECT-ORIENTED LANGUAGE FOR IMAGE ANALYSIS AND PROCESSING

**Nikolaos G. Bourbakis**
and
**Richard Andel**

## 1. Introduction

Image processing is one of the most interesting and exponentially growing research area. Several image processing libraries and languages have been developed to facilitate the efficient processing of images [1-9]. The libraries are serial based one at a time image processing tools with no flexibility and friendliness for the user. On the other hand, the existing image processing languages, either they are specialized to a certain image processing sub-areas, such accessing [6], graphics [8],etc., or they are based on general purpose programming languages with the ability to manipulate images [9]. There is also a category of more flexible image processing languages, called object oriented image processing languages [2]. In this category, languages, such as OLIVE, HIPS and Eikones, are for machine vision and/or image processing and friendly to the user for portable applications.

In this paper, we present the recent version of the Eikones language. The basic idea behind the Eikones language was introduced in 1985 [1] in a hardware form, since the PC computing power at that time for real-time applications was very limited. In 1986, the first software version of Eikones was developed with a limited set of image processing methods. The recent version of the language includes many image processing methods classified in a small number of major categories. In addition, it combines object oriented, programming, and library capabilities. Its main structure is based on a simple context-free grammar.

## 2. The Eikones Language

### 2.1. Definitions

Definition 1: The Eikones language is generated by a context free grammar G = (Ns, Ts, PR, S), where,

Ns = {S, T, L, K, I, O} U {operations} is the set of non-terminal symbols,

Ts = 3 U {i/i0Z} U {e} U {@, #} is the set of terminal symbols,

PR is the following set of production rules

S ---> T

S ---> ToS

T ---> Lk

T ---> e (empty)

L ---> /qL1(**a**)/qL2(**a**)/.../qLt(**a**)/

o ---> # (synthesis operator)

q ---> @ saving operator

k ---> $i_k j_m$

**a** ---> /(x1,y1),(x2,y2)/(x1,y1),(x2,y2),degree/(x1,y1),(x2,y2),values/(x1,y1),
(x2,y2),patterns/(x1,y1),(x2,y2),level/threshold(s)/image(size)/templates/
pattems/color/

S is the starting symbol

The language generated by the grammar G is

$$\mathbf{L(G)} = \{Wt/\ Wt=qLi_1j_1(\mathbf{a}i_1j_1)\#qLi_2j_2(\mathbf{a}i_2j_2)\#...\#qLi_tj_t(\mathbf{a}i_tj_t),\ Lij03,\ t0Z\}$$

where,

. Wt represents words of the language,

. t represents the length of a word Wt,

. $i_k$ represents the category in which a letter Likjm belongs to,

. $j_m$ represents the rank of the letter inside the category,

. $ai_kj_m$ represents the specs and the parameters required by each letter (image processing method) to perform a particular task on an image selected by the user or on an image output. **a** is specifically defined in table 1.

### 2.2. The Alphabet

The alphabet of the Eikones language is defined by the selection of the methods used for image processing. For instance, invert, binarization, histogram, identity, reverse, thresholding, multiplication, addition, subtraction, division, logical (AND, OR, NOT, XOR...), window, convolution, low pass filter, high pass filter, dilasion, erosion, skeletonization, transformation, etc. The image processing methods are classified in a small number of categories to provide flexibility to the parsing process and to the user as well. In particular, the categories used here are described in the table 1, which includes a sample of the methods used by the language.

**TABLE 1.**

| Letter | Symbol | Category | Rank | Specs |
|---|---|---|---|---|
| L00 | WIN | Windows | 1 = create a window | coordinates (x1,y 1/x2,y2) |
| | | | 2 = rotate a window | coordinates (x1,y 1/x2,y2),degrees |
| | | | 3 = delete a window | coordinates (x1,y 1/x2,y2),value |
| | | | 4 = insert a window | coordinates (x1,y 1/x2,y2),pattern |
| | | | 5 = zoom a window | coordinates (x1,y 1/x2,y2),level |
| | | | 6 = focus a window | coordinates (x1,y 1/x2,y2),level |
| L10 | MON | Monadic | 0 = copy (equality) | |
| | | | 1 = inverse | |
| | | | 2 = binarize | threshold |

| | | | | |
|---|---|---|---|---|
| | | | 3 = threshold | threshold pairs |
| | | | 4 = stretch | threshold pairs |
| L20 | DYA | Dyadic | 1 = add | image |
| | | | 2 = subtract | image |
| | | | 3 = multiply | image |
| | | | 4 = divide | image |
| | | | 5 = logical and | image |
| | | | 6 = logical or | image |
| | | | 7 = logical xor | image |
| | | | 8 = logical nand | image |
| L3x | FLT | Filters | | |
| | FLT_HP | High-Pass | 1 = x | template |
| | FLT_LP | Low-Pass | 1 = x | template |
| | COV | Convolution | 1 = x | template |
| L4x | MFL | Morpho | | |
| | MFL_ER | | 1 = erosion | pattern |
| | MFL_DL | | 2 = dilision | pattern |
| | MFL_OP | | 3 = opening | pattern |
| | MFL_CL | | 4 = closing | pattern |
| | HST | | 1 = histogram | |
| | ENH | | 2 = enhancement | |
| | SEG | | 1 = x | template |
| | EGD | | 1 = x | template |

| | | | |
|---|---|---|---|
| SKE | | 1 = x | templates |
| L50 | TRA | Transforms | 1 = FFT |
| | | | 2 = Walsh |
| | | | 3 = Hadamard |
| | | | 4 = Haar |
| | | | 5 = DCT |
| | | | 6 = SLT |
| | | | 7 = KLT |

Thus, the language's alphabet

$$3 = \{Li_kj_m / Li_kj_m \text{ is an image processing method, } i,j,k,m \in Z\}$$

The words present two forms. The first form is

$$Wt = qLi_1j_1(ai_1j_1)\#...\#qLi_{j_t}(ai_{j_t}),$$

which represents the sequence of image processing tasks to be perfomed, and

$$Wt[P] = qLi_1j_1(ai_1j_1)\#...\#qLi_{j_t}(ai_{j_t})\ [P],$$

which represents the image processed by the sequence

## 3. Implementation of Eikones

The Eikones language was implemented by C++ for a PC environment. The basic architectural scheme is shown in figure 1.

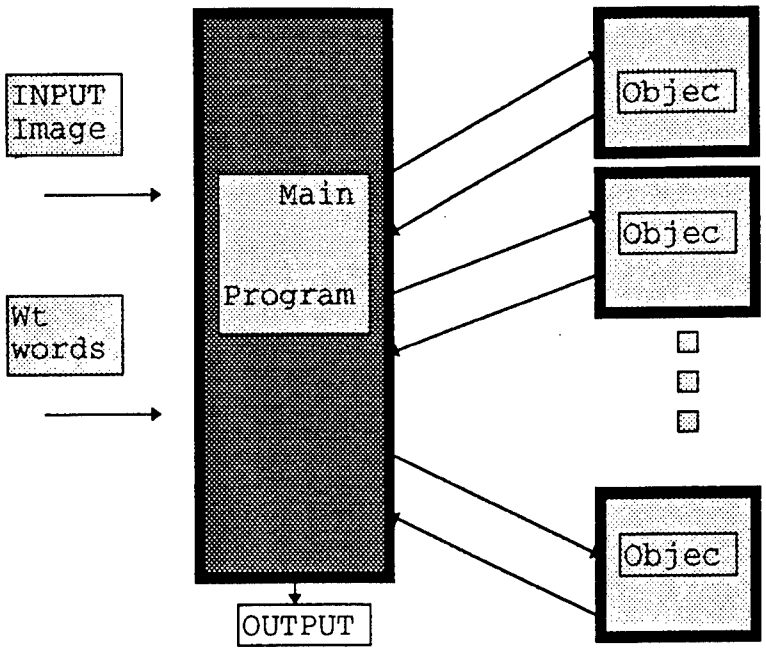



Fig.1:Global Configuration of the EIKONES Language architecture

The Objects of the language are the image processing methods. Each object has its own specs, such as {input (nxn), type of input (color, or grey, or binary), output (mxm), type of output, etc.}

## 4. Illustrative Examples

In this section we present two illustrative examples to demonstrate some of the language's capabilities.

**. Example -1**

The language word is Wt = MON(2)#WIN(2)#MON(1)

which means inverse the original image, rotate it, and binarize it. Thus, the results from each process are:

The original image

The Inverted Image

The Rotated Image

The Binarized Image

**. Example -2**

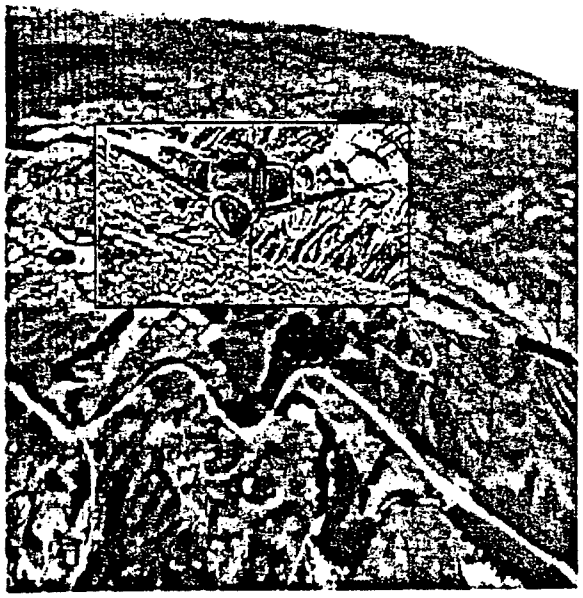
The same original image was used. The language word selected was

Wt1 = FLT-HP#WIN(2)#WIN(1)

Wt2= WIN(4)

It means, that, create a window from the original image, rotate it, and apply high pass filter on it.

Then the result insert it into the original image.



The final result

## 5. Conclusion

In this paper an object oriented image processing language implemented by C++ in a PC environment was presented. The code for the implementation of Eikones is 3K lines. The basic idea behind this language is to develop an efficient, flexible and user's friendly tool for image processing. The language presented here, combines these features. Extension of this work is the development additional features for the language, such as statistical measurments for image processing, learning capabilities, and natural language interfaced.

**. Acknowledgment**

This work is partially supported by an AIS grant 1996, and summer research at AIRFORSE-IRR-RL 1997. The authors wish to express many thanks to all the research personel at the IRR division and especial to A.Hall and S.Faar for their support.

## References

[1] N.Bourbakis et.al. High performance architectures for real-time, multilevel picture information systems, Proc. IEEE Workshop on LFA, JUne 1985.

[2] M.Lavin and M.Flickner, An object oriented language for machine vison and image processing, Proc. IEEE Workshop on TAI, 1989, VA.

[3] 3m Corp., Image processing tool, TR-1988, VA

[4] B.Batchelor et.al.,Command language for interactive image analysis,IEE Proc. vol.127,5,1980

[5] M.Bohm et. al,A simple dialog language for the processing of image sequence, IEEE Symp. on Medical Imaging & Image Interpretation, Berlin 1982

[6] N.Bourbakis, SCAN A language for efficient accessing of 2-D arrays, Proc. IEEE Workshop on LFA, Singapore 1986

[7] H.Tamura et al.,Design and implementation of SPIDER: a transportable image, CVGIP, 23, 1982

[8] L.Guibas and J.Stolfi, A language for bitmap manipulation, ACM T-Graphics, 1,3,1982

[9] J.Wilson et al.,The image algebra Fortran language, TR-1988-03, UFL

[10] M.Landy, et. al.,"HIPS-A UNIX based image processing system, CVGIP, 25,331-347,1984

# AN EXAMINATION OF JAVA AND CORBA SECURITY

**Pat Garrity**
**Graduate Student**
**Department of Computer Science**

**Central Michigan University**
**Pearce Hall**
**Mt. Pleasant, MI 48859**

**Final Report for:**
**Graduate Student Research Program**
**Rome Laboratory**

**Sponsored by:**
**Air Force Office of Scientific Research**
**Bolling Air Force Base, Washington, DC**

**And**

**Rome Laboratory**

**August 1997**

# An Examination of Java and CORBA Security

**Pat Garrity**
Graduate Student

and

**Michael Stinson**
Professor

Department of Computer Science
Pearce Hall
Central Michigan University
Mt. Pleasant, MI 48859

Abstract

*As the power of computer networking continues to expand, a greater effort is being made to harness that power. Two areas of intense activity are executable content, which refers to code which is automatically downloaded and run on a user's machine over a network, and distributed computing, which refers to the distribution over a network of computer tools and applications. In this paper, we look at the security measures associated with two of the approaches taken to distributed computing and executable content. First, in the case of distributed computing, we examine the built-in security associated with Java™ applets. Second, in the case of executable content, we look at the security measure associated with the CORBA™ standard for object request brokers.*

# An Examination of Java and CORBA Security

**Pat Garrity**
Graduate Student

and

**Michael Stinson**
Professor

## Introduction

Network security is a concern that can be approached in many different ways. The trend towards object-oriented programming and object based technology suggests that we need to study destructive behavior at the object level. (The reader is referred to our accompanying paper on destructive objects.)

Objects in cyberspace have a greater potential for destructive behavior than objects on an isolated system because of their mobility, by which we mean movement across a network. Two primary ways in which objects move through cyberspace are via executable content and distributed computing.

Executable content refers to the ability of cyberspace documents such as web pages to contain embedded programs that are automatically downloaded and run on a user's machine. There are several mechanisms available for executable content, the most ubiquitous of which are Java™ applets. The danger of automatically downloaded code is that it can originate anywhere, perhaps from someone with malicious intent. While most network security is based in some way on trusted code, Java seeks to ensure the safety of even untrusted code. In the first part of this paper we will examine how Java does this.

Distributed computing refers to the use of a network to divide a process among several objects, perhaps residing on different machines or even contained on different networks. Distributed computing allows for a document centered rather than application centered approach to a task. The centerpiece for the user in a distributed computing environment will be the document being working on - not the application the user needs to create the document. Any tools needed to work on the document will be distributed over the LAN (local area network) or WAN (wide area network) to which the user is connected. Distributed computing is inherently object-oriented. Code from programming languages that are not object based is forced to simulate an object-oriented approach. The danger from distributed computing is that in order for it to be worthwhile, objects foreign to a user's machine must have access to sensitive parts of the it, such as the file and/or operating systems.

In this work, we will look at examples of how security is being addressed in these two areas of network computing. In the case of executable content we will examine how the Java programming language developed by Sun Microsystems tries to guarantee safety for the user, and in the case of distributed computing will we see how the Common Object Request Broker Architecture (CORBA™) standard produced by the Object Management Group (OMG) supplies security services to protect its users.

It should be noted that neither Java™ nor CORBA™ are the only available tools in their respective domains. Other tools, such as COM™ and ActiveX™, while providing a different approach to distributed computing face many of the same security concerns outlined in this work. Therefore, an examination of Java's and CORBA's approaches to security is valuable to the broader field of managing destructive objects.

## Java™

Java™ is a programming language that was developed by Sun Microsystems with executable content in mind. Its developers built security into the language in two ways:

- As a stand-alone language Java incorporates strong inherent safety features such as strong type checking and no user-accessible pointers [MF97a]. This is partially a reaction to languages such as C (from which Java gets much of its syntax). C allowed so much user freedom that it is virtually an assembly language. Pointer arithmetic and the ability to cast pointers from one type of data structure to another affords a large potential for abuse by malicious programmers. Java attempts to avoid this by allowing the programmer far less freedom at the memory-space level.
- In addition, Java has built-in security designed specifically for safe executable content. While Java is a full-featured programming language in which one can write complete applications, it is Java applets which are designed to make it the mechanism of choice for executable content. Applets are (usually) small programs that are tightly controlled in the actions they may perform on a user's machine. For this reason the Java security model is often referred to as the sandbox, since applets are only allowed to act within the confines of a virtual sandbox which restricts the destruction they can cause [MF97b]. These restrictions also have

---

™ *Java* is a registered trademark of *Sun Microsystems, Inc.*
™ *CORBA* is a trademark of the *Object Management Group, Inc.*
™*Com* and *ActiveX* are trademarks of *Microsoft, Inc.*

the effect, however, of making Java™ less powerful than desired when it comes to distributed computing.

## The Java Security Model

There are three main parts to the Java security model: the Bytecode Verifier, the Class Loader, and the Security manager. This three-way breakdown leads some sources to refer to it as a three-layered model, but this terminology gives a false impression of redundancy. As McGraw states: "Though such a label is commonly encountered in the Java security literature, it is misleading. ... if any of the three prongs is exploited, the entire security system breaks." Following his lead we will refer to it as a three-pronged model [MF97a].

Figure 1 below illustrates how the security model fits in the Java universe. Java is an interpreted language, whose source code is compiled into bytecode that will run on any machine that has a Java execution environment. A standard Java execution environment consists of the Java Virtual Machine (VM) that interprets Java bytecode and translates it into machine specific instructions, along with base classes that provide multi-platform primitives. This platform independence is one of the main features of Java™ [Fla97]

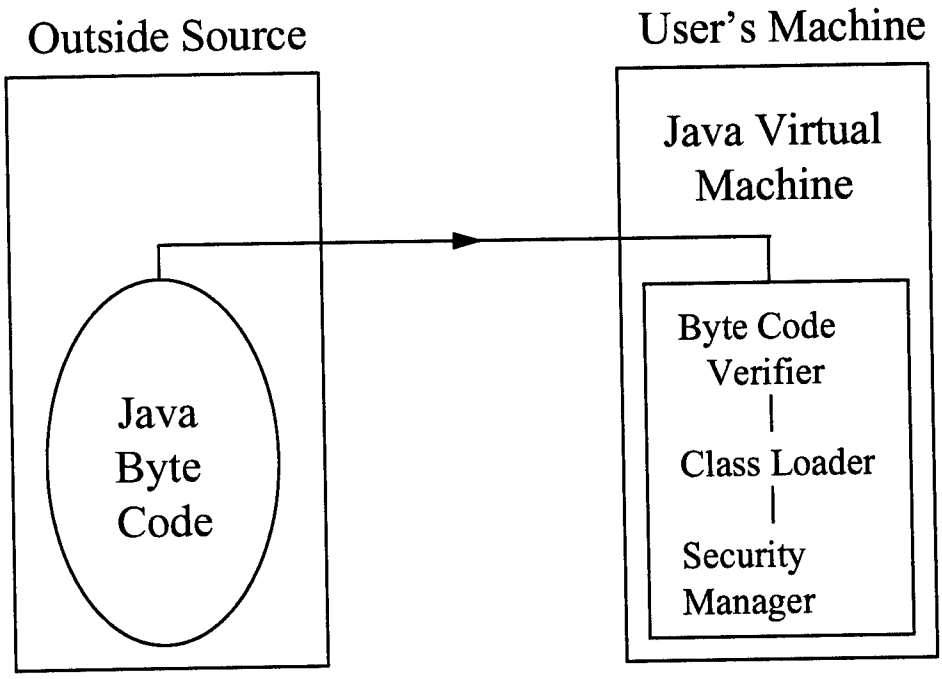



Fig. 1 – How Java™ handles code from an outside source.

The VM is where the three prongs of the security model reside. When a user accesses a web page which contains an applet, the applet's bytecode is downloaded and run on the user's machine. Each part of the model plays a role in making sure the bytecode that's imported does no harm to the user's machine. First the applet class loader places the applet's code in its own namespace. Then the bytecode verifier examines the bytecode's format. If any line of bytecode is not in an approved form, or appears to perform an illegal operation, the applet will not run. While the applet runs the security manager is given veto power over any method* considered dangerous, while the class loader makes sure that any class called by the applet will be put in the appropriate namespace. We will examine each prong of the model in turn.

Bytecode Verifier - The most basic function of the bytecode verifier is to check the format of the bytecode to ensure that that it conforms to Java's specifications. After all, there is a chance the code was generated by an malicious compiler. Anyone who cares to could easily write bytecode that would break the VM if the verifier did not check that each bytecode command was acceptable. In addition, the verifier uses a basic theorem prover to ensure that the bytecode doesn't circumvent the inherent safety features of the language. The theorem checker ensures that the bytecode doesn't overflow or underflow the stack, use registers incorrectly, convert data types illegally, or allow access to restricted classes , all of which helps insure that the code will not cause memory space violations [Ven97].

Java™ is not the first language to use bytecode to attempt to achieve platform independence. (One previous attempt was p-code which was developed at the University of California, San Diego in the mid '70s). One of the problems encountered in previous tries however, was that the code ran too slow. More powerful computers have helped a great deal, and the developers of Java™ have tried to maximize execution speed, but a large applet can still bog down the bytecode verifier enough to amount to a denial of service attack [MF97a].

Class Loader - The second prong in the Java™ security model is the class loader. Java™ is an object-oriented language, and so everything is a class except for primitive data types. When a class is added to the Java run-time environment it is loaded into a namespace. Which namespace a particular class gets loaded into depends on its origin relative to the user's machine. The namespace the class resides in determines the access it has to other parts of the run-time environment. Namespaces also useful prevent name collisions, since the programmer writing the applet needn't worry about duplicating names in other classes. Each class in the run-time environment is tagged with the class loader that installed it. Figure 2 shows how namespaces are split up on a machine on a local network [MF97a].

---

* A *method* is the object-oriented term for a procedure or function.

Objects in the local class namespace include those built in to the Java language and those originating on the user's machine. Any other classes loaded would be coming over the net as applets or classes referenced by applets. It is the job of the Applet Class Loader to determine which name space these classes enter. This is important, since an applet can only see the classes in its own namespace, which helps prevent a malicious applet from convincing the VM that it is actually a different class, presumably one with greater access privileges. It also prevents an applet from replacing standard Java™ classes with its own versions [MF97a].

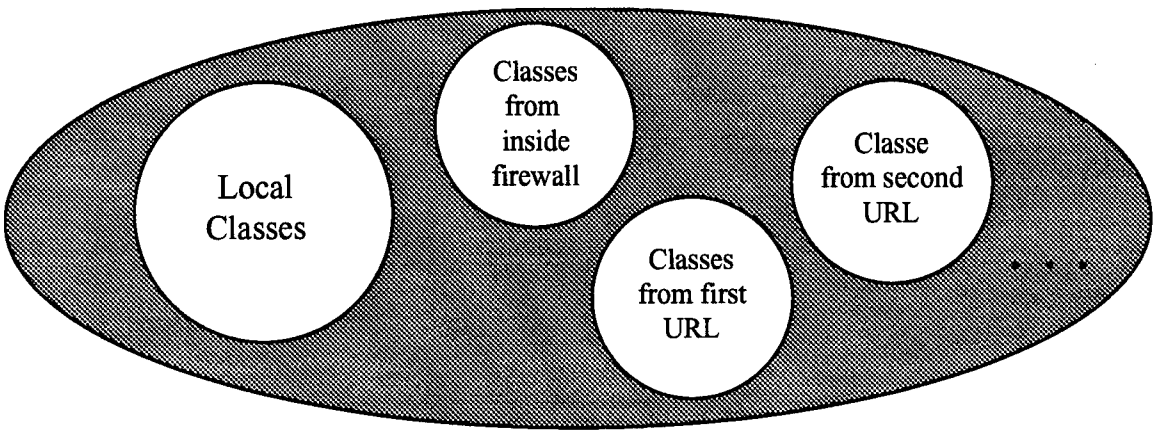



Fig. 2 – Classes are kept distinct from each other on the basis of origin.

An applet is not allowed to install a new class loader. The Applet Class Loader is the applet's only choice for loading new classes. An application, however, can set up its own class loader (or class loaders), customized to suit its needs. The developer who wishes to do this must be very careful. If the customized loader they use is not correctly written, they could open a security hole to hostile applets coming over the network [MF97a].

Security Manager - The security manager module was added to Java with the intent of providing safety for users who would find their machines running untrusted code. As such it forms a large part of the Java security model.

The security manager performs run time checks on dangerous methods. It is consulted whenever a potentially hazardous operation is attempted. Some of the operations which the security manager routinely checks include [MF97a]:

- all access requests, including access to threads, the operating system, the file system, the network (via sockets), and other Java components
- the installation of new class loaders
- the creation of operating system level programs

- the interaction of threads and thread groups.

It should be clear that any method seeking to carry out one of the above operations poses a danger to the user's machine. It is not clear, however, that the above list is exhaustive. In fact, the security manager is allowed to veto any operation referred to it. Like the class loader, the security manager is customizable. Application developers can design a security manger that checks any method they consider a threat. Applets, of course, are forbidden from instantiating a security manager. The Java™ virtual machine comes with a default security manager which checks all of the above operations. If you access a web page and an applet downloads, it is the security manager of the VM residing in your web browser that controls what the applet can do. The security manager takes into consideration the origin of the applet, as indicated by the class loader that installed it. Both major browsers (Netscape™ and Microsoft Explorer™) are currently using the default security manager supplied with the Java™ VM, but that could change in the future.

The security manager vetoes an operation by throwing a security exception. Java has several types of exception objects, all of which are instances of some subclass of the class throwable. Although it is not technically correct, we will follow standard usage and refer to them all as exceptions. A Java.lang.Error exception cannot usually be recovered from. A Java.lang.Exception exception can usually be recovered from, providing the developer has written code to catch the exception. A Java.lang.SecurityException exception is used to indicate an operation not permitted for security reasons. It cannot be caught and should always halt execution of the program [Fla97].

## Signature Checking in Java™

Java's idea that any code could be allowed to run on a user's machine if it were tightly controlled stands in contrast to the approach chosen by others, such as ActiveX™ [Sec97]. Security in ActiveX™ relies entirely on signature checking. Code is considered safe because it carries a digital signature that confirms its origin, and the user can take care to run code from a trusted origin. There are some problems with relying solely on signature checking, however. The user may make a poor decision about who to trust, or he may know nothing about the source of the code that is running on his machine. The latter case is exactly the situation involved in a large part of web browsing. Still, signature checking does provide additional security, and the developers of Java decided to add it to their product.

---

™ *Netscape* is a trademark of *Netscape, Inc.*

By using a JAR file it is now possible to add a digitized signature to an applet. A JAR (Java ARchive) file is a single compressed file that contains all the classes an applet needs to run. (Besides allowing for signature checking, JAR files speed up applets since all of the applet's classes come from a single source.) The browser or other application running the VM on the user's machine may then grant the applet special privileges, such as file system access, if it recognizes the signature. Of course, the application running the VM must have some way of recognizing the digital signature. It may have a database of signatures to which it can refer, or it is possible to use public key encryption to decode the JAR file's signature [Ber97]

It should be noted that signature checking in Java only verifies the origin of the applet and ensures it has not been altered en-route, it does not protect the user's system from invasion. It should also be noted, however, that although the applet may have more privileges, the Java security manager must still approve all dangerous operations, and the security manager does not have to give the applet access to the user's entire system.

## Distributed Computing and the Java™ Security Model

Although a Java applet is an excellent tool for executable content, the Java security model has some limitations. In order to do true distributed computing, applications not residing on the user's machine must be given access to sensitive parts of the user's system. While signature checking in Java™ allows this type of access somewhat, additional security assurances are desirable. Although a user may be sure of an applet's origin with signature checking, the user must decide on his own whether that origin is trustworthy, and he can't be sure that a third party hasn't tampered with the applet after it left that origin. In addition, it is desirable in distributed computing to ensure that the object providing a service over the network does not actually decrease the efficiency of the user's work environment. Such a decrease in efficiency will often take the form of a denial of service attack. Some ways in which denial of services attack may work include: completely filling a file system, allocating all of a system's memory, or using all of a machine's CPU cycles. In Java™ applets that will not cause permanent damage, but are merely troublesome at the time they strike are not dealt with [MF97a]. This means that denial of service attacks are not guarded against. An applet may play by all the rules of the security model and still consume enough resources or irritate the user enough to force them to shut down their browser. Under certain circumstances this may be more than a mere nuisance, it may cause the user to lose a great deal of work. It is also possible to design applets which kill other applets, which could also make a user reluctant to trust important work to an applet.

## CORBA™

While executable content in the form of applets on a web pages is fine for small tasks, it is not really distributed computing. True distributed computing involves a major change in the way client/server interactions are viewed. As figure 3 illustrates, distributed computing means doing away with the monolithic server that sits apart from the users of network resources. Instead, the services are distributed among the machines on the network. The distinction between client and server is blurred, and this presents new network administration problems [Adl95].

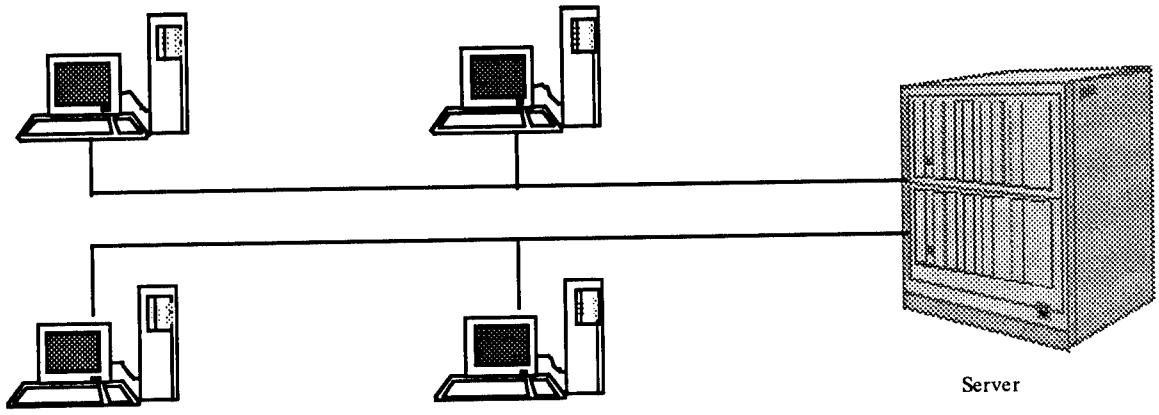



Fig. 3a - Traditional Client/Server System

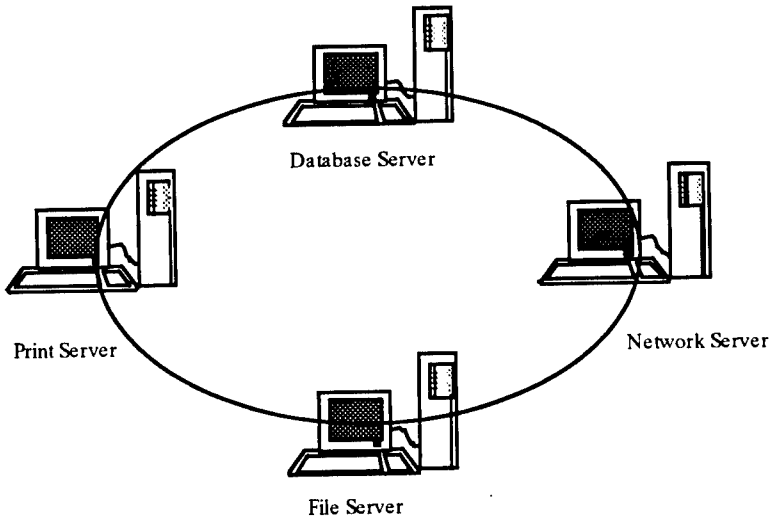



Fig. 3b - Distributed Client/Server System

Distributed computing requires something as powerful as an object request broker (ORB) to manage client/server interactions and insure network security. An ORB is the mechanism by which objects scattered across a network can communicate and interact, and the Object Management Group's (OMG) Common Request Broker Architecture (CORBA™) is rapidly becoming the standard that ORBs must adhere to in order to function effectively. Among other things, a CORBA-compliant ORB provides the security services necessary for distributed computing [Lew95].

## ORB Security Problems

Figure 3b. shows how services are spread out in a distributed computing environment. It is important to note that while one's individual machine may act as both client and server, at the ORB level individual objects are filling that dual role. The objects interact with a CRBA-compliant ORB via OMG's Interface Definition Language (IDL). This allows objects written in different languages to communicate.

In a traditional client/server setup the server is usually trusted while the client is not. For example, a client may assume that his database server (which is maintained by the network administrator) has no malicious intent, but that administrator is not going to assume the same thing about anyone who happens to log on to his server. In distributed computing the database server may be anyone who has registered with the ORB as such. If not for the protection afforded by the ORB, the client would be foolish to assume this object/server is trustworthy [OHE97].

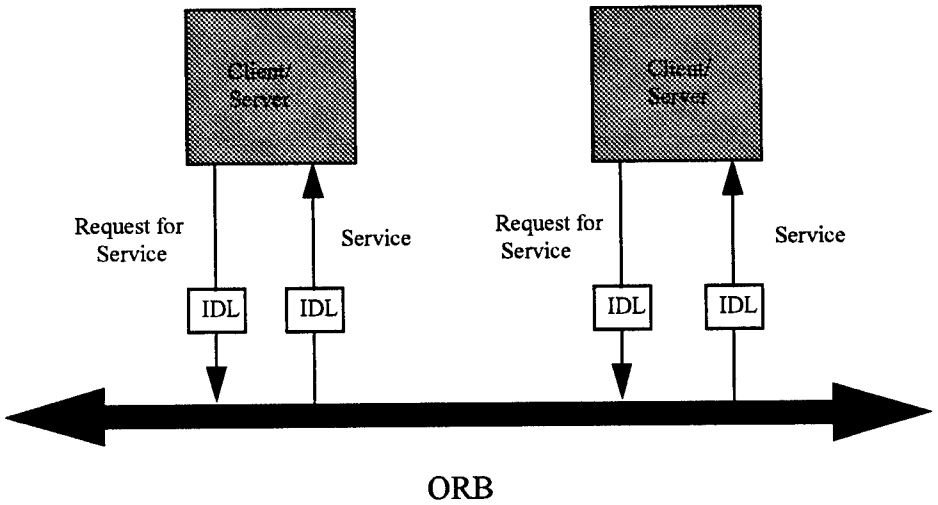



Fig. 4 - An ORB mediates interactions between client/servers on a distributed system.

Figure 4 illustrates how an ORB mediates client/server interactions. Any object offering a service may register with the ORB. A client wanting that service puts in a request to the ORB, which handles the transaction between the two objects. Note that an object registered as a server may also act as a client.

Further security difficulties are presented by the fact that a distributed computing environment is inherently anarchistic. A client's call for service may end up involving several objects. Distributed objects evolve over time, and the aggregate of objects that combined to supply a service today may be different tomorrow, without the client's being aware of the change [OHE97]. For these reasons, the ORB must offer strong security services.

## CORBA Security Services

How does CORBA's security model differ from that offered by Java™? One analogy is that of the security for a bank. Java security is akin to that afforded to the bank by an automated teller machine. Allowing customers to use an ATM does not put the assets inside the bank at risk. The trade-off is that the customer is greatly restricted in the banking transactions he can complete. CORBA security, on the other hand, is closer to that provided by an armed guard (one with extraordinary powers, however, as we will see). Anyone is allowed to come into the bank and interact, and the guard insures that there will be no foul play. CORBA offers a slate of security services designed to insure that objects interacting over a CORBA-compliant ORB behave themselves.

Authentication - An object that wishes to use an ORB must first establish its identity. When a user logs on to the ORB for the first time they are given an authenticated ID which spells out what privileges they have. A principal user need only do this once. Afterward, that user or an object acting on their behalf can utilize the ORB services by showing their digitized signature [OHE97].

Credentials - When a principal user logs on they are given credentials which contain the user's ID and spell out the privileges they have on the ORB. These privileges are based in part on the objects identity, group affiliation (if any), and security clearance. The authenticator actually returns a credentials object, which serves as the user's security ticket that must be shown whenever the user wishes to perform an ORB-related operation.

Delegation - As mentioned in the section on ORB security problems, services in a distributed computing environment are often provide by aggregates rather than single objects. An object that is requested to perform a service may call on other objects to supply part of that service. This chain of

object calls must be regulated by the ORB. In particular, the ORB must decide whose credentials are being used. The transfer of credentials from one object to another is called delegation, and CORBA allows this to be done three different ways [OHE97]:

- *no delegation* means an intermediate object inherits no credentials from the object that called it
- *simple delegation* means an intermediate object receives the calling object's credentials
- *composite delegation* means the intermediate object will have both its own and the calling object's credentials.

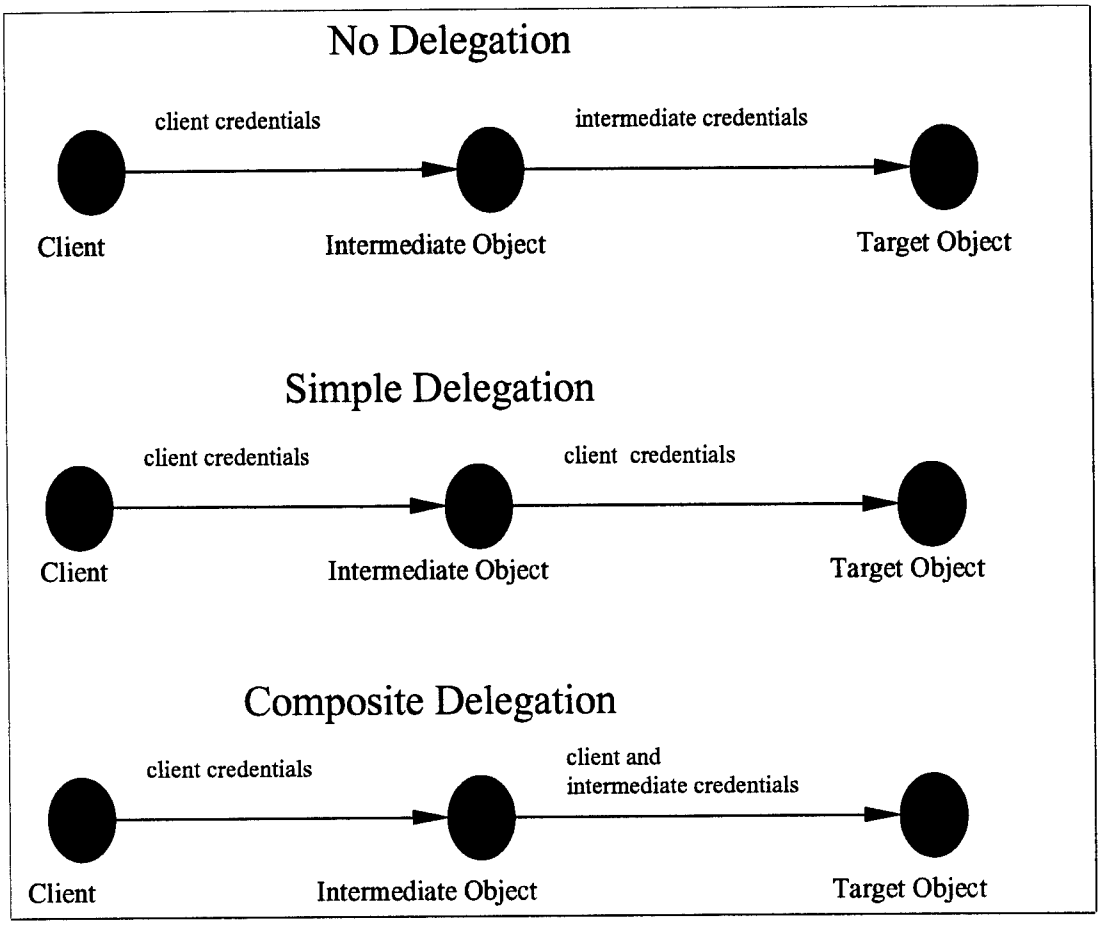



Fig. 5 - There are three ways credentials can be delegated.

Audit Services - The way in which CORBA™ handles authorization and credentials requires a CORBA-compliant ORB to keep track of at least some ORB-object interactions. In fact, CORBA's audit services offer much more in the way of record keeping. CORBA™ will record and store all ORB

interactions, including attempted log-on. Events can be audited by object or object type, operation, time, principal attributes, or success or failure [OHE97].

Non-repudiation - The detailed records kept by CORBA's audit services are necessary for CORBA to offer non-repudiation. This means providing irrefutable evidence that two parties were involved in a client/server interaction. Neither party should be able to deny this evidence in a court of law. CORBA™ attempts to ensure this by offering the following services for any message that passes between two objects via the ORB [OHE97]:

- *evidence of message creation* can be used to prove that the originator created the message
- *evidence of message receipt* can be used to prove that the message was delivered
- an *action timestamp* records the date and time an event took place
- an *evidence long-term storage facility* is provided to store transaction records
- an *adjudicator* is provided to settle disputes based on stored evidence.

Encryption and Checksums - Messages between objects can be encrypted to ensure a private conversation, or a cryptographic checksum can be appended to a message to ensure that it was not tampered with en-route.

Security Domains - According to CORBA™, "a *security domain* is a set of objects to which a security policy applies for a set of security-related activities administered by a *security authority*". The security policies define the rules for access control, authentication, privilege delegation, non-repudiation, and auditability. Allowing objects to group themselves into trusted domains helps keep down the overhead of running an ORB [OHE97].

Security Interfaces - While the majority of objects will be oblivious to the security services being performed by the ORB, it is possible for an object to be security-aware. Such an object can request certain security services from the ORB. A security-aware object may specify a particular type of credential delegation, request specific non-repudiation services, or enforce its own audit policies. This all depends of course, on the object having the right credentials in the first place [OHE97].

## Java™ and CORBA™ Interaction

How do CORBA and Java (or more generally, executable content and distributed computing) fit together on the web? The brass ring for network managers these days appears to be distributed computing. This would seem to indicate that ORBs are the wave the future. While this is probably true, it doesn't mean that ORBs are the answer to everything. Home users want their web browsers, and don't want to have to worry about CORBA-compliance. Java, meanwhile, has a tight relationship with web browsers, and Java applets run with no effort on the part of the user [Res97].

One possible scenario has Java applets acting as the entry point to ORBs. In this scheme the user (who may be on a local network) would click on an applet's icon to request a service. The applet would then request the service of the ORB. At that point, all of CORBA's security would come into play. In particular, the applet would have to be recognized by the ORB. This would allow a nice combination of web browsing along with the possibility of powerful distributed computing [OHE97.

## Future Work

The methods and standards involved in sharing resources across a network is one of the fastest growing fields in computer science. In the area of executable content, more work needs to be done in the area of security against applets from untrusted sources. Research needs to be done on ways to combat applets that consume enough resources or annoy a user enough to force him to shut down his network connection. A sort of "guardian" applet that could detect and kill malicious applets would be very useful

In the field of distributed computing, ways should be sought to make ORBs more ubiquitous. Having one standard such as CORBA™ that is readily accepted is a giant step in this direction. If ORBs become as universal as web browsers, they could be used to fill many roles. For example, an ORB might ensure to parents the suitability of material for children from any site registered with that ORB. For system managers, an ORB could guarantee the security of objects downloaded through that ORB. Research should be done on ways to empower an ORB to automatically detect and restrict the behavior of destructive objects.

*Acknowledgements*

The authors would like to thank everyone at Rome Labs for their help in preparing this paper, especially:

Dr. Heather Dussault, Head of the Information Institute, for her guidance and accurate answers to many questions,

Mr. Joseph Giordano, Head of the Information Warfare Team, for appropriate suggestions and ideas,

Mr. Chris Harrison, Network Manager, for making our job easier with helpful advice and patient solutions to myriad problems.

## References


[Adl95] R. M. Adler, "Distributed Coordination Models for Client/Server Computing", *Computer,* pp. 14 - 22, April, 1995.

[Ber97] C. Berg, "How Do I Create a Signed Applet", *Dr. Dobb's Journal,* pp. 109 – 111, August, 1997.

[Fla97] D. Flanagan, *Java In A Nutshell,* Cambridge: O'Reilly, 1997.

[Lew95] T. G. Lewis, "Where Is Client/Server Software Headed?", *Computer,* pp. 49 - 55, April, 1995.

[MF97a] G. McGraw and E. W. Felten, *Java Security,* New York: Wiley, 1997.

[MF97b] G. McGraw and E. W. Felten, "Understanding the keys to Java security -- the sandbox and authentication", *JavaWorld,* May, 1997.

[OHE9] R. Orfali, D. Harkey, and J. Edwards, *Instant CORBA,* New York: Wiley, 1997.

[Res97] R. I. Resnick, (2 July, 1997), "Bringing Distributed Objects to the World Wide Web", [WWW Document], URL *http://www.interlog.com/~resnick/ron.html*

[Sec97] Secure Internet Programming Team, (28 April, 1997), Princeton University Department of Computer Science, "Security Tradeoffs: Java vs. ActiveX", [WWW Document], URL *http://www.cs.princeton.edu/sip/java-vs-activex.html*

[Ven97] B. Venners, "Java's Security Architecture", *JavaWorld,* August, 1997.

**Associate did not participate in the program.**

# OPERATIONAL ANALYSIS OF AN ACTIVELY MODE-LOCKED FIBER LASER

Walter Kaechele
Graduate Student
Physics Department

Rensselaer Polytechnic Institute
110 8th Street
Troy, NY 12180

Final Report for:
Graduate Student Research Programs
Rome Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Rome Laboratory

December 1997

# OPERATIONAL ANALYSIS OF AN ACTIVELY MODE-LOCKED FIBER LASER

Walter Kaechele
Graduate Student
Physics Department
Rensselaer Polytechnic Institute

**Abstract**

This paper describes the construction and operation of an actively mode-locked fiber ring laser. The inclusion of a lithium niobate Mach-Zehnder modulator into a fiber ring cavity permits self-starting and external control of the repetition rates of the laser. The modulator and a sythesized frequency generator mode-locked the laser at frequencies ranging from 1.5 MHz to as high as 5 GHz. An experimental and theoretical analysis of the pulse shaping mechanisms is included as is an examination of the output noise at both low and high repetition rates. The paper concludes with an investigation of rational harmonic mode-locking, a novel method of increasing repetition rates while applying only modest drive signals.

# OPERATIONAL ANALYSIS OF AN ACTIVELY MODE-LOCKED FIBER LASER

Walter Kaechele

## 1 Introduction

As the name implies, active mode-locking occurs when some form of external drive (either optical or electrical) produces a modulation within the cavity which causes the phases of the longitudinal modes to lock together. Typically the modulation derives from some bulk modulator placed within the cavity which is driven by an electronic frequency synthesizer. The principal electrically driven modulators that have been used are the acousto-optic modulator, due to its ease of operation and ready availability, and the fiber pigtailed lithium niobate ($LiNbO_3$) optic modulator closely compatible with the fiber nature of the laser. The lithium niobate modulator is particularly useful since it provides a completely integrated package without resorting to bulk optical components. The synchronization of the pulse train to an electrical drive allows the laser to be referenced to an external phenomena. This has particular advantages when the laser pulse train must be modulated by an electrical signal as is the case with digital communications

Active mode-locking of erbium doped fiber lasers has been an area of intense interest since 1989 when it was realized that picosecond pulses could be obtained[1]. Operating as an actively mode-locked laser, the laser generated 50 ps pulses with sufficient bandwidth to support 2 ps pulses at a repetition rate of 90 MHz. The addition of 2 km of standard telecommunications fiber allowed the pulses to access the additional soliton pulse shaping mechanisms which occur in the anomalous dispersion regime reducing the pulse widths to 4 ps. Subsequent improvements in cavity design eliminated the need for these long cavities as soliton pulse-shaping in an actively mode-locked laser produced pulse widths below 2 ps in a harmonically mode-locked laser operating at 420 MHz[2]. The laser had a Fabry-Perot cavity with only ten meters of fiber which resulted in a cavity mode spacing of approximately 10 MHz. This laser was tunable over the entire erbium gain bandwidth from 1.521 - 1.58 $\mu$m, a property of erbium lasers which makes them attractive for a number of applications.

Following these initial experiments the performance of actively mode-locked fiber lasers has continued to improve. Erbium doped fiber lasers routinely provide transform limited

picosecond pulses with repetition rates in the gigahertz range.[3] Stabilization techniques have reduced environmental fluctuations and thereby allowed extended operation with stable, equal amplitude pulse trains. Phase locking techniques to lock the pulse phase to a stable external source have been developed.[4] The inclusion of a pulse-rate étalon in a ring cavity stabilized a harmonically mode-locked laser when the free-spectral range of the cavity matched the operating repetition rate.[5] More recently, polarization preserving components have been included to eliminate the polarization mode dispersion which occurs in standard fiber lasers. This laser was capable of producing 6 ps pulses at repetition rates up to 40 GHZ while being tunable over 50 nm of the erbium gain bandwidth.[6]

Traditional operation of these lasers at harmonics of the cavity frequency has been well documented and characterized. If the drive signal to the modulator falls on specific frequencies between the cavity harmonics, it is possible to induce the laser to pulse at much higher repetition rates in a state known as rational harmonic mode-locking. By exploiting this method very high repetition rates are possible and have been demonstrated as high as 200 GHz.[7] The mechanisms of this method are limited by the available gain and relative linewidths of the axial modes as will be shown later. This method has renewed interest in the mode-locking mechanisms and potential applications of actively mode-locked fiber ring lasers.

This chapter deals with the development and characterization of an actively mode-locked fiber ring laser. Operation in a number of frequency regimes ranging from megahertz to gigahertz will be described. This laser also serves as the source or master signal for the synchronization experiments described later. Characteristics of this laser which will be considered are the pulse width, spectrum, energy and noise. The chapter concludes with an examination of extending the laser capabilities via rational harmonic mode-locking, a novel method enabling very high repetition rates to be obtained at relatively low drive frequencies.

## 2 Active Mode-Locking Theory

Active mode-locking is achieved by directly modulating the optical field inside the cavity at a frequency $f$ equal to (or a multiple of) the mode spacing $\Delta\nu$. A modulator located within the cavity drives the modulation by becoming transparent at regular intervals occurring at rates equal to the fundamental cavity mode spacing $f_0 = \Delta\nu$ or at a multiple or harmonic of this frequency $f_N = N\Delta\nu$. In the latter case the coupling occurs not between adjacent cavity

modes but between $N$th adjacent modes and is referred to as harmonic mode-locking. The two types of active mode-locking are frequency modulation (FM) and amplitude modulation (AM) depending on the what aspect of the field is modulated. The modulation creates sidebands, which overlap adjacent modes separated by the modulation frequency $f$, allowing the coupling between the modes to occur and the synchronization of their respective phases.

In the case of FM mode-locking the modulator couples the cavity modes by periodically modulating the phase of the optical field inside the cavity. As a means of describing this method of mode-locking, a Fourier component of the optical field will be taken through the three elements of the laser cavity: loss, gain and the modulator.

The loss of the cavity comes about from various sources within the resonator. The most obvious source of loss is the output coupler whether it is a partially reflecting mirror or a fiber optic output coupler. Additional sources of loss occurring within the resonator are due to scattering or index mismatches from components such as the modulator or lenses found within the cavity. All the sources of loss are lumped into a single parameter $\ell$ so that the $n$th mode amplitude $a_n$ is multiplied by $1 - \ell$ on one pass through the resonator.

The gain in the cavity has a definite spectral profile and linewidth determined by the particular gain medium. A common gain profile used to model the medium is a Lorentzian function even though the actual shape of the gain profile may differ considerably from a Lorentzian. This assumption greatly simplifies calculations and captures the essential behavior. The $n$th Fourier component passing through the gain element experiences the change

$$G(n)a_n = (1 + g_n)a_n = \left[1 + \frac{g}{1 + \frac{(n\omega_m)^2}{\Omega_g^2}}\right] a_n \simeq \left\{1 + g\left[1 - \frac{(n\omega_m)^2}{\Omega_g^2}\right]\right\} a_n \qquad (1)$$

where the spectrum is centered at the $n = 0$ Fourier component. The frequency of the $n$th spectral component is $\omega_n = \omega_0 + n\omega_m$, where $\omega_0$ is the carrier frequency. The driving or modulation frequency $\omega_m$ is typically the cavity mode frequency or a multiple thereof for harmonic mode-locking to occur.

The final element of the cavity is the modulator which will be assumed to pass the radiation perfectly at perfectly timed instants spaced $t_m = 2\pi/\omega_m$ apart. This means the component $a_n$ will be multiplied by the modulation $M[1 - cos(\omega_m t)]$, where $M$ is the modulation depth.

$$M[1 - cos(\omega_m t)]a_n e^{i(\omega_0 + n\omega_m)t} = \\ -Ma_n e^{i\omega_0 t}\left[\frac{1}{2}e^{i(n-1)\omega_m t} - e^{in\omega_m t} + \frac{1}{2}e^{i(n+1)\omega_m t}\right] \qquad (2)$$

The Fourier component $a_n^{(k+1)}$, having passed $k$ times through the resonator, acquires the following form on it subsequent e-entry into the cavity.

$$a_n^{(k+1)} = a_n^{(k)} + g\left(1 - \frac{(n\omega_m)^2}{\Omega_g^2}\right) a_n^{(k)} - \ell a_n^{(k)} + \frac{M}{2}(a_{n+1}^{(k)} - 2a_n^{(k)} + a_{n-1}^{(k)}) \tag{3}$$

In the steady state, $a_n^{(k+1)} - a_n^{(k)} = 0$, and Eq. (3) equation reduces to a second order difference equation. Since the spectrum of a mode-locked laser necessarily contains many modes, typically tens of thousands, a second order derivative may be substituted in place of the second order difference in Eq. (3). With this assumption the steady state reduces to:

$$0 = \left[g\left(1 - \frac{\omega^2}{\Omega_g^2}\right) - \ell + \frac{M\omega_m^2}{2}\frac{d^2}{d\omega^2}\right] a(\omega) \tag{4}$$

where the Fourier component $a_n$ has been replaced by the function $a(\omega)$ of the continuous frequency variable $\omega = n\omega_m$. The solution to Eq. (4) is a simple Hermite-Gaussian function

$$a(\omega) = H_\nu(\omega\tau)e^{-\omega^2\tau^2/2} \tag{5}$$

where $H_\nu(\omega\tau)$ is a Hermite polynomial. The solution has the constraints

$$\frac{1}{\tau^4} = \frac{M\omega_m^2\Omega_g^2}{2g} \tag{6}$$

$$g - \ell = M\omega_m^2\tau^2(\nu + \frac{1}{2}) \tag{7}$$

The constraints in Eqs. (6) (7), especially the gain compensation for the dispersion of the spectral components in Eq. (7), stabilize the lowest order Hermite-Gaussian function against the growth of higher order Hermite Gaussians. The spectrum has a Gaussian shape, and the Fourier transform of the spectrum, Eq. (5), provides the temporal shape of the periodic pulses ($\nu = 0$).

$$a(t) = \frac{\sqrt{2\pi}}{\tau}Ae^{-t^2/2\tau^2} \tag{8}$$

The amplitude of the pulse is determined by setting an arbitrary amplitude $A$ for the simple Gaussian solution.

## 3 Laser Configuration

This laser was based on the traditional design of actively mode-locked fiber ring lasers employing a fiber pigtailed Mach-Zehnder modulator.[3, 5, 8] The essential features of these

lasers are an isolator, output coupler and modulator. Since the modulator is a polarization sensitive element, at least one polarization controller is required in the cavity to account for the birefringence of the cavity. The layout for this laser is shown in Fig. (1).

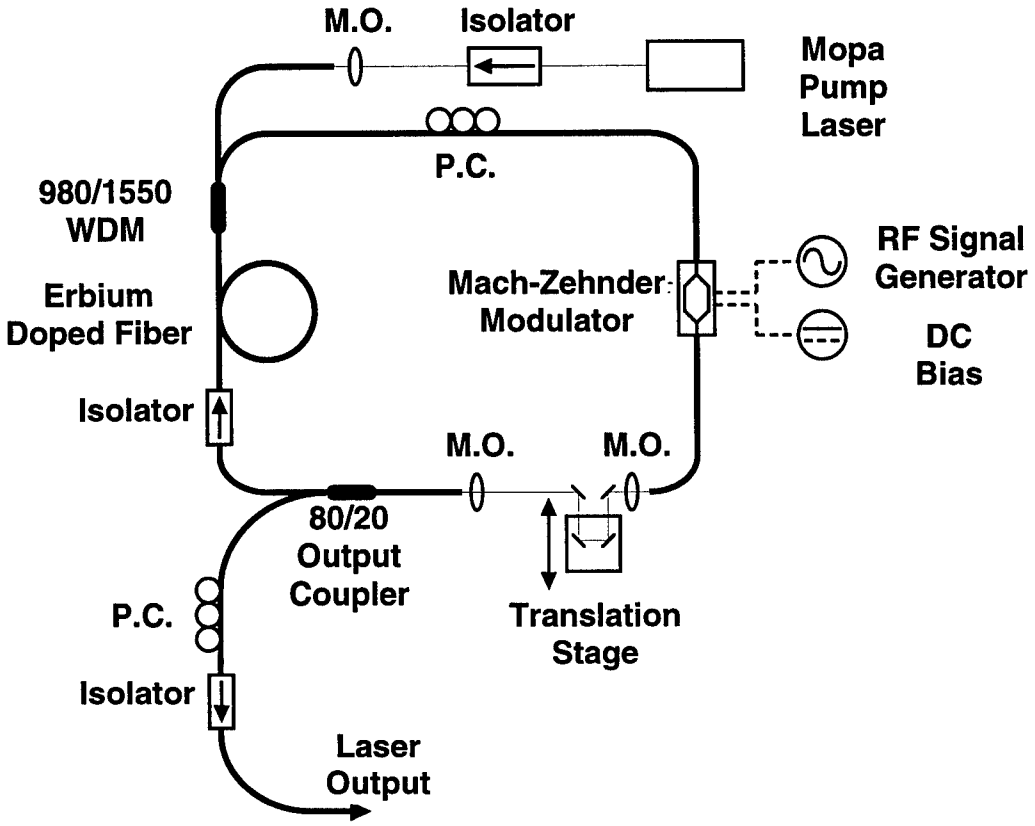



Figure 1: Schematic of the actively mode-locked ring laser.

Four meters of erbium fiber provided by Mike Dennis and Irl Duling of the Naval Research Laboratory acted as the gain medium and was specifically designed to be compatible with standard monomode fiber, such as the industry standard Corning SMF-28. This simplified the splicing of this fiber length to the wavelength division multiplexer (WDM) as there was no significant difference in core diameters. The characteristic values provided with the fiber were listed as $\sim$1.0 db/mW for the absorption with a dispersion of 11$\pm$3 ps/nm/km. The erbium was pumped by an SDL master oscillator/power amplifier (MOPA) diode laser which provided approximately 950 mW of power at a wavelength of 990 nm slightly off from the peak absorption value of 980 nm. The MOPA emits light in a diffraction limited $TEM_{00}$ mode required for efficient coupling to single mode fiber. The pump light was focused onto a flat cleaved strand of Corning Flexcor 1060 fiber which is single mode at 990 nm. The

fiber is single mode at the pump wavelength to prevent higher order transverse modes which could reduce efficiency by producing an intensity profile within the erbium fiber which is not Gaussian. Once optimized the pump delivered about 400 mW of power into the fiber core. The pump then passed through a Gould 980/1550 nm WDM which had an insertion loss of 0.33 dB providing 380 mW directly to the erbium fiber. Approximately 17 mW of pump light was measured coming out of the erbium fiber giving an absorption of 13.5 dB or 3.37 dB/m.

The erbium was fusion spliced to an Optics for Research (OFR) 1550 nm polarization independent optical isolator. The isolator had an insertion loss of - 0.90 dB with an isolation of - 36 dB. The isolator served two purposes in the laser cavity. The first was to ensure unidirectional lasing thereby preventing a significant loss of power caused by lasing in the both directions. The second role of the isolator was to remove any excess pump power from propagating in the cavity which could potentially damage the modulator. Beyond the isolator is the 80/20 coupler through which the laser output is emitted.

The primary component of the laser, except for the erbium doped fiber, was the lithium niobate Mach-Zehnder modulator. The Uniphase Telecommunications Products 1550 nm modulator has an operating bandwidth of 12 GHz with an insertion loss of 3.2 dB. A DC bias on the modulator allows the transmittance to be varied as shown in Fig. (2). The modulator exhibited an incredibly high extinction ratio in excess of 60 dB. The operation of the modulator requires only a single polarization and came fiber pigtailed on both the input and output with Fujikura SM-15-P-8/125-UV/UV-400 polarization maintaining fiber. The input polarization crosstalk was given as 35.9 dB with an output polarization crosstalk of 41.6 dB. The highly polarization selective nature of the modulator allowed it to act as a fast saturable absorber permitting nonlinear pulse shaping to further reduce pulse widths.

The free space delay line allowed fine adjustments to be made to the round trip time. This allowed the laser to be operated at a single frequency with environmental changes being corrected by the delay line rather than by altering the drive frequency. Fiber lasers suffer from environmental changes to the cavity length due to temperature fluctuations and often require electronic stabilization mechanisms to achieve long term stability. With minor adjustments to the delay line, stable operation could be maintained for periods of time up to approximately 30 minutes, sufficient for measurements.

As mentioned earlier, the polarization controller located before the Mach-Zehnder mod-

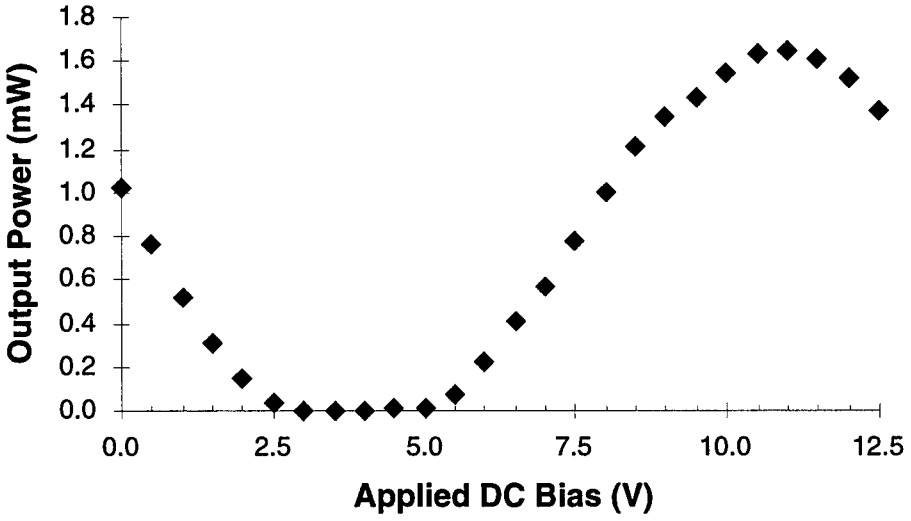



Figure 2: Transmission of the modulator as a function of applied voltage.

ulator permitted direct control of the birefringence while at the same time providing a wavelength selective element. The wavelength selectivity arises from the fact that the polarization controller alters the birefringence within the cavity. By changing the birefringence within the cavity, the optical path length changes as well. Any change in the optical path length prevents the pulses from arriving at the modulator at the proper time. The laser is then self-correcting and adjusts its preferred wavelength to fit the new optical path length. This provides a means of fine adjustment to the operating wavelength.

Additional lengths of fiber could be added to the cavity to provide a coarse adjustment to the cavity length to enable operation at a specific frequency with fine adjustments provided via the delay line. At low repetition rates additional fiber also allowed the nonlinear shaping mechanisms of APM to take place significantly reducing the pulse width. Without the additional fiber picosecond pulses were not possible at rates in the megahertz regime.

With this general geometry the laser made it possible to observe multiple operating regimes with only minor changes to the cavity. The traditional method of harmonic mode-locking provides interesting regimes depending on whether the modulator is driven at low harmonics ($1 \leq n < 100$) or very high harmonics ($n > 500$). Shifting the drive frequency off from a harmonic of the cavity by a predetermined amount allowed the laser to operate in a rationally mode-locked state producing pulses at rates as much as 10 times the applied frequency.

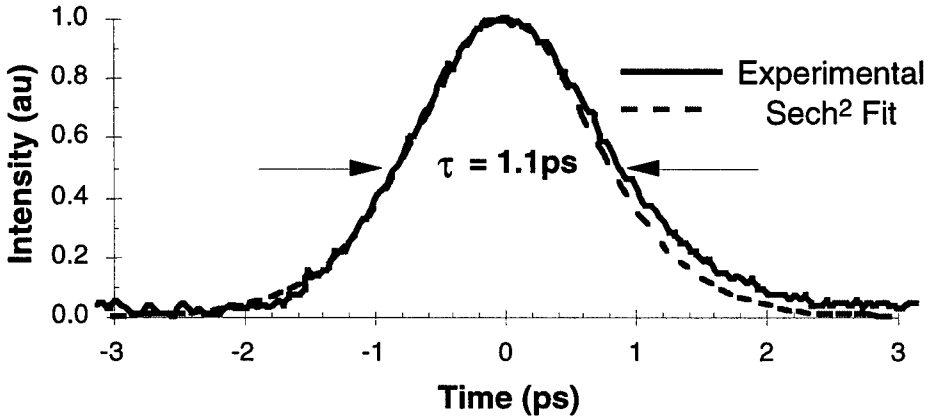
## 4 Mode-Locked Operation

As stated earlier the laser may operate in a number of different regimes producing different outputs. In all cases the output is taken from the output coupler and passed through another polarization controller and OFR polarization insensitive isolator. The isolator prevents any unwanted reflections from re-entering the laser cavity and subsequently altering the laser operation. The polarization controllers provide adequate adjustment of the polarization state to enhance second harmonic autocorrelation traces. In all regimes operational wavelength, power, repetition rates and pulse widths are collected and compared using the following diagnostic system.

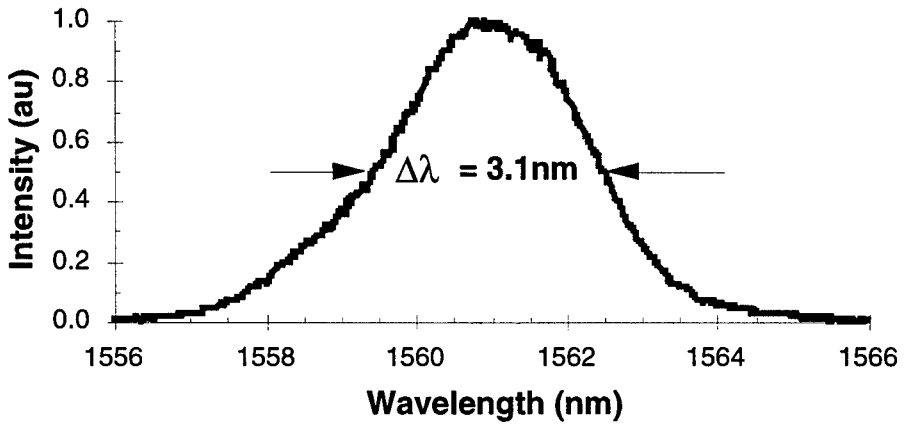
An Anritsu MS9001B1 optical spectrum analyzer measured the operational wavelength of the laser. The analyzer had a resolution bandwidth of 0.1 nm allowing a facile means of monitoring the mode-locked operation by observing spectral broadening characteristics of mode-locking. When operated in cw mode the laser output would mode hope from wavelength to wavelength but was always had a linewidth limited by the resolution of the spectrometer. When the laser entered a mode-locking regime, the spectrometer settled on a single peak wavelength and the linewidth increased beyond the resolution limit indicating multiple modes had been successfully locked. The optical power was monitored using a Newport model 818-IR detector connected to a model 835 power meter. The pulse trains were observed using a New Focus model 1537 6 GHz photodetector or Antel ARD-28 photodetector. Both photodetectors were capable of resolving pulses with pulse widths $\sim$ 50 ps. In both cases the resulting current was directed to various oscilloscopes including a Tektronix SD-24 TDR sampling head. The dual-channel sampling head had an operating bandwidth of 20 GHz allowing structures too large to be observed with the autocorrelator to be recorded. In addition to the oscilloscopes the photocurrent was examined on a Tektronix 2755AP programmable spectrum analyzer where the various components of the pulse train power spectrum could be examined. Finally an Inrad model 5-14-LD non-collinear autocorrelator with a 5 mm $LiNbO_3$ crystal was used to directly measure the pulse width via second harmonic generation. The autocorrelator sampled the pulses over a long period allowing low intensity pulses to be measured effectively. The required value to ensure sufficient signal for measurement was 1 $mW^2$ calculated by multiplying the peak pulse power by the average pulse train power.

### 4.1 Mode-Locking at the Fundamental Frequency

Initial operation of the ring laser was conducted at the fundamental cavity spacing which varied from 4.88 MHz $> \Delta\nu >$ 1.52 MHz. These frequencies correspond to cavity lengths 41 m $<$ L $<$ 136 m. At these low repetition rates the drive frequency window for stable mode-locked operation is very narrow. Characteristics of a typical pulse at a frequency, $\Delta\nu$ = 1.519 MHz are shown in Fig. (3). The pulse has a temporal profile best fit by a hyperbolic secant with a FWHM of 1.1 ps. The corresponding spectral width is 3.1 nm centered at 1561 nm. This width corresponds to a frequency bandwidth of 380 GHz. Assuming a hyperbolic secant shape, the time-bandwidth product of this pulse, $\tau\Delta\nu = 0.408$, exceeds




(a)




(b)

Figure 3: The temporal and spectral output of the ring laser operating at the fundamental frequency, 1.519 MHz. The long cavity length provides sufficient anomolous dispersion to allow 'soliton' pulse shaping as evidenced by the agreement between the observed pulse and a hyperbolic secant squared curve.

the transform limited minima of 0.315. This value indicates additional pulse shortening might be achieved through better management of the cavity dispersion balanced by SPM. The output power of the laser was 5.9 mW which can provide the peak pulse power assuming a negligible background continuum via Eq. (9).

$$P_{peak} = P_{ave} \frac{T}{\tau_p} \tag{9}$$

The pulse shown in Fig. (3) has a peak power of 3.565 kW. The average pulse energy is 3.88 nJ.

It is important to note that pulse widths of this magnitude are much smaller than would be expected for an actively mode-locked erbium fiber laser. Using Eq. (6) the calculated pulse width is approximately 500 ps assuming the value $(\frac{2g}{M})^{\frac{1}{4}}$ is of the order unity. The actual pulse is two orders of magnitude shorter than the predicted width for purely active mode-locking implying that there is substantial nonlinear pulse shaping taking place. The high nonlinear effects are a result of the high peak pulse power within the cavity. The internal peak power is four times the exiting pulse power, 14.26 kW, providing an intensity of 26.36 GW/cm$^2$ in Corning SMF-28 fiber with a core diameter of 8.3 $\mu$m. Although values vary for the nonlinear coefficient for fibers the earliest measured value for a silica glass fiber of $3.2 \times 10^{-16}$ cm$^2$/W is still almost exclusively used.[9] The acquired phase shift for the peak of the pulse due to the nonlinear index of refraction is $\Delta\phi = 0.34$ rad/cm, a very large value for a cavity with a length greater than 100 m. This nonlinear polarization rotation allows the linearly polarizing modulator to act as an intensity discriminator eliminating the less intense edges of the pulse.

The power spectrum of the pulse train could be analyzed using the theory proposed by von der Linde[10] to provide a measure of the timing-jitter. Fig. (4) shows the Dirac $\delta$ peak and surrounding structure for the second and the tenth cavity harmonic to illustrate the differences between harmonics. To determine the nature of the noise, amplitude or temporal fluctuations, the dependence of the particular structure on the corresponding harmonic number must be determined. The sideband structures identified in Fig. (4) exhibited a low frequency noise component whose area followed an n$^2$ dependence, indicative of timing-jitter. Analyzing the higher harmonic spectra, the calculated timing-jitter of the pulse train proved to be 70 ps. This rather large timing jitter is a characteristic of passive rather than active mode-locking. Passive mode-locking dominates the pulse shaping since the effective transmission window of the modulator is much longer at low repetition rates when driven by a

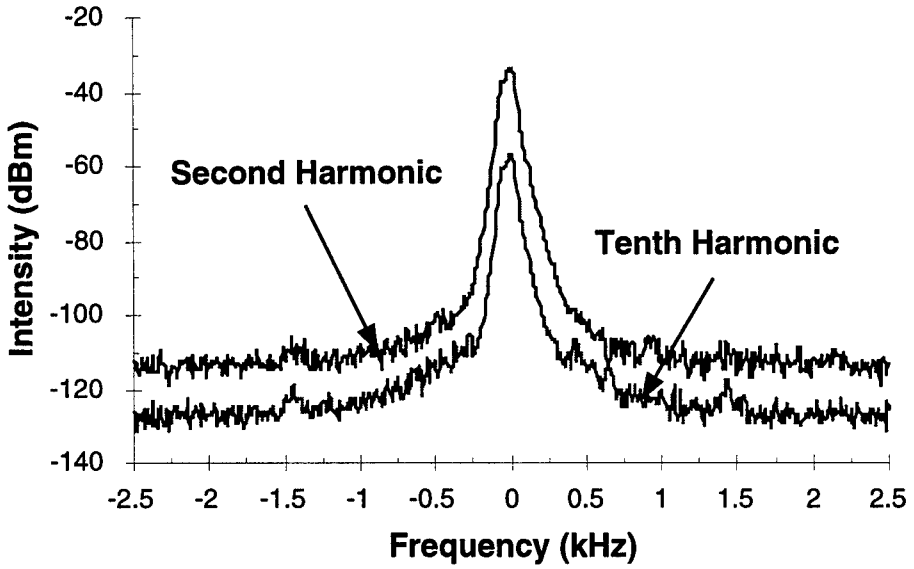



Figure 4: Fourier components of the power spectrum of the ring laser operating at the fundamental cavity frequency. The spectra exhibit the characteristic rise in sideband noise and reduction of the peak powers as the harmonic number increases.

sinusoidal signal. This value could be improved by reducing that transmission window by using a pulse generator rather than a frequency synthesizer.

The ring laser can be operated at low repetition rates and still produce picosecond pulses. These pulse are not shaped by the action of the modulator as in pure active mode-locking, but by passive effects shaping the pulse within the cavity. The laser has the advantage of high output power but suffers from large timing-jitter. The jitter makes this system an unattractive choice for the synchronization experiments or applications which require at great deal of temporal stability.

### 4.2 Harmonic Mode-Locking at High Harmonics

The laser was then set to be operated at much higher harmonics which were generated by a Hewlett Packard 8672A synthesized signal generator capable of generating signals at rates up to 18 GHz. The cavity was shortened to have a fundamental cavity spacing of $\Delta\nu \sim 5$ MHz. The modulator was driven by a 5.006 GHz signal from the generator swinging a voltage of 1.24 $V_{p-p}$. This voltage translates to a modulation depth of 25%. Running at this frequency corresponds to laser operation at the 1024 harmonic. Again using Eq. (6) the predicted pulse width comes out to be on the order of 5 ps which will be shown to be in much closer agreement with the observed pulse widths.

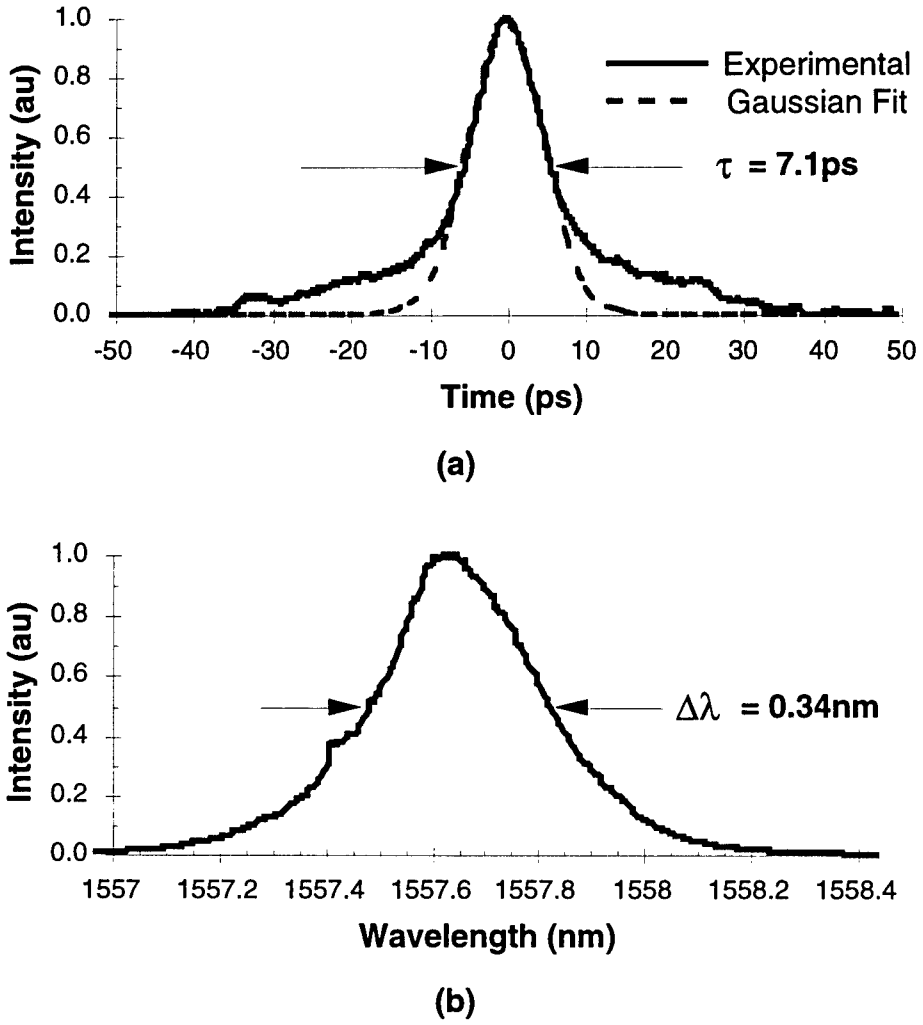



Figure 5: Output of the ring laser operating at 5.006 GHz which corresponds to approximately the 1024 harmonic. The pedestal is due to insufficient nonlinear pulse shaping occuring within the cavity.

The spectral and temporal profiles of the laser output are shown in Fig. (5). The observed pulse width of 7.13 ps and corresponding spectral width of 0.336 nm provide a time-bandwidth product of 0.296 which is slightly lower than the value of 0.315 for hyperbolic secant pulses. At this high repetition rate, the output power of the laser dropped to 0.6 mW. By again employing Eq. (9), the peak pulse power was calculated to be 16.8 mW. The high repetition rate significantly drops the peak power of the pulses decreasing the effectiveness of nonlinear pulse shaping within the cavity. This is evidenced by the wings present on the sides of the pulse. These wings mask the lower portions of the pulse where Gaussian and hyperbolic secant pulses diverge making it difficult to discern an accurate pulse shape. The

mode-locking theory of Chapter 2 would predict a Gaussian pulse shape.

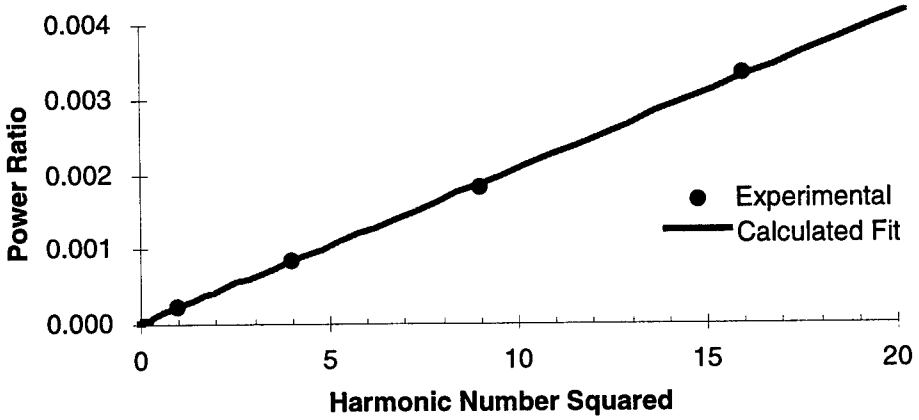



Figure 6: Plot of the ratio of the peak power levels of the delta and noise sidebands at the harmonic Fourier components of the pulse train. The ratio exhibits the $n^2$ dependence characteristic of timing jitter.

The final measure of this systems operation was the noise analysis of the output. The low frequency noise sidebands observed in the power spectra again provided a measure of the timing jitter of the pulses. Only four data points were available for measurement as the bandwidth of the spectrum analyzer extended only out to 21 GHz. The measured timing-jitter was approximately 1 ps which is significantly lower than the jitter observed at low repetition rates. The timing-jitter was most likely sub-picosecond, but because of the inability to examine harmonics higher than the fourth it could not be precisely determined. Since amplitude and timing fluctuations are comparable at the low harmonics, amplitude fluctuations could be contributing to the observed sideband energy.

The ability to operate at repetition rates above a gigahertz combined with low inherent timing jitter allows this laser to be used in a high speed communication architecture as a source or diagnostic tool for testing high speed photonic communication devices. A drawback to the present system is the low peak pulse power which typically requires significant amplification before the pulses can be effectively exploited in an application.

## 5 Rational Mode-Locking

Active mode-locking of a fiber ring laser routinely provides pulse trains at repetition rates in the gigahertz regime. Commercially available Mach-Zehnder $LiNbO_3$ modulators are capable of operating up to 18 GHz. State of the art modulators have been developed which allow

operation at rates as high as 100 GHz[11], but even these are orders of magnitude beneath the carrying capacity of telecommunications fiber. To use as much of the available bandwidth as possible very elaborate WDM and TDM systems have been developed combining large numbers of pulsed laser sources operating at rates much lower than the final bit stream. The large number of expensive components to achieve these high rates makes these systems cost prohibitive. Rational mode-locking provides a means of achieving repetition rates as high as 200 GHZ from a single laser source.[7]

Rational mode-locking operates on the principal that pulsed operation can be achieved with the external signal not operating at a cavity harmonic, $f_{signal} \neq n\Delta\nu$, but rather when $f_{signal} = (n + \frac{1}{p})\Delta\nu$ where $n$ and $p$ are integers.[12] Driven by this signal the laser operates at repetition rates given by $f_{laser} = (np \pm 1)\Delta\nu$. The parameter $p$ greatly increases the repetition rate with negligible demands placed on the modulator response. It will be shown that there is a limit to how large a value $p$ can become before other effects hamper performance. Despite those limitations, this method provides a method of achieving very high repetition rates with components capable of only modest repetition rates.

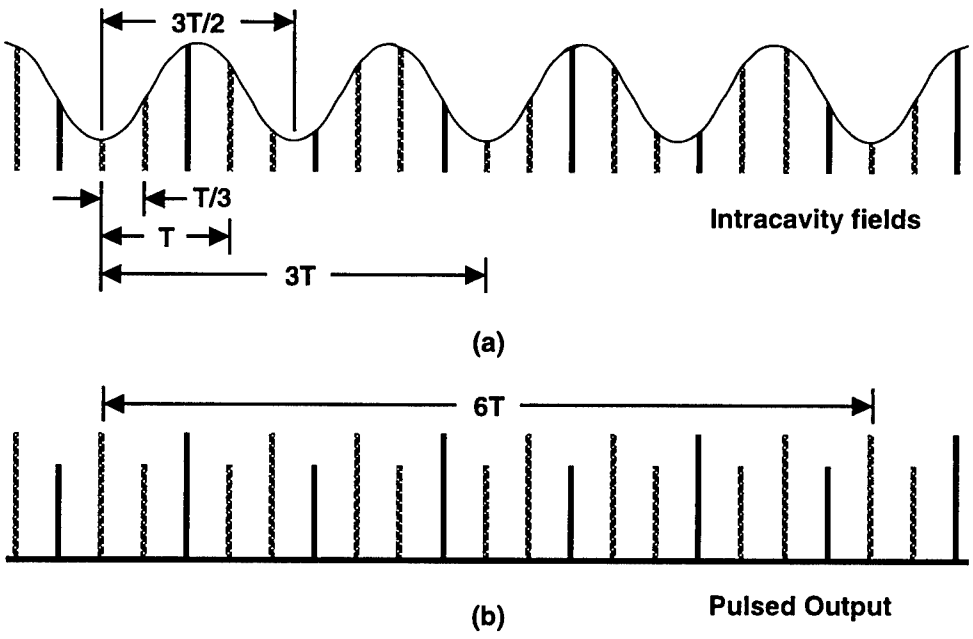



Figure 7: A graphical representation of the intracavity pulses (a) and the output pulse train (b) of a rationally mode-locked laser with $n = 1$ and $p = 2$. The intracavity pulses undergo an irregular modulation which is periodic in time with a period of $np + 1$T or 3T. The observed output is modulated with a period of $p$T or 2T which means the overall pulse series is periodic at 6T.

Fig. (7) illustrates the propagation of a single pulse through the cavity. The pulse amplitude is attenuated by the partially transparent modulator producing the periodically varying pulse train. Within the cavity at any one time, there are $np \pm 1$ pulses providing an equivalent output pulse train. The modulator is transparent for only one of the pulses, however, and the subsequent pulses all suffer attenuation to varying degrees. The amount of attenuation depends on the relative position of the pulse relative to the pulse that passes through the modulator unattenuated.

The experiments relating to rational mode-locking were carried out with the ring laser having a fundamental cavity mode spacing of 1.52 MHz. The actual experiments were carried out by simply choosing a value for $n$ and $p$ and tuning the signal generator to the required frequency. Once the frequency was set, the output was viewed on a Tektronix 2467B 400 MHz analog oscilloscope. In addition to the oscilloscope, the spectrum analyzer monitored the output for spectral broadening to ensure the laser remained mode-locked. Fig. (8) shows the operation of the laser with a drive frequency of only 1.82 MHz which corresponds to values of $n = 1$ and $p = 5$. At this drive frequency the laser operated at a repetition rate of 9.12 MHz. The modulator experienced a change in frequency of only 300 kHz or 19.7% while the repetition rate of the laser increased by a factor of 6. The increase in repetition rate comes with a cost in energy which can be seen in the resulting variations in pulse amplitudes due to the individual pulses experiencing unequal attenuation within the cavity. It is also interesting to note that as the drive frequency decreases, nearing the value of the actual harmonic frequency, the repetition rates increase dramatically.

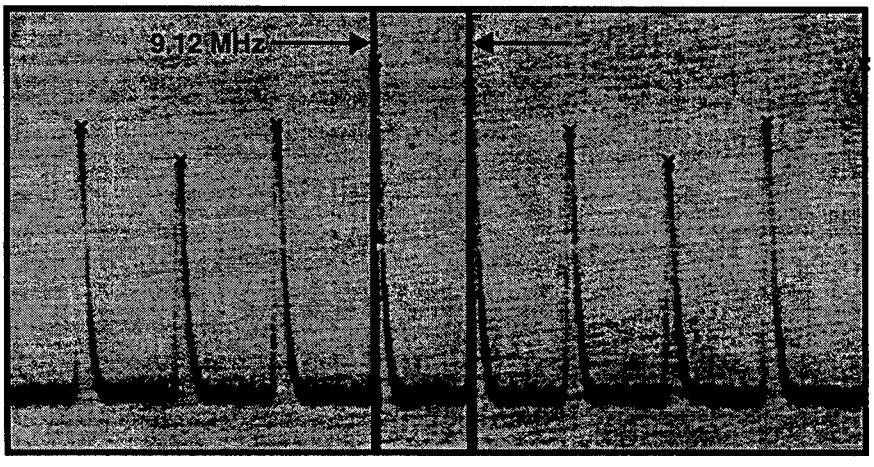



Figure 8: Rational mode-locking pulse train when $n = 1$ and $p = 5$.

As the value of $p$ increases, two effects come to the forefront which limit or reduce the effectiveness of the mode-locking. The first is the attenuation due to the modulator. In every round-trip time a pulse is emitted which has experienced $p$ trips around the cavity without the modulator being completely transparent. This effect is clearly visible in Fig. (9) where $n = 1$ and $p = 10$. The large envelope containing the pulses is characteristic of the sinusoidal drive signal applied to the modulator. The pulses nearest the transmission maxima experience little loss compared to those further away. This is an unavoidable affect observed even at the relatively low values of $p$ as seen in Fig. (8).

The second problem which limits the maximum value $p$ can assume is the competition between rational and harmonic mode-locking regimes. At large values of $p$ the coupling between the modes falls very closely to the nearest neighbor as well as the mode $(np \pm 1)\Delta\nu$ away. As $p$ increases the coupling between nearest neighbors takes over and the laser begins mode-locking at the harmonic frequency $n\Delta\nu$. To date the highest value of $p$ which provided stable pulsed output is 16 by Jeon *et al.*[13]

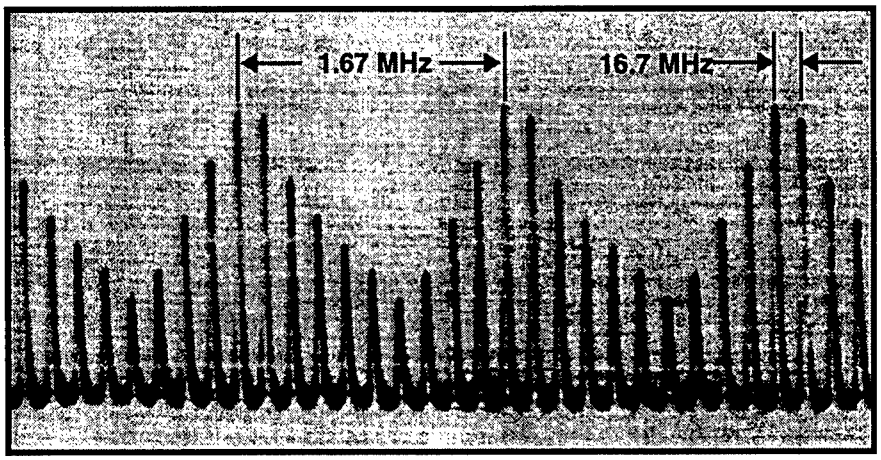



Figure 9: Rational mode-locking when $n = 1$ and $p = 10$. The pulse amplitudes are periodically attenuated with a frequency of 1.671 MHz which corresponds to the drive signal applied to the modulator.

## 6 Summary

Operation of an actively mode-locked erbium fiber laser has been demonstrated in three operating regimes. The laser was capable of producing picosecond pulses at repetition rates ranging from 1 MHz - 5 GHz with output powers on the order of 1 mW. Observed pulse widths

were shorter than those predicted by simple active mode-locking theory indicating additional soliton shaping was occuring within the cavity. The cavity dispersion was not optimized for the pulse energies since the pulses were not transform-limited hyperbolic-secant pulses.

Rational harmonic mode-locking was demonstrated as a means to achieve higher repetition rates than the bandwidth of the modulator. These higher rates come at the cost of additional cavity loss and unequal pulse amplitudes. It has been demonstrated that the unequal amplitudes can be removed by circulating the pulses in an external cavity which is a sub-harmonic of the pulse repetition rate.[13]

The laser provides a suitable source for a number of applications, most notably those relating to optical fiber communications. Examination of this laser operating in a synchronized system will be provided later. This synchronized system could be further expanded to model a true clock-recovery system crucial to any high speed communication systems. The wavelength tunability, variable repetition rates and narrow pulse width are important factors which make actively mode-locked fiber ring lasers attractive sources for communication applications.

# References


[1] J. D. Kafka, T. Baer and D. W. Hall, 'Mode-locked erbium-doped fiber laser with soliton pulse shaping,' *Opt. Lett.*, **14**, 1269 (1989).

[2] J. K. Smith, J. R. Armitage, R. Wyatt, N. J. Doran and S. M. J. Kelly, 'Erbium fibre soliton laser,' *Elec. Lett.*, **26**, 1149 (1990).

[3] H. Takara, S. Kawanishi, M. Saruwatari and K. Noguchi, 'Generation of highly stable 20 GHz transform-limited optical pulses from actively mode-locked $Er^{3+}$-doped fiber lasers with an all-polarisation maintaining ring cavity,' *Elec. Lett.*, **28**, 2095 (1992).

[4] X. Shan, D. Cleland and A. Elis, 'Stabilising Er fibre soliton laser with pulse phase locking,' *Elec. Lett.*, **18**, 107 (1992).

[5] G.T. Harvey and L.F. Mollenauer, 'Harmonically mode-locked fiber ring laser with an internal Fabry-Perot stabilizer for soliton transmission,' *Opt. Lett.*, **18**, 107 (1993).

[6] T. Pfeiffer and G. Veith, '40 GHz pulse generation using a widely tunable all-polarisation preserving erbium fibre ring laser,' *Elec. Lett.*, **29**, 1849 (1993).

[7] E. Yoshida and M. Nakazawa, '80~200 GHz erbium doped fibre laser using a rational harmonic mode-locking technique,' *elec. Lett.*, **32**, 1370 (1996).

[8] R. P. Davey, K. Smith and A. McGuire, 'High-speed, mode-locked, tunable, integrated erbium fibre laser,' *Elec. Lett.*, **28**, 482 (1992).

[9] R. H. Stolen and C. Lin, 'Self-phase modulation in silica optical fibers,' *Phys. Rev. A*, **17**, 1448 (1978)

[10] D. von der Linde, 'Characterization of the noise in continuously operating mode-locked lasers,' *Appl. Phys. B*, **39**, 201 (1986).

[11] K. Noguchi, O. Mitomi, H. Miyazawa, 'Low-voltage and broadband Ti:$LiNbO_3$ modulators operating in the millimeter wavelength region,' *Conference on Optical Fiber Communication*, San Jose, CA, paper ThB2, 1996.

[12] N. Onodera, A. J. Lowery, L. Zhai, Z. Ahmed and R. S. Tucker, 'Frequency multiplication in actively mode-locked semiconductor lasers,' *Appl. Phys. Lett.*, **62**, 1329 (1993).

[13] M. Y. Jeon, H. K. Lee, J. T. Ahn, D. S. Lim, H. Y. Kim, K. H. Kim and E. Lee, '24 GHz high repetition-rate optical pulse generation from an actively modelocked fiber ring laser,' *Optical Society of America Annual Meeting*, Long Beach, CA, paper TuNN3, 1997.

**William Kostis's report was unavailable at the time of publication.**

# A SIMULATION STUDY ON A PARTITIONING PROCEDURE FOR RADAR SIGNAL PROCESSING PROBLEMS

Helen Lau
Graduate Student
Department of Mathematics

Syracuse University
215 Carnegie
Syracuse, NY 13244

Final Report for:
Graduate Student Research Program
Rome Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Rome Laboratory

August 1997

# A SIMULATION STUDY ON A PARTITIONING PROCEDURE FOR RADAR SIGNAL PROCESSING PROBLEMS

Helen Lau
Graduate Student
Department of Mathematics
Syracuse University

Abstract

This paper contains the simulation results on trials of a partitioning procedure developed in a paper by Chen, Melvin and Wicks (1995), for screening k multivariate complex normal populations, with respect to a control population for homogeneous noise covariance matrices. The probability of a correct partition under the least favorable configuration, was used as a performance measure on these simulation trials. Other performance measures evaluated were expected subset sizes, expected number of good populations for each partition, expected number of bad populations for each partition, probability of false alarm and the probability of detection.

# A SIMULATION STUDY ON A PARTITIONING PROCEDURE FOR RADAR SIGNAL PROCESSING PROBLEMS

Helen Lau

## 1. Introduction

In the standard radar detection problem, the goal is to accurately detect a known target in the presence of some noise or interference. In order to do that, it is usually assumed that for some px1 measured signal vector, X, $X=(X_1,X_2,\ldots,X_p)$ is a complex normal (Gaussian) vector with mean $\mu$, and noise covariance $\Sigma$. If a known target $\delta=(\delta_1, \delta_2,\ldots, \delta_p)$ is present in the measured signal, then the expected value of $\mu$ is $\delta$, i.e.: we expect that $E(\mu) = \delta$, where $\delta$ is the px1 vector that characterizes that specific target formation. If there is no target present, then it is expected that $E(\mu) = 0$. It is assumed that all measured signals are normally distributed and share the same pxp noise covariance structure, $\Sigma$.

Under the assumption that all measured signal data share the same noise covariance structure, target detection methods have been thoroughly investigated by several authors. A signal to noise criterion was developed by Reed, Mallett and Brennan in a 1974 paper. Kelly (1986), further developed this signal to noise ratio into a statistical decision theory procedure using a likelihood ratio function of signal to noise. Khatri and Rao (1987) developed a similar t-test of Kelly's likelihood ratio.

As reported in Chen, Melvin and Wicks (1995), serious problems arise in these traditional detection methods if target signals are actually imbedded in non-homogeneous interference. One can expect a significant loss in the power of these detection methods if interference is indeed non-homogeneous. In this 1995 paper, Chen et.al. developed a partitioning procedure from a classical ranking and selection approach in order to select a subset of populations that share the same noise covariance structure with the control population. Their method is characteristic of a subset selection approach as formulated by Gupta in 1956. It is suspected that application of this partitioning procedure should greatly increase the performance of current detection methods.

Using a test statistic and partitioning procedure developed by Chen, et.al., the purpose of this paper is to conduct simulation trials to investigate the effectiveness of this method. Theoretical and empirical probabilities of a correct partition (CP) will be compared as a performance measure. Expected subset sizes and the expected number of similar and dissimilar populations within the subset will also be considered. The probability of false alarm, Pr(FA), and the probability of detection, Pr(D) is also examined for partitioned and non-partitioned data.

This report has been organized in the following way:

Section 2: Summary of Notation

Section 3: A Brief Summary of Ranking and Selection Techniques

Section 4: Partitioning Radar Signal Data

Section 5: An Explanation of the Probability of a Correct Partition

Section 6: Other Performance Measures (e.g.: expected subgroup sizes, Pr(FA), Pr(D).)

Section 7: Simulation Results

Section 8: Concluding Remarks.

## 2. Summary of Notation

The following notations are used in this paper. A* denotes the conjugate transpose of a complex matrix A. In the case of real elements, A* will denote the transpose of A. |a| denotes the complex modulus of a complex number, a. When a, is a real number, |a| is just the absolute value of that number. ||A|| is the determinant of the matrix A. tr(A) is the trace of a matrix A. An nxn identity matrix is denoted $I_n$. $X \sim CN_p(\mu,\Sigma)$ signifies that X is a random px1 vector of a multivariate (p-variate) complex normal distribution with mean, $\mu$, and covariance, $\Sigma$. Hence, X has the following probability density function (p.d.f.):

$$f(X) = \Pi^{-p} \|\Sigma\|^{-1} \exp[-(X-\mu)^* \Sigma^{-1}(X-\mu)].$$

$S \sim CW_p(f, \Sigma)$ denotes that S is a complex pxp positive definite matrix that has a complex Wishart distribution with f degrees of freedom, covariance matrix $\Sigma$, and the following p.d.f.:

$$g(S) = \left[\tilde{\Gamma}_p(f)\right]^{-1} \|\Sigma\|^{-f} \|V\|^{f-p} \exp[-\mathrm{tr}\,\Sigma^{-1} S],$$

where, $\tilde{\Gamma}_p(a) = \Pi^{p(p-1)/2} \prod_{i=1}^{p} (a-i+1)$. In general, a random variable Y with a distribution function $F_Y(y/\theta)$, will be denoted $Y \sim F_Y(y/\theta)$, where $\theta$ is a parameter of the distribution function.

## 3. A Brief Summary of Ranking and Selection Techniques

Historically, there have been two approaches taken within ranking and selection theory, namely, the indifference zone approach and the subset selection approach. In the indifference zone approach, the objective is to find the population that has the "best" characteristic of interest. Alternatively, in subset selection, our goal is to find a subset that contains the "best" population. The partitioning procedure developed by Chen, et.al. (1995), is actually a modification of classical subset selection methods that reflects a process of "screening" populations with respect to a control.

If we consider the set of k independent populations $\Pi = \{\pi_1, \pi_2, ..., \pi_k\}$, where each $\pi_i \sim F(x|\theta_i)$, and $\theta_i$ is an unknown parameter of interest for each population $\pi_i$, i=1,2,...,k, then by the subset selection approach, we would like to select a subset $S$ of $\Pi$ such that the probability that the best population is in $S$, has a minimum probability, $P^*$, where $\frac{1}{k} < P^* < 1$. Any subset which contains the best population results in a correct selection (CS). In the case that we have several subsets that meet the minimum probability requirement, then any one of these subsets may be arbitrarily chosen over all other possible subsets. Thus

any procedure $R$ that defines the subset selection process, must meet the condition that $Pr(CS|R) \geq P^*$, for all possible $\theta$ in $\Theta$, where $\Theta$ is the parameter space. It should be noted that subset sizes will be determined based on analyzed data.

In the indifference zone approach, the goal is to find the "best" population based on the parameter $\theta$. If $\theta_{[1]} \leq \theta_{[2]} \leq \ldots \leq \theta_{[k]}$ represents the ordered values of all $\theta_i$, i=1,2,...,k, then we say that $\pi_i$ is better than $\pi_j$, if $\theta_i$ is greater than $\theta_j$. In general, $\theta_{[i]} \neq \theta_i$. Thus by this method, the purpose is to select the population associated with the largest parameter, $\theta_{[k]}$. Similar to the subset selection approach, any procedure R that defines the selection process, must satisfy some minimum probability requirement, $P^*$, $\frac{1}{k} < P^* < 1$. In addition, the distance between $\theta_{[k]}$ and $\theta_{[k-1]}$ must be greater than some minimum specified value. If we say that $d(\theta_{[i]},\theta_{[j]})$ is a distance function that measures the separation between two parameters $\theta_{[i]}$ and $\theta_{[j]}$ and if $d^*$, is the minimum specified distance allowable between two parameters, then the preference zone is the parameter space defined as $\Theta_{d^*}$ where $\Theta_{d^*} = \{\theta \mid d(\theta_{[k]},\theta_{[k-1]}) \geq d^*\}$, $\Theta_{d^*} \subseteq \Theta$. The complement of the preference zone is called the indifference zone. Thus for any valid procedure R, $\inf\{Pr(CS|R)\} \geq P^*$ for all $\theta$ in $\Theta_{d^*}$. The least favorable configuration is the parameter configuration that achieves the minimum probability of a correct partition over $\Theta_{d^*}$ for the indifference zone approach, and over $\Theta$ for the subset selection approach. For more information on ranking and selection theory, see Gupta et.al. (1979).

## 4. **Partitioning Radar Signal Data**

We know that given a random sample of f signal vectors, $Y_1, Y_2, \ldots, Y_f$, where each $Y_i \sim CN_p(0,\Sigma)$, the detection methods developed by Reed et.al. (1974), Kelly (1986) and Khatri and Rao (1987) are effective in accurately detecting an independent signal vector X where $X \sim CN_p(\mu,\Sigma)$. In the case when noise covariance structures are not the same among sampled signal vectors $Y_1, Y_2, \ldots, Y_f$, we can suppose that there exists k different noise covariance structures among the signal vectors. These f signal vectors can then be partitioned into k subgroups, $\{Y_{11}, Y_{12}, \ldots, Y_{1n_1}\}$, $\{Y_{21}, Y_{22}, \ldots, Y_{2n_2}\}$,..., $\{Y_{k1}, Y_{k2}, \ldots, Y_{kn_k}\}$ which represent a sampling from k distinct populations $\pi_1$, $\pi_2$, ..., $\pi_k$, respectively where each $\pi_i \sim CN_p(0,\Sigma_i)$, i=1,2,...,k. If an independent signal vector X is sampled from a control population $\pi_0$ where $\pi_0 \sim CN_p(\mu_0,\Sigma_0)$, then among these k populations, $\pi_1$, $\pi_2$, ..., $\pi_k$, there should be some population(s) , $\pi_i$, that have noise covariance structures, $\Sigma_i$, similar to $\Sigma_0$. The partitioning procedure developed by Chen et. al.(1996), describes a statistic that allows one to distinguish which populations among these k populations, $\pi_1$, $\pi_2$, ..., $\pi_k$, have covariance matrices, $\Sigma_i$ "close enough" to $\Sigma_0$. In order to define "close enough," the authors looked at the univariate comparison of variances for normal populations with zero means.

Recall that the similarity of the variances from two normal distributions with zero means, is often measured by the ratio distance function d(x,y)=x/y. The variances of these populations are said to be more similar as d(x,y) approaches unity. For the covariance matrices of the multivariate normal case, similar reasoning can be applied to compare multivariate normal matrices since covariance matrices share similar properties with their univariate counterpart. Thus the ratio of covariance matrices, $\Sigma_0\Sigma_i^{-1}$, can be used to measure the similarity between the matrices of the control population, $\pi_0$ and some population, $\pi_j$. If we define two distance functions, $d_1(\Sigma_0,\Sigma_i)$ and $d_2,(\Sigma_0,\Sigma_i)$ that reflect how much deviation we will be allowed between $\Sigma_i$ and $\Sigma_0$, then one can define a partitioning of the populations.

**4.1 Formulation of the Problem:**

Let $\Pi = \{\pi_1, \pi_2, ..., \pi_k\}$, be the set of k populations from which we have k independent samples $Y_{11}, Y_{12}, \ldots, Y_{1n}$, $Y_{21}, Y_{22}, \ldots, Y_{2n}, \ldots$, $Y_{k1}, Y_{k2}, \ldots, Y_{kn}$, of size n. We wish to define two exhaustive subsets, $\Pi_G$ and $\Pi_B$, where $\Pi_G$ will be the subset of populations $\pi_i$, where $\Sigma_i$ is *similar* to $\Sigma_0$, for i=1,2,...,k, and $\Pi_B = \Pi - \Pi_G$. The notation G and B are used here to denote the similar ('good') populations and the dissimilar ('bad') populations, respectively. If we say that $0 \le \lambda_{i1} \le \lambda_{i2} \le \ldots \le \lambda_{ip}$ are the ordered eigenvalues of $\Sigma_0\Sigma_i^{-1}$, and let $d_1(\Sigma_0,\Sigma_i) = \lambda_{i1}$ and $d_2(\Sigma_0,\Sigma_i) = \lambda_{ip}$, then for some $\delta_2 > \delta_1 > 0$, we can define $\Pi_G$ and $\Pi_B$, in the following way:

$$\Pi_B = \{ \pi_i \mid d_1(\Sigma_0, \Sigma_i) > \delta_2 \text{ or } d_2(\Sigma_0, \Sigma_i) < \delta_1 \} \text{ and} \qquad (4\text{-}1)$$

$$\Pi_G = \Pi - \Pi_B. \qquad (4\text{-}2)$$

The values of $\delta_1 > 0$ and $\delta_2 > 0$, are positive real numbers that define the cutoff values which differentiate similar and dissimilar populations. Theoretically, if $\delta_1 = \delta_2 = 1$, then $\Pi_G$ defines the subset of populations $\pi_i$, that are exactly the same as $\pi_0$. Since we are looking for similar populations, $\delta_1$ should be a value less than 1 and $\delta_2$ should be greater than 1.

**4.2 Partitioning Procedure:**

Using the sample data collected from each population, $\pi_i$, i=1,2,...,k, we wish to define a subset $S_G$ based on sample data such that $S_G \subset \Pi_G$. Thus $S_G$ will be a subset that contains only good populations whose eigenvalues fall within the interval range of $[\delta_1, \delta_2]$. The following is the procedure $R_C$ defined by Chen, et.al.(1996), such that a minimum probability requirement, $P^*$, is satisfied for each partitioning.

***Procedure $R_C$***: Let $X \sim CN_p(0,\Sigma_0)$, be an independent sample from the control population $\pi_0$, and let $Y_{i1}$, $Y_{i2}, \ldots, Y_{in} \sim CN_p(0,\Sigma_i)$, i=1,2,...,k be independent samples of n random variables from $\pi_i$. If

$$S = \frac{1}{n}\sum_{j=1}^{n} Y_{ij} Y_{ij}^{*} \qquad (4\text{-}3)$$

is the sample covariance matrix associated with population $\pi_i$, then there exists the following associated test statistic, $T_i$ for each population $\pi_i$, where

$$T_i = \frac{X^* S_i^{-1} X}{n}. \tag{4-4}$$

Then the set of populations $\Pi=\{\pi_1,\pi_2,...,\pi_k\}$ can be partitioned into exhaustive subsets $S_G$ and $S_B$ such that

$$S_G = \{\, \pi_i \mid C \le T_i \le D \,\} \text{ and} \tag{4-5}$$

$$S_B = \{\, \pi_i \mid T_i < C \text{ or } T_i > D \,\}, \tag{4-6}$$

where C and D are chosen such that $\Pr(CP|R) \ge P^*$, where $\frac{1}{k} < P^* < 1$.

## 5. The Probability of Correct Partition

Recall that from subset selection theory, the procedure $R_C$ is valid only if a guaranteed minimum probability of a correct partition, $P^*$, is achieved for all possible $\Sigma_i$ in $\Omega$, where $\Omega$ is the parameter space of the covariance matrix measures. Thus a good performance measure of $R_C$ is the minimum probability of a correct partition, Pr(CP). In particular, we will be interested to see if the theoretical minimum probability of a correct partition is achieved in simulated data. As Chen, et. al. (1996) showed, the minPr(CP) is a function of the LFC of all possible parameters in the parameter space $\Omega$. Thus, the following Pr(CP) is derived given the LFC.

Recall that the probability of a correct partition is the chance that all the bad populations are eliminated. By definition (4-1), those populations whose eigenvalues are less than the minimum cutoff value, $\delta_1$, and those populations whose eigenvalues are all greater than the maximum cutoff value, $\delta_2$, will be considered to be the bad populations. Without loss of generality, let populations $\{\, \pi_1,\ldots,\pi_{k_1} \,\}$ and $\{\, \pi_{k_k+k_2+1},\ldots,\pi_k \,\}$, where $\pi_1,\ldots,\pi_{k_1}$ have covariance matrices $\Sigma_1,\Sigma_2,\ldots,\Sigma_{k_1}$ where $\Sigma_0\Sigma_i^{-1}$, i=1,2,...,$k_1$, have eigenvalues which are all less than $\delta_1$, and $\pi_{k_k+k_2+1},\ldots,\pi_k$ have covariance matrices $\Sigma_{k_1+k_2+1},\Sigma_{k_1+k_2+2},\ldots,\Sigma_k$, where $\Sigma_0\Sigma_j^{-1}$, j=$k_1$+$k_2$+1,...,k, have eigenvalues which are all greater than $\delta_2$, denote those populations which are bad. Then in their 1996 paper, Chen, et.al. showed that the probability of a correct partition is as follows:

$$\Pr(CP) = \Pr(\, T_i < C \text{ and } T_j > D \,), \text{ for i=1,2,...,}k_1 \text{ and j=}k_1+k_2+1,...,k \tag{5-1}$$

$$= \Pr\left( \frac{X^* S_i^{-1} X}{n} < C, \frac{X^* S_j^{-1} X}{n} > D \right), \text{ for i=1,2,...,}k_1 \text{ and j=}k_1+k_2+1,...,k \tag{5-2}$$

$$= \Pr\left( \frac{1}{\left(\frac{X^* S_i^{-1} X}{n}\right)} > \frac{1}{C}, \frac{1}{\left(\frac{X^* S_j^{-1} X}{n}\right)} < \frac{1}{D} \right), \text{ for i=1,2,...,}k_1 \text{ and j=}k_1+k_2+1,...,k \tag{5-3}$$

$$= \Pr\left( \frac{X^* \Lambda_i X}{\frac{X^* S_i^{-1} X}{n}} > \frac{X^* \Lambda_i X}{C}, \frac{X^* \Lambda_j X}{\frac{X^* S_j^{-1} X}{n}} < \frac{X^* \Lambda_j X}{D} \right), \text{ for } i=1,2,\ldots,k_1 \text{ and } j=k_1+k_2+1,\ldots,k \qquad (5\text{-}4)$$

where $\Lambda_i = \Sigma_0 \Sigma_i^{-1}$. Rao (1973) found the distribution function of $Y_i = \frac{X^* \Lambda_i X}{\frac{X^* S_i^{-1} X}{n}}$ to be a Chi-squared distribution with 2(n-p+1) degrees of freedom. Since the partitioning is dependent upon the relation of $\Lambda_i = \Sigma_0 \Sigma_i^{-1}$, the covariance matrix of the control population, $\Sigma_0$, is assumed to be a pxp identity matrix, $I_p$, and $\Sigma_i$ are diagonal matrices of the reciprocal eigenvalues of the matrix $\Sigma_0 \Sigma_i^{-1}$. Thus, if $\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{ip}$, are the eigenvalues of $\Sigma_0 \Sigma_i^{-1}$, and $\Sigma_0 = I_p$, then

$$(5\text{-}5) \qquad \Sigma_i = \begin{bmatrix} \frac{1}{\lambda_{i1}} & & & 0 \\ & \frac{1}{\lambda_{i2}} & & \\ & & \cdots & \\ 0 & & & \frac{1}{\lambda_{ip}} \end{bmatrix}, \text{ and } \Lambda_i = \begin{bmatrix} \lambda_{i1} & & & 0 \\ & \lambda_{i2} & & \\ & & \cdots & \\ 0 & & & \lambda_{ip} \end{bmatrix}. \qquad (5\text{-}6)$$

Recall that if $\delta_1$ is the lower bound and $\delta_2$ is the upper bound of the distance functions $d_1(\Sigma_0, \Sigma_i) = \lambda_{i1}$ and $d_2(\Sigma_0, \Sigma_i) = \lambda_{ip}$ where (4-1) and (4-2) are satisfied, then given the Pr(CP), (5-1), it is clear that the minimum probability of a correct partition is achieved at the point where $\delta_1 = \lambda_{i1}$ and where $\delta_2 = \lambda_{ip}$, since the value of $X^* \Lambda_i X$ will increase as any diagonal element of $\Lambda_i$ is increased, while all other diagonal elements are held fixed. Then for populations $\pi_1, \ldots, \pi_{k_1}$, which will have a test statistic $T_i < C$, then the probability of a correct partition as defined in (5-4), is minimized when the diagonal elements of $\Lambda_i$, i=1,2,..., $k_1$, reach the value $\delta_1$. By the same reasoning, the probability of a correct partition is minimized for populations $\pi_{k_k+k_2+1}, \ldots, \pi_k$, when the diagonal elements of $\Lambda_i$, i= $k_1 + k_2 + 1$,...,k, have reached the value $\delta_2$. It is also true that the probability of a correct partition is minimized when <u>all</u> populations are dissimilar to the control population, hence all populations have been eliminated.

Chen, et.al. showed that the LFC of any parameter vector $\{\lambda_{11}, \lambda_{12}, \ldots, \lambda_{1p}\}$, $\{\lambda_{21}, \lambda_{22}, \ldots, \lambda_{2p}\}, \ldots, \{\lambda_{k1}, \lambda_{k2}, \ldots, \lambda_{kp}\}$ under $R_C$ is when

$$\lambda_{i1} = \lambda_{i2} = \ldots = \lambda_{ip} = \delta_1, \qquad i=1,2,\ldots,m \text{ and} \qquad (5\text{-}7)$$

$$\lambda_{j1} = \lambda_{j2} = \ldots = \lambda_{jp} = \delta_2, \qquad j=m+1,\ldots,k, \qquad (5\text{-}8)$$

where m is an integer between 0 and k that minimizes Pr(CP). Thus, the LFC is where $\delta_1 = \lambda_{i1}$ and $\delta_2 = \lambda_{ip}$. Since we know that $Y_i \sim \chi^2_{2(n-p+1)}$, then Chen, et. al. (1996) showed that the probability of a correct partition given the least favorable configuration is minimized by the function:

$$\min \Pr(CP|LFC) = \int_0^{\infty} \left( \chi^2 \left( \frac{\delta_2}{D} y \right) \right)^{k-m} \left( 1 - \chi^2 \left( \frac{\delta_1}{C} y \right) \right)^{m} f(y) dy , \qquad (5\text{-}9)$$

where $\chi^2$ is the cumulative density function (c.d.f.) of a Chi-square distribution with 2(n-p+1) degrees of freedom and f(y) is the p.d.f. of a Chi-square distribution with 2p degrees of freedom.

Note that if the ratios of $\frac{\delta_1}{C}$ and $\frac{\delta_2}{D}$ are fixed, then minPr(CP|LFC) stays the same. The proof of this is straight forward if we suppose that $r_1$ and $r_2$ are fixed ratios where $\frac{\delta_1}{C} = r_1$ and $\frac{\delta_2}{D} = r_2$. Another interesting point is that if the value of C is increased, then the Pr(CP|LFC) increases since the value of $\frac{\delta_1}{C} y$, will decrease, hence the c.d.f., $\chi^2(\frac{\delta_1}{C} y)$ will decrease. By similar reasoning, as the value of D decreases, $\frac{\delta_2}{D} y$ will increase, so $\chi^2(\frac{\delta_2}{D} y)$ will increase and the minPr(CP|LFC) will increase. Thus, if one is having difficulty finding adequate boundaries, C and D, for the test statistic $T_i$, one might try increasing the value of C or decreasing the value of D. Of course one does not want very large intervals, $(-\infty, C] \cup [D, \infty)$, for the test statistic $T_i$, to be rejected since this will result in a loss of power in the procedure. Another interesting phenomenon that needs further investigation is the occurrence of the maximum minPr(CP|LFC) at the point near n = 2p. That is, for fixed k, p, C, D, $\delta_1$ and $\delta_2$, the greatest probability of a correct partition given the least favorable configuration appears to occur at n=2p or somewhere near that value. Solving for zero in the derivative of (5-9) is difficult since this reduces to finding the roots of the derivative of the $\chi^2$ p.d.f. that has 2(n-p+1) degrees of freedom. From the cases considered however, we do know that minPr(CP|LFC) is not a monotonic increasing function of n.

## 6. Other Performance Measurements

Other performance measures considered in these trials were expected subset sizes, expected number of good populations in a subset and the expected number of bad populations in a subset. We also looked at the probability of detection and the probability of a false alarm in partitioned data versus non-partitioned data in the simulated trials. We wanted to measure the differences in results for partitioned data and non-partitioned data when the random signal vector contains a known target formation using the traditional detection methods. In statistical terms, it is clear that the probability of detection, Pr(D), is an analogous measure of the power of a testing method and the probability of a false alarm, Pr(FA), is an equivalent measure of the level of significance achieved by a testing procedure when the null hypothesis assumes that there is no target present in some independent measured signal vector. We suspect that there should be significant differences in the simulated results for this performance measure. Kelly's likelihood ratio test statistic was used in the simulation trials to test for target detection. It should be noted that the use of Khatri and Rao's t-test statistic will also produce similar results if used as a testing criterion. In fact,

simulated trials showed that the results are exactly the same when all elements are real random variables since a unique functional relationship can be shown to exist between Kelly's likelihood ratio test statistic and Khatri and Rao's t-test statistic. In the case of complex random variables, the results were similar, but not exact. This may be due to the fact that Khatri and Rao's t-statistic is only measured by the real component of the test statistic, t, where,

$$t = \frac{\{2(f-p+1)\}^{\frac{1}{2}}\delta^* S^{-1}X}{\{(a^{-1}+X^*S^{-1}X)\delta^*S^{-1}\delta - |\delta^*S^{-1}X|^2\}^{\frac{1}{2}}}. \tag{6-1}$$

It should be noted that f is the total number of observations, p is the number of signal variates measured, $\delta$ is the px1 complex target formation vector, a is some known scalar where $X \sim CN_p(\mu, a^{-1}\Sigma)$ is the random signal vector, and $S \sim CW_p(f,\Sigma)$ is the sample covariance matrix where $S = \sum_{j=1}^{f} Y_j Y_j^*$.

From Kelly's paper, we have that the likelihood ratio test for $\zeta > \zeta_0$ where

$$\zeta = \frac{1+(X^*S^{-1}X)}{1+(X^*S^{-1}X) - \dfrac{|\delta^*S^{-1}X|^2}{\delta^*S^{-1}\delta}}, \tag{6-2}$$

is equivalent to the likelihood ratio test for $\eta > \eta_0$ where

$$\eta = \frac{|\delta^*S^{-1}X|}{(\delta^*S^{-1}\delta)[1+X^*S^{-1}X]}. \tag{6-3}$$

(The test statistic $\eta$ is preferred for its simplicity over $\zeta$.) After some algebraic manipulation, we can see that $\zeta = \frac{1}{1-\eta}$ and $\eta = \frac{|t|^2}{\upsilon+|t|^2}$, where t is Khatri and Rao's test statistic and t has a Student t-distribution with $\nu$ degrees of freedom and a=1. Thus, $\eta = \frac{\zeta-1}{\zeta}$ and $\frac{|t|^2}{\upsilon} = \frac{\eta}{1-\eta}$.

*Lemma 1*: In the case for real random variables, if $t \sim t_\nu$, with $\nu = f-p+1$, then $\eta$ has a Beta distribution with parameters $\frac{1}{2}$ and $\frac{\upsilon}{2}$, i.e.: $\eta \sim \text{Beta}(\frac{1}{2}, \frac{\upsilon}{2})$.

*Proof:* The proof of this is straight forward since $|t|^2 = t^2$ has an F distribution with 1 and $\nu$ degrees of freedom and $\eta = \dfrac{\frac{|t|^2}{\upsilon}}{1+\frac{|t|^2}{\upsilon}} = \frac{|t|^2}{\upsilon+|t|^2}$. (Q.E.D.)

*Lemma 2*: In the case for complex random variables, then $\eta \sim \text{Beta}(1, \frac{\upsilon}{2})$, where the real component of the test statistic t has $\chi_\nu^2$, where $\nu = 2(f-p+1)$.

*Proof:* The proof of this follows from Kelly's paper. Since $\zeta = 1+\frac{|w|^2}{T} = \frac{T+|w|^2}{T}$ , where $T \sim \chi^2$ distribution with $\nu = 2(f-p+1)$ degrees of freedom and $w^{**} \sim CN_1(0,I_1)$, are independent random variables. Then it follows that, $|w|^2 \sim \chi^2$ distribution with 2 degrees of freedom. From Johnson and Kotz, (1970), (p. 38), we have that $\eta \sim$ Beta $(1,\frac{\upsilon}{2})$, since $\eta = \frac{\zeta-1}{\zeta} = \frac{|w|^2}{|w|^2+T}$. (Q.E.D.)

**Note: The distribution of w as defined above, is a conditional distribution. However, the condition is not a condition of T according to Kelly's article. Then since T and w are independent random variables, the covariance of w and T must be zero, hence the conditional distribution of w is not affected by T.

As mentioned earlier, the differences that arise between Khatri and Rao's t-statistic and Kelly's $\eta$ - statistic in the complex case may be due to the effect of the complex component of Khatri and Rao's t-statistic. Recall that the functional relation between $\eta$ and t is that

$$\frac{|t|^2}{\upsilon} = \frac{\eta}{1-\eta}, \tag{6-4}$$

where $\nu = 2(f-p+1)$ are the degrees of free for the real component of t in the complex case. From Johnson and Kotz, (1970), (p.51), we know that $\frac{|t|^2}{2}$ has a Pearson Type IV distribution with 1 and $\frac{\upsilon}{2}$ degrees of freedom, since that $\eta \sim$ Beta $(1,\frac{\upsilon}{2})$,

*Lemma 3*: Then $\frac{|t|^2}{2}$ will have an F distribution with 2 and $\nu$ degrees of freedom.

*Proof:* Using some basic transformations and substitutions, this proof is straight forward. (Q.E.D.)

Taking $\frac{|t|^2}{2}$ as a test statistic, we should get exactly the same results as Kelly's $\eta$ likelihood ratio statistic. This poses a very interesting question about the complex component of Khatri and Rao's test statistic and the effects of leaving this measure out of the detection method. If the complex component of a vector is very significant for a case in study, then one might consider Kelly's detection method (or the equivalent F-statistic, $\frac{|t|^2}{2}$) over Khatri and Rao's method for complex data. This would be an interesting topic for further research.

## 7. Simulation Trial Results

Using MATLAB, v.5.0, , simulated trials were conducted for the case when we had k=4 populations, p=20 signal variates and n=33 observations for each population $\pi_i$, i=1,2,3,4. Taking the minimum eigenvalue cutoff ($\delta_1$) to be 1/6 and the maximum eigenvalue cutoff ($\delta_2$) to be 6, we get that for C=1/2 and D=2, the theoretical minimum probability of a correct partition is 0.9246. For this particular

case, we examined 3 specific trials where in each trial, we assumed that of 4 populations, $\pi_1$, $\pi_2$, $\pi_3$, $\pi_4$, there were 2 similar populations, $\pi_1$ and $\pi_2$, and 2 dissimilar populations, $\pi_3$ and $\pi_4$, to the control, $\pi_0$. For each trial, $\pi_1$ and $\pi_2$, were the same. These "good" populations were defined to be the populations, where $\Sigma_0\Sigma_1^{-1}$ had eigenvalues, $\lambda_{1j}$, j=1,2,...,20, where $\lambda_{1,1}= \lambda_{1,2}= \ldots =\lambda_{1,20} = 1$ and $\Sigma_0\Sigma_2^{-1}$ had eigenvalues, $\lambda_{2,j}$, j=1,2,...,20, where $0.5 \leq \lambda_{2,1} \leq \lambda_{2,2} \leq \ldots \leq \lambda_{2,20} \leq 2$. Note, $\Sigma_1$ is exactly the same as $\Sigma_0$.

In Trial 1, we assumed that the 2 dissimilar populations, $\pi_3$ and $\pi_4$, had covariance matrices where the eigenvalues for $\Sigma_0\Sigma_j^{-1}$, j=3,4, that were very large. For $\pi_3$, the eigenvalues of the distance relation $\Sigma_0\Sigma_3^{-1}$ were in the interval (6,105). For $\pi_4$, the eigenvalues ranged from 12 to 30. In the 2nd Trial, we assumed that both populations, $\pi_3$ and $\pi_4$, had covariance matrices where the eigenvalues for $\Sigma_0\Sigma_j^{-1}$, j=3,4, were both very small. For $\Sigma_0\Sigma_j^{-1}$, j=3,4, the eigenvalues defined to be in the intervals (0.03, 0.09) and (0.005, 0.155) respectively, for j=3 and j=4. For Trial 3, we considered the case where $\Sigma_0\Sigma_3^{-1}$, had eigenvalues in the interval (0.03, 0.09) and where $\Sigma_0\Sigma_4^{-1}$, had eigenvalues from (6.0, 105.0). Thus, in this trial, one population had significantly larger eigenvalues for the distance function $\Sigma_0\Sigma_j^{-1}$, and one population had significantly smaller eigenvalues for $\Sigma_0\Sigma_j^{-1}$.

An algorithm was derived using MATLAB to generate a random sample vector $X \sim CN_{20}(0, \Sigma_0)$ and four independent sets $\mathbf{Y_i}=\{Y_{i,1}, Y_{i,2},\ldots,Y_{i,33}\}$, of 33 random sample vectors where $Y_{i,j} \sim CN_{20}(0, \Sigma_i)$, i=1,2,3,4 and j=1,2,...,33. Then for each set of data $\mathbf{Y_i}$, Chen's partitioning procedure was applied to determine which set of sample data had a covariance matrix significantly different from the control covariance matrix, $\Sigma_0$. Then using the partitioned data, we tested all the data from the "good" samples against the null hypothesis, $H_o$, that there was no target present in the random signal vector X using Kelly's detection method at 0.05 level of significance. We then tested the same null hypothesis using all the generated data from $\pi_1$, $\pi_2$, $\pi_3$, $\pi_4$, also at the same level of significance. This procedure was repeated 1000 times and counters kept track of how often the null hypothesis was rejected and accepted. A similar procedure was performed to evaluate the effectiveness of Chen's procedure for detecting targets in non-homogeneous data. The simulation results are summarized in the following Table 1 for 1000 iterations.

From these results, we can see that for each simulated trial, the theoretical probability of a correct partition was ALWAYS achieved. In fact, these empirical trials always found the Pr(CP) to be almost always 1. In addition, we can see that the average subgroup sizes were fairly consistent for all 3 cases where the partitioning method found the average "good" population subgroup size to be approximately 1.7. Only once out of 6000 trials was a bad population included in the subgroup of populations similar to the control. These results clearly show that there is high precision in Chen's partitioning procedure for this particular example and indicates a high probability of a correct partitioning for good critical values C, D, $\delta_1$ and $\delta_2$.

Next, the false alarm and detection probabilities, Pr(FA) and Pr(D), respectively, for partitioned data based on Chen's statistic $T_i$, versus cumulative non-homogeneous data, were evaluated using Kelly's

**********

| | TRIAL 1 | | TRIAL 2 | | TRIAL 3 | |
|---|---|---|---|---|---|---|
| **$E(S_G)$*** | 1.7270 | | 1.7460 | | 1.7395 | |
| **$E(G_S)$*** | 1.7270 | | 1.7460 | | 1.7395 | |
| **$E(B_S)$ *** | 0 | | 0 | | 0.0005 | |
| **Empirical Pr(CP)** | 1 | | 1 | | 0.9995 | |
| | **False Alarm** | **Detection** | **False Alarm** | **Detection** | **False Alarm** | **Detection** |
| **Partitioned Data** | 0.0205 | 0.7209 | 0.0265 | 0.7255 | 0.0370 | 0.7159 |
| **Cumulative Data** | 0.2160 | 0.9560 | 0.0000 | 0.0010 | 0.0060 | 0.3200 |

**Table 1: Simulation Results**

*$E(S_G)$ denotes the average subgroup size, $S_G$, of the partitioned data for 2000 trials.

*$E(G_S)$ represents the average number of GOOD populations in the subgroup, $S_G$, for 2000 trials.

*$E(B_S)$ represents the average number of BAD populations in the subgroup, $S_G$, for 2000 trials.

**********

detection method. In Trial 1, we have a good power of detection but a poor false alarm rate of 21.60% for cumulative data. It is possible that Pr(D) is good for cumulative data since the Pr(FA) is high. For the partitioned data, we see that Pr(FA) is 0.0206 which is better than the 0.05 specified level of significance. In Trial 1, we also see that Pr(D) = 0.7209, is not as good as in the case for cumulative data, although a detection rate of 72.09% may still be adequate for a radar system. In Trial 2, we have that for both partitioned data and cumulative data, we get good false alarm rates which are both less than the fixed 0.05 level of significance, but there are clear distinctions in the detection rates for partitioned data versus cumulative data. For this particular trial, the simulation results indicated that the detection rate for cumulative data is almost zero at 0.10%, where as the rate of detection for partitioned data is significantly better at the rate of 72.55%. In Trial 3, we see that as in the 2nd Trial, false alarm probabilities are good for both the partitioned and cumulative data. The probability of detection is also significantly better and has an adequate chance of detection for partitioned data versus the cumulative non-homogeneous data.

Thus we see that without partitioning data, there is always some trade off when using Kelly's criterion as a detection method for cumulative non-homogeneous data. When non-homogeneous data has been partitioned to find a subset of data that may be similar enough to the control population, we see that the false alarm rate is always better than the 5% level of significance in this example. Detection rates were also usually better for partitioned data as we saw in Trial 2 and 3. In the case where the detection rate for cumulative data was better than the Pr(D) for the partitioned data, there was a significant trade off for this higher detection rate as the Pr(FA) for cumulative data became very high and it is quite possible that this

detection rate is so high because the false alarm rate is so high. In all cases considered, the probabilities of detection for partitioned data were always found to be acceptable.

## 8. Concluding Remarks

Results from these trials demonstrate that the test statistic $T_i$, (4-4) and the procedure $R_C$ formulated by Chen, et. al. (1996) may be a definitive measure for discriminating non-homogeneous populations in radar signal data. In cases when non-homogeneous data has not been screened/partitioned, these simulation results indicate a significant trade off between the false alarm rate and the rate of target detection. When data was been partitioned via Chen's procedure, these simulation results showed that the false alarm rate for partitioned data always is better the fixed 5% level of significance. Detection rates for screened data were also shown to always be at least 70%, which was often better than the detection rate for cumulative data in this simulation. In the one case where the detection rate for cumulative non-homogeneous data was better than the detection rate for the partitioned data, there was a considerable trade off in its corresponding false alarm rate. In general, we can see that by screening non-homogeneous data, there was always an improvement in the false alarm rate while a good probability of detection measure was attained. The Pr(D) for partitioned data was often significantly better than the detection rate for cumulative non-homogeneous data in the case examined for this paper.

We see that Chen's et. al. detection procedure does improve detection performance while maintaining a specific false alarm measure when signal data is non-homogeneous for the case presented here. One might consider it worthwhile to investigate other cases where the situation variables (k, n, p, etc.) are different, although it is expected that one should obtain similar, if not better, results. The principal results of this paper affirm what was hypothesized in Chen et. al. (1995), since we generally saw an improvement in detection rates after screening data.

**Acknowledgements**

The author wishes to thank the following people for their support and guidance through this summer research experience: Professor Pinyuen Chen, Syracuse University, Capt. William L. Melvin, U. S. Air Force Rome Laboratory and Dr. Michael C. Wicks, U. S. Air Force Rome Laboratory.

**References**

Anderson, T. W. (1984). *An introduction to multivariate statistical analysis*, second edition, Wiley, New York.

Chen, P. (1994). On testing the equality of covariance matrices under singularity, *Final Report*, 1994, AFOSR Summer Faculty Research Program.

Chen, P. Melvin, W., and Wicks, M. J. (1995). Partitioning procedure in radar signal processing problems, *Final Report*, 1995, AFOSR Summer Faculty Research Program.

Chen, P. Melvin, W., and Wicks, M. J. (1996). Screening among multivariate normal data, *Technical Report.*

Goodman, N. R., (1963). Statistical analysis based on certain multivariate complex Gaussian distribution, *Annuals of Mathematical Statistics*, vol. 34, Mar. 1963, 152-177.

Gupta, S. S. and Panchapakesan, S. (1979). *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations*, John Wiley & Sons, New York, New York.

Johnson, N. L. and Kotz, S. (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, John Wiley & Sons, Inc., New York.

Kelly, E. J. (1986). An adaptive detection algorithm, *IEEE Transactions on Aerospace and Electronic Systems,* vol. AES-22, #1, 115-127.

Khatri, C. G. and C. R. Rao (1987). Test for a specified signal when the noise covariance matrix is unknown, *Journal of Multivariate Analysis*, vol. 22, number 2, 177-188.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications* ($2^{nd}$ ed.), John Wiley & Sons, New York.

Reed, I. S., Mallett, J. D. and Brennan, L. E. (1974). Rapid convergence rate in adaptive arrays, *IEEE Transactions on Aerospace and Electronic Systems,* vol. AES-10, #6, 853-863.

# SIMULATION OF A ROBUST LOCALLY OPTIMUM RECEIVER IN CORRELATED NOISE USING AUTOREGRESSIVE MODELING

**Myron R Mychal**
Research Assistant
Electrical and Computer Engineering Department
Illinois Institute of Technology

Illinois Institute of Technology
3300, S. Federal Street
Chicago, IL 60616-3793

Final Report for:
Summer Faculty Research Program
Rome Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Rome Laboratory

September 1997

# SIMULATION OF A ROBUST LOCALLY OPTIMUM RECEIVER IN CORRELATED NOISE USING AUTOREGRESSIVE MODELING

**Myron R Mychal**
Research Assistant
Electrical and Computer Engineering Department
Illinois Institute of Technology

Abstract

This report presents the simulation of a robust *locally optimum* (LO) non-linear spread spectrum receiver. The signaling environment consists of the desired received signal in correlated interference and thermal noise. Autoregressive (AR) spectral modeling methods and a histogram approximation of the probability density function are employed. Preliminary results for transmission in the presence of *continuous wave* (CW) and *partial band* (PB) interference are presented and discussed. A comparison of this method to a similar nonlinear processing algorithm is performed. Preliminary results for the performance of a binary phase-shift keyed communications system indicate that applying AR modeling to the environment improves performance substantially, especially in the case of partial band interference.

**SIMULATION OF A ROBUST LOCALLY OPTIMUM RECEIVER IN CORRELATED NOISE USING AUTOREGRESSIVE MODELING**

Myron R. Mychal

# The AR LO Detector

## Introduction

Many facets are involved in the design of a spread spectrum communications system [20]. One important consideration is determining a method to best recover the transmitted signal when it is subjected to interference in the transmission path. This interference is often highly correlated and not necessarily Gaussian as in typical interference models [12]. The linear correlator realization of the matched filter is no longer optimal for this environment. *Locally optimum* (LO) detection provides a method for circumventing this problem if the *probability density function* (pdf) of the interference is known. However, in cases where the interference exhibits strong self-correlation, traditional LO methods exhibit poor performance [24].

LO detectors with memory more successfully combat this type of disturbance. The disadvantage of a memory-based processor is the rapid increase in the dimensionality of the *joint probability density function* (jpdf) noise vector making the LO detector nonlinearity unwieldy [22]. To alleviate this problem, frequency domain methods are used to determine a $P^{th}$ *-order autoregressive* (AR(P)) model of the channel disturbance. The AR methodology reduces

the dimensionality of the underlying jpdf to that of the model order. The statistics of the jpdf remain an issue. To this end, pdf estimation techniques for *independent identically distributed* (iid) noise samples prove useful (11) since the input sequence to the AR model is iid, albeit characterized by some unknown pdf. The necessary AR model parameters are determined by well-known spectral estimation techniques [12]. The following sections describe the basic LO detector, develop the *autoregressive locally optimum* (ARLO) detector, discuss the simulation results, and present the conclusion.

# The Locally Optimum Detector

In communicating real, binary data in additive noise, the receiver must distinguish between two possible information signals. For example, in a *binary phase shift keyed* (BPSK) communications system, the receiver must decide whether a value of $\pm 1$ was transmitted.

More formally, the detector must choose correctly between one of the following hypotheses:

$$\begin{aligned} H_1&: \text{Signal } \mathbf{s}_1 \text{ is present} \\ H_0&: \text{Signal } \mathbf{s}_0 \text{ is present,} \end{aligned} \tag{1.1}$$

where the notation $\mathbf{x} = [x_1 \cdots x_N]^T$ denotes a vector of length $N$ and $\mathbf{s}_1$ and $\mathbf{s}_0$ represent the two possible signal vectors. The value of $N$, in general, denotes the number of signal samples for a given symbol (for BPSK, a symbol is represented by a bit). To derive the corresponding LO detector with memory, begin by observing the received signal vector as:

$$\mathbf{r} = \mathbf{s}_m + \mathbf{n}, \tag{1.2}$$

where $m = 0$ or 1 and $\mathbf{n}$ is a random noise vector, with jpdf, $f_{\mathbf{n}}(\boldsymbol{\eta})$. Given the observation, $\mathbf{r} = \boldsymbol{\rho}$, the optimum detector is of the form,

$$L(\rho) \underset{\text{choose } H_0}{\overset{\text{choose } H_1}{\gtrless}} \tilde{\gamma}. \tag{1.3}$$

where $\tilde{\gamma}$ is the decision threshold and is zero for *maximum likelihood* (ML) detection. In Eq. (1.3), $L(\rho)$ is the log-likelihood ratio, given by,

$$L(\rho) = \ln\left[\frac{f_{\mathbf{r}}(\rho \mid H_1)}{f_{\mathbf{r}}(\rho \mid H_0)}\right] = \ln\left[\frac{f_{\mathbf{n}}(\rho - \mathbf{s}_1)}{f_{\mathbf{n}}(\rho - \mathbf{s}_0)}\right], \tag{1.4}$$

where $f_{\mathbf{r}|\mathbf{H}}$ is the jpdf of the received signal under hypothesis $H$. Using a first-order Taylor series expansion around the signal points, and assuming a large *interference-to-signal* (ISR), the log-likelihood ratio can be approximated by $L(\rho) \approx l(\rho)$ [12] where,

$$l(\rho) = \sum_{i=1}^{N} (s_{1i} - s_{0i}) g_i(\rho). \tag{1.5}$$

The $s_{ji}$ in Eq. (1.5) represents the $i^{th}$ sample of the $j^{th}$ signal and

$$g_i(\rho) \triangleq -\frac{\partial}{\partial \rho_i} \ln[f_{\mathbf{n}}(\rho)] \equiv \frac{-\frac{\partial}{\partial \rho_i} f_{\mathbf{n}}(\rho)}{f_{\mathbf{n}}(\rho)}, \tag{1.6}$$

is the LO nonlinearity. Thus, the LO detector for the detection of known binary signals in additive noise is,

$$l(\rho) = \textstyle\sum_{i=1}^{N} (s_{1i} - s_{0i}) g_i(\rho) \underset{\text{choose } H_0}{\overset{\text{choose } H_1}{\gtrless}} \tilde{\gamma}. \tag{1.7}$$

# The LO Detector for AR Noise of Order $P$

Suppose the noise sample, $n_i$, at a discrete time instant, $i$, can be modeled as a $P^{th}$-order Markov process. Then the autoregressive noise model is [24]

$$n_i = \begin{cases} \sum_{j=1}^{P} a_j n_{i-j} + w_i, & \text{for } i \in [1, N], \\ 0, & \text{for } i \leq 0 \end{cases} \tag{1.8}$$

where the $\{a_i\}$ are the AR($P$) coefficients, $w_i$ represents an iid random process, and the vectors contain $N$ samples. In this case, the noise pdf is,

$$f_{\mathbf{n}}(\boldsymbol{\eta}) = f_{\mathbf{n}}(\eta_1, \cdots, \eta_N) = \prod_{i=1}^{N} f_{n_i}(\eta_i \mid \eta_{i-1}, \cdots, \eta_{i-P}), \tag{1.9}$$

where,

$$\begin{aligned} f_{n_i}(\eta_i \mid \eta_{i-1}, \cdots, \eta_{i-P}) &\triangleq f_{n_i}(\eta_i) && \text{for } i = 1, \\ f_{n_i}(\eta_i \mid \eta_{i-1}, \cdots, \eta_{i-P}) &\triangleq f_{n_i}(\eta_i \mid \eta_{i-1}, \cdots, \eta_1) && \text{for } i = 2, \cdots, P. \end{aligned} \tag{1.10}$$

If a "block approximation" [12] [1] is assumed, then $f_{\mathbf{n}}(\boldsymbol{\eta})$ can be written as,

$$f_{\mathbf{n}}(\boldsymbol{\eta}) = \prod_{i=1}^{N} f_w \left( - \sum_{j=0}^{P} a_j \eta_{i-j} \right), \tag{1.11}$$

where $f_w(\omega)$ is the pdf of the white noise process and $a_0 \triangleq -1$. For this case, the form of the nonlinearity is,

$$g_i(\boldsymbol{\nu}) \triangleq \sum_{l=0}^{\min(P, N-i)} a_l h' \left( - \sum_{j=0}^{P} a_j \nu_{i+l-j} \right) \text{ for } i \in [1, N], \tag{1.12}$$

where,

[1] In general, $\eta_1$ depends on the last "$P$" previous samples, which are assumed to be zero so that, $\eta_1 = w_1$. This only affects a small fraction of the terms for $N \gg P$.

$$\nu_k = \begin{cases} 0, & \text{for } k \leq 0, \\ \rho_k & \text{for } k \in [1, N], \end{cases} \tag{1.13}$$

$h'(\omega)$ is the derivative of $h(\omega)$, and

$$h(\omega) \triangleq \ln [f_w(\omega)]. \tag{1.14}$$

# Histogram and AR Estimation Techniques

Essential to the computation of the LO detector nonlinearity is the noise pdf, $f_w(\omega)$, and the autoregressive coefficients, $\{a_i\}$. Since neither the source statistics of the driving white noise, nor the AR coefficients are known *a priori*, they must be estimated.

The estimation of the pdf is performed via a histogram method employed in the LO detector without memory [12]. Pure noise data is not available, so the received data must be used. Then, a three-point derivative is used to to compute $h'(\rho)$. The histogram approach is chosen because of its ease and simplicity of implementation and generally acceptable performance.

The AR coefficients are estimated using the *modified covariance algorithm* (MCA) [5]. This is a non-windowing method similar to the covariance method, but it differs from the latter in that it minimizes the sum of the squares of the forward and backward predictor errors. The MCA is chosen since, when compared to other AR estimation methods, it often provides a) stable, high resolution spectral estimates; b) exhibits lower sensitivity to phase and decreased peak shifting, and, c) is not subject to spectral line splitting.

## Simulation Results

Simulation of an ARLO detector in a spread spectrum communications system comprised of a transmitter, channel, and (assumed) synchronized receiver was performed. In this system, the transmitter emits $B$ BPSK symbols sampled at $N_b$ samples per symbol. This sample waveform is then multiplied by a *pseudo-noise* (PN) sequencer with a chipping rate of $N_c$ samples per chip. Thus, the SS system processing gain is $PG = N_b/N_c$. A scale parameter for the corresponding *Signal-To-Noise Ratio* (SNR) is used to generate probability of error curves.

The channel is assumed to be an *Additive White Gaussian Noise* (AWGN) channel that is additionally corrupted by a self-correlated interfering signal which is either a *continuous wave* (CW) interferer or a *partial band* (PB) interferer. The parameter for the CW interferer is its frequency whereas for the PB interferer the passband is specified. Both interferers are scaled to a pre-determined ISR. The receiver is assumed to be synchronized with ideal filtering at the front end. The received signal samples are then passed to the ARLO detector.

The ARLO detector estimates the statistics of the received signal using a histogram method [9] to obtain an approximation of the pdf. A three-point derivative of the likelihood function, taking advantage of the properties of natural logarithms, is made. Next, an estimate of the AR model parameters is obtained using the modified covariance algorithm [5] with $P$ chosen as a parameter. The received signal is then filtered using the $\{a_i\}$ as a *finite impulse response* (FIR) filter to whiten the spectrum. Then, Eq. (1.12) is applied to this whitened signal to determine the LO nonlinearity. From this, a decision is made using a standard PN decoder.

In the CW case, the interference is modeled as a sinusoid at a given frequency within the main lobe of the signal spectrum. Typical spectra for

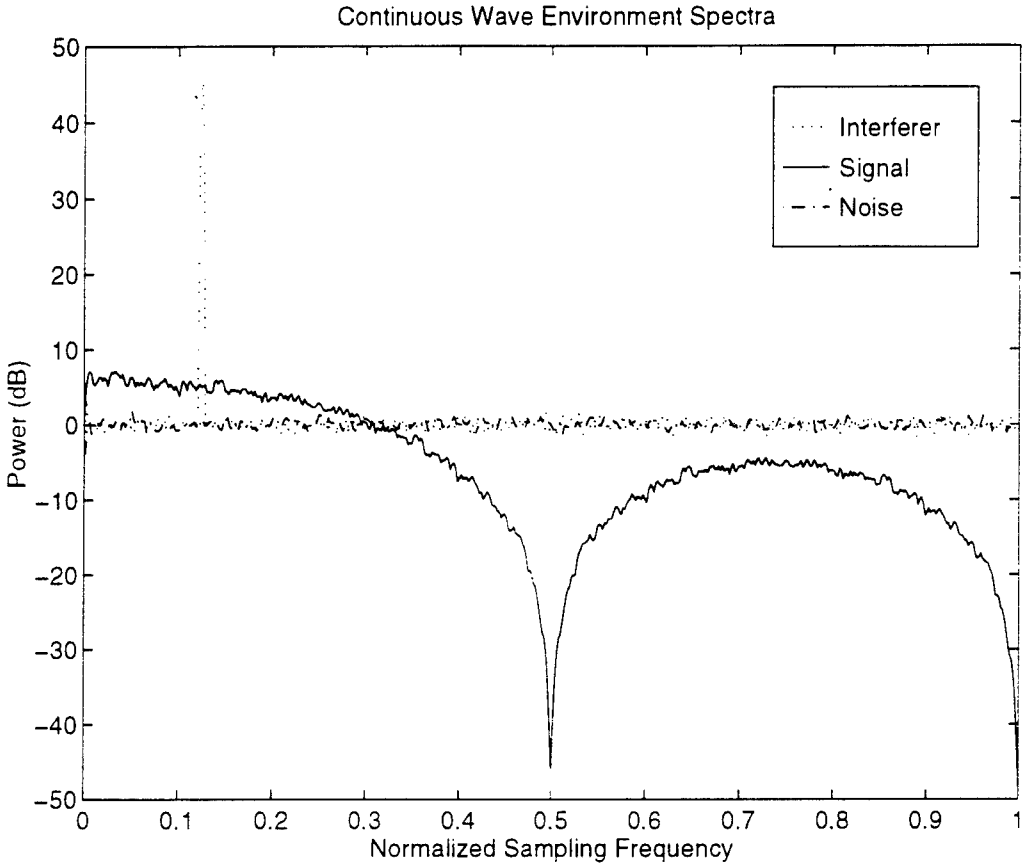



Figure 1.1: CW environment for the case of ISR = 20dB with PG = 12dB

this environment are depicted in Figure 1.1.

A typical bit error performance curve for this environment is shown in Figure 1.2, for $P = 16$. Bit error curves are provided for a standard linear correlator, the LO detector without memory, a linear correlator prefixed by a whitening filter, an LO detector prefixed by the same whitener and the ARLO detector. The whitening filters are created from the same AR coefficient estimates used in the ARLO detector.

In the PB case, the interference is modeled as an iid random process passing through an all-pole filter of pre-determined coefficients. Figure 1.3 depicts typical spectra in the CW environment.

A typical bit error performance curve for the PB environment is shown

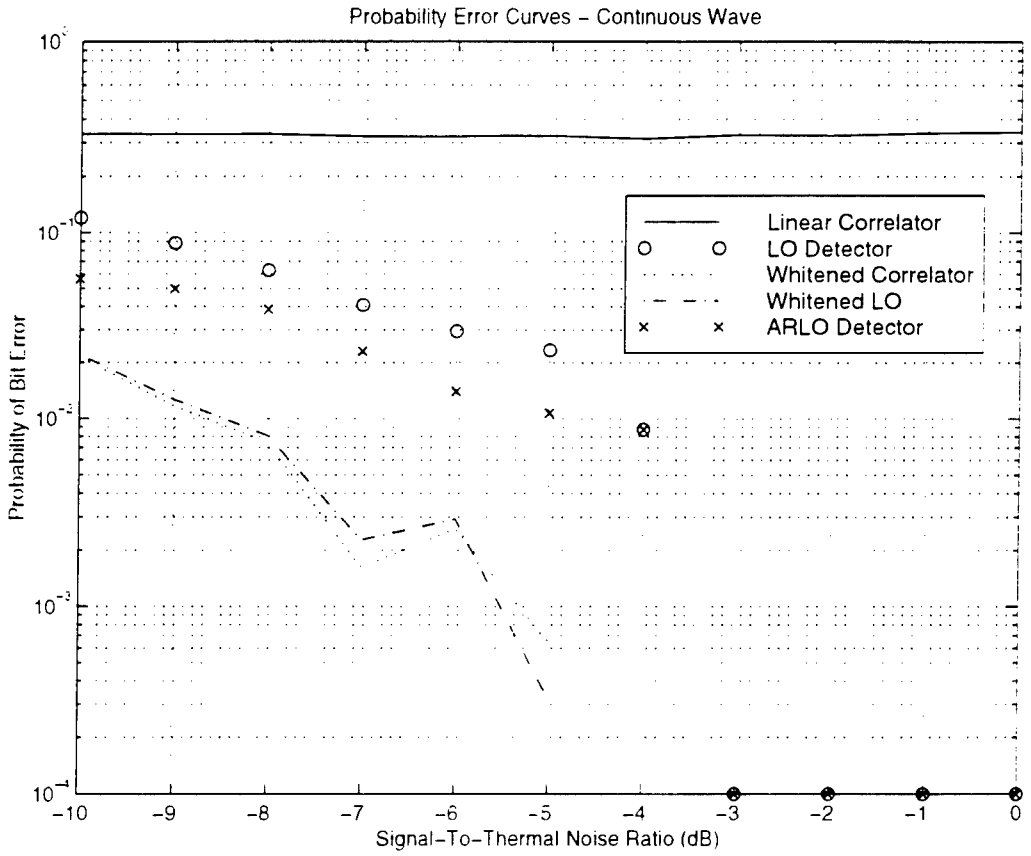



Figure 1.2: CW environment for the case of ISR = 20dB with PG = 12dB

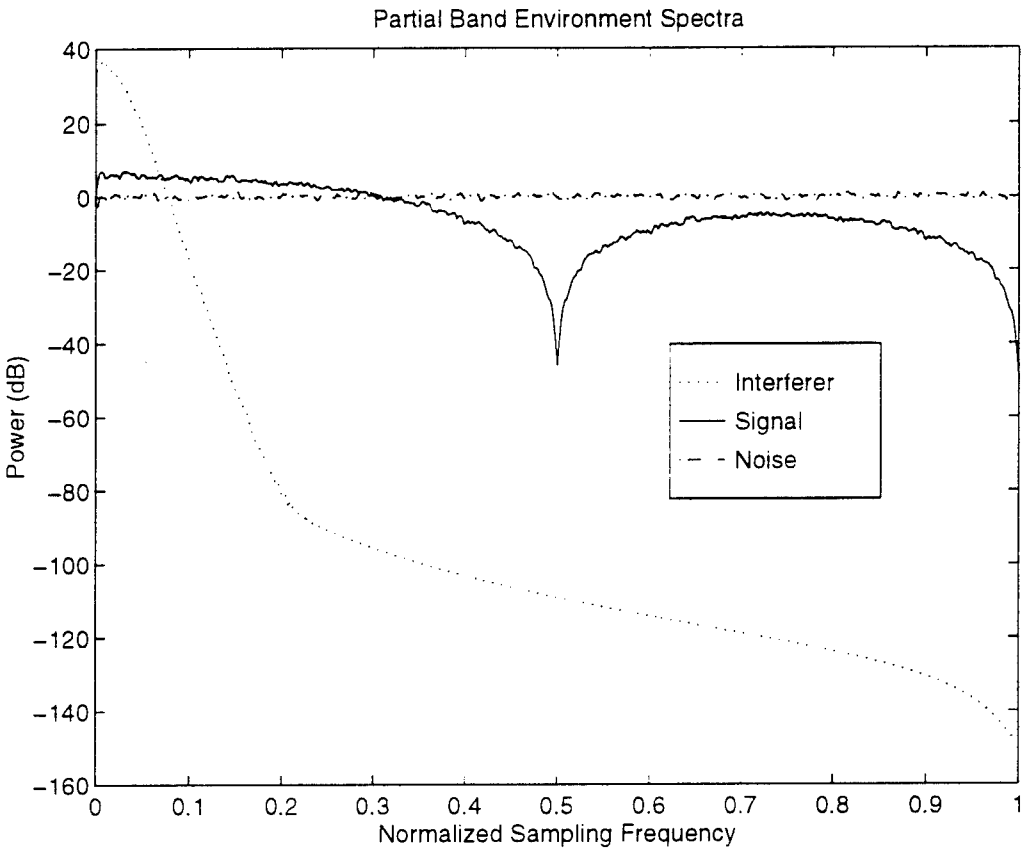



Figure 1.3: PB environment For the case of ISR = 20dB with PG = 12dB

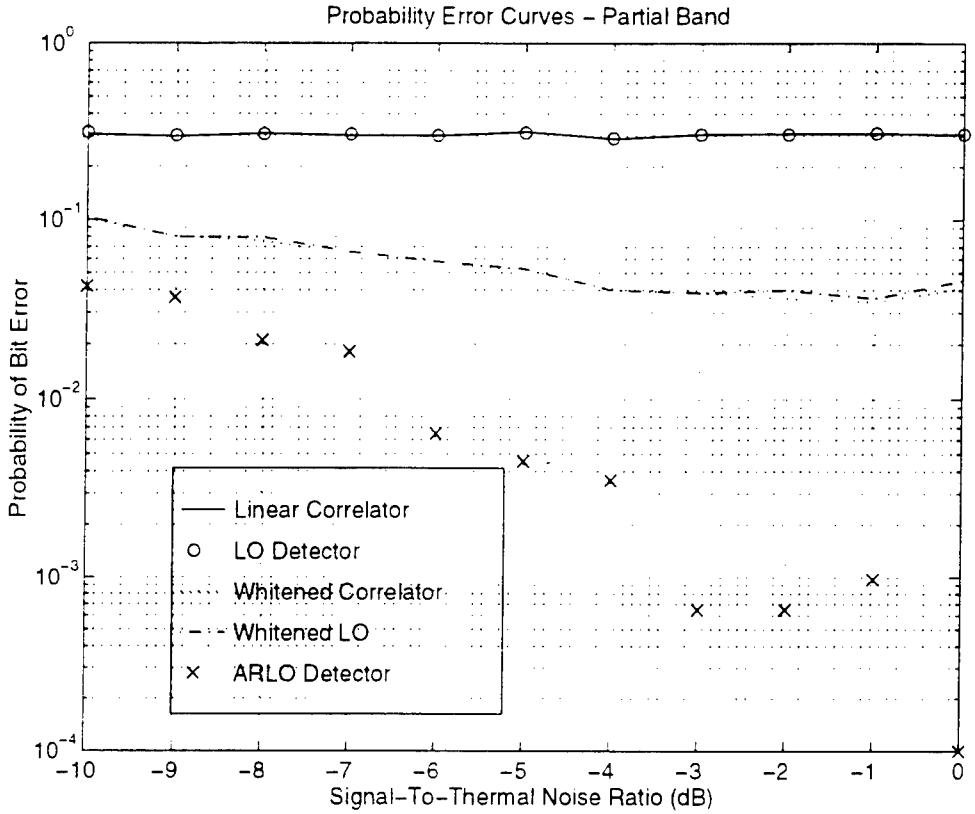



Figure 1.4: PB environment for the case of ISR = 20dB with PG = 12dB

in Figure 1.4. The parameters for this particular observation are $P = 16$ AR coefficients for the model and a pass band roughly equivalent to half of the main lobe of the signal spectrum. The performance for the same five detectors studied in the CW environment are illustrated.

# Conclusions

Analyzing the bit error curves in the previous figures for CW interference, we see that auto-regressive techniques do not seem to provide an advantage over traditional LO detection methods. In fact, it appears that whitening filters provide slightly better performance. However, the difference in performance is not significant. It is obvious that the linear correlator performs much more

poorly than any of the other techniques. These results match those obtained in [12] for similar environments.

However, in studying the cases for PB interferers, we note a dramatic increase in performance of the ARLO detector. For a given value of bit error probablity, the ARLO detector requires a significantly lower SNR than other LO detection schemes. Even the whitening methods cannot excise the wider-band interferer. Thus, the ARLO detector is more robust than any of the other receviers studied.

# Bibliography


[1] P.M. Clarkson, *Optimal and Adaptive Signal Processing*, CRC Press, Inc., 1993.

[2] P.M. Clarkson abd H. Stark, eds., *Signal Processing Methods for Audio, Images and Telecommunications*, London: Academic Press, 1995.

[3] J.F. Doherty and H. Stark, "Direct-Sequence Spread Spectrum Narrowband Interference Rejection Using Property Restoration," *IEEE Transactions on Communications*, vol. COM-44, pp. 1197-1204, Sep. 1986.

[4] J.H. Grimm, et. al., "Continuous Polynomial Approximation," *Proceedings of the 1993 IEEE MILCOM Conference*, pp. 283-287, 1993.

[5] M.H. Hayes, *Statistical Digital Signal Processing and Modeling*, New York: John Wiley and Sons, 1996.

[6] Hazeltine Report No. 6662, *Adaptive Nonlinear Coherent Processor Development*, Final Technical Report for Rome Air Development Center, USAF Report No. RADC-TR-89-387, 1990.

[7] J.H. Higbie, "Adaptive Nonlinear Suppression of Interference," *Proceedings of the 1988 MILCOM Conference*, pp. 23.3.1-9, 1988.

[8] D.M. Hummels and J. Ying, "Locally Optimum Detection of Unknown in Non-Gaussian Markov Noise," *Proceedings of the IEEE Midwest Symposium on Circuits and Systems*, pp. 1098-1101, 1991.

[9] W.E. Jacklin, J.H. Grimm, and D.R. Ucci, "The Simulation of a Two-Dimensional Spread Spectrum System With Locally Optimum Processing," *Proceedings of the 1993 IEEE MILCOM Conference*, pp. 288-292, 1993.

[10] W.E. Jacklin and D.R. Ucci, "The Fourier Series Implementation of a Locally Optimum Detector," *Proceedings of the 1994 IEEE MILCOM Conference*, pp. 992-996, 1994.

[11] W.E. Jacklin and D.R. Ucci, "A Comparison of Performance Metrics for Two Robust Locally Optimum Detectors," *Proceedings of the 1995 IEEE MILCOM Conference*, pp. 165-169, 1995.

[12] W.E. Jacklin, "Statistical Methods for Robust Locally Optimum Signal Detection," Ph. D. Dissertation , The Illinois Institute of Technology, Jul. 1996.

[13] S.A. Kassam, *Signal Detection in Non-Gaussian Noise*, New York: Springer-Verlag, 1988.

[14] J. Ketchum and J. Proakis, "Adaptive Algorithms for Estimating and Suppressing Narrowband Interference in PN Spread Spectrum Systems," *IEEE Transactions on Communications*, vol. COM-30 , pp. 913-923, May 1982.

[15] A.W. Maras, "Locally Optimum Bayes Detection in Ergodic Markov Noise," *IEEE Transactions on Information Theory*, vol. 40, pp. 41-45, Jan. 1994.

[16] J.L Melsa and D.L. Cohn, *Decision and Estimation Theory*, McGraw-Hill, Inc. 1978.

[17] D. Middleton, "Canonically Optimum Threshold Detection," *IEEE Transactions on Information Theory*, vol IT-12, pp. 230-243, Apr. 1966.

[18] L. Milstein and R. Iltis, "Signal Processing for Interference Rejection in Spread Spectrum Communications," *IEEE Acoustics, Speech and Signal Processing Magazine*, vol. 3, pp. 18-31, Apr. 1986.

[19] H.V. Poor, *An Introduction to Signal Detection and Estimation*, $2^{nd}$ Ed., New York: Springer-Verlag, 1994.

[20] D. Schilling et. al., "Spread Spectrum for Commercial Communications," *IEEE Communication Society Magazine*, pp. 66-79, Apr. 1991.

[21] D.R. Ucci, W.E. Jacklin, and J.H. Grimm, *Investigation and Simulation of a Nonlinear Processor for Spread Spectrum Receivers* , Final Technical Report for Laboratory, USAF Report No. RL-TR-93-258, 1993.

[22] D.R. Ucci, W.E. Jacklin, J.H. Grimm, *A Spread Spectrum Receiver With Nonlinear Processing*, Final Technical Report for Rome Laboratory, USAF Report No. RL-TR-93-50, 1993.

[23] D.R. Ucci, W.E. Jacklin, and J. Tamas, *Quasi-Optimal Processing in Spread Spectrum Environments* , Final Technical Report for Rome Laboratory, USAF RL Contract No. F30602-93-C-0099, 1994.

[24] D.R. Ucci and W.E. Jacklin *Robust Locally Optimum Detection*, Final Technical Report for Rome Laboratory, USAF RL Contract No. F30602-93-C-0099, 1995.

**M. Nagy**
**Report not available at time of publication.**

# A TECHNIQUE FOR LOCATING AND CHARACTERIZING CRYSTALLINE REGIONS IN SIMULATED SOLIDS

C Matthew Palmer
Graduate Student
Department of Physics

The George Washington University
Washington, DC 20052

Final Report for:
Graduate Student Research Program
Rome Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling AFB
Washington, DC

and

Rome Laboratory

August 1997

# A TECHNIQUE FOR LOCATING AND CHARACTERIZING CRYSTALLINE REGIONS IN SIMULATED SOLIDS

C Matthew Palmer
Graduate Student
Department of Physics
The George Washington University

Abstract

Software was developed which examines simulated solids and searches for crystalline regions in them. This is accomplished by forming a set of three lattice vectors, and using them to determine which atoms in the sample should be included in a particular crystalline region. Software designed in this manner can detect crystalline regions of arbitrary symmetry, and can even be made to accommodate lattices which are imperfect. The lattice vectors can also be used to determine if a particular crystalline region is of one of the common lattice types (FCC, BCC, etc). This information provides a useful means of characterizing the results of molecular dynamics simulations.

# A TECHNIQUE FOR LOCATING AND CHARACTERIZING CRYSTALLINE REGIONS IN SIMULATED SOLIDS

C Matthew Palmer

## I. Introduction

In analyzing the results of molecular dynamics simulations, one of the things that is of interest is to determine the size, extent, and nature of the crystalline regions which form in simulated solids, as well as the grain boundaries that occur between them. Because molecular dynamics is a numerical technique which requires that the positions and velocities of every individual atom in the sample be known, such an analysis can be performed through the use of lattice vectors and the positions of the atoms in the sample.

In Section II, the principles used in the software are explained. Section III explains in some detail the procedures used by the program to perform a search for crystalline regions. Section IV presents some information regarding analyses that have already been conducted using this software, and Section V presents some final conclusions regarding this software, as well as indicating means of obtaining the source code and executable files, for those who may be interested.

## II. Discussion of Problem

Crystal lattices are typically defined as repeating, periodic arrays of atoms whose positions can be characterized by a set of three lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$, as follows:

$$\mathbf{r} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3$$

where $n_1$, $n_2$, and $n_3$ are integers. By allowing the three n's to take on different integer values, one can produce the required periodic form. There are more complicated crystal lattices which have more than a simple one-atom basis and require additional vectors to characterize them, but we shall deal with such lattices later in this section.

In addition to these lattice vectors (the "direct lattice") there is also a set of "reciprocal lattice" vectors, which are important in the propagation of waves along the lattices. There exists a general prescription for determining the reciprocal lattice vectors, which is as follows:

$$\mathbf{b}_1 = \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

$$\mathbf{b}_2 = \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

$$\mathbf{b}_3 = \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} .$$

Usually, these vectors are defined with a factor of $2\pi$ in front of them. However, for the purposes of this program, this factor is unnecessary and will be neglected. The reciprocal lattice vectors defined above have a very useful orthogonality relation involving the direct lattice vectors:

$$\mathbf{a}_i \cdot \mathbf{b}_j = \delta_{ij} ,$$

a property which can be easily verified.

This relation provides us with a way to decide whether or not a particular atom is a part of the crystalline region characterized by the three lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$. If an atom is a part of the region in question, then its position vector will have the form $\mathbf{r}$ indicated above. If one takes the dot product of such a position vector with the three reciprocal lattice vectors, the following results are obtained:

$$\mathbf{r} \cdot \mathbf{b}_1 = n_1$$

$$\mathbf{r} \cdot \mathbf{b}_2 = n_2$$

$$\mathbf{r} \cdot \mathbf{b}_3 = n_3$$

Hence, if our atom is a part of the region in question, all three dot products will result in integers of some sort, whereas if the atom is not a part of the region, the results of one or more of these dot products will be non-integral.

As was mentioned earlier, there are some lattices, such as the hexagonal close-packed (HCP) and diamond lattices, which cannot be characterized simply by three lattice vectors. For these sorts of structures, one must specify not only the three lattice vectors but the positions of two basis atoms as well. In the case of the diamond lattice, for example, the lattice vectors are taken to be those of the face-centered cubic (FCC) lattice with basis atoms whose positions, relative to the FCC lattice sites, are given to be

$$\mathbf{0} \text{ (the null vector)}$$

and

$$\frac{a}{4}(\mathbf{i}+\mathbf{j}+\mathbf{k})$$

where a is the length of the side of the FCC cube and **i**, **j**, and **k** are the unit vectors in the x, y, and z directions, respectively.

Thus, any program which seeks to perform the desired analysis needs to do a variety of things. It must determine a set of lattice vectors, decide if there is a second basis atom, and then use that information to search the sample and determine which atoms should be included as a part of that crystalline region. The program also needs to be able to perform this procedure repeatedly throughout the sample, since samples of any great size will tend to be polycrystalline.

## III. Methodology

In the previous section, the principles involved in searching for crystalline regions were explained. What follows is a more detailed discussion of how these principles are implemented in the developed software.

The first step in the search procedure is to create a list of near neighbors for each atom. This is done for two reasons. First, it speeds the search procedure a great deal, particularly for large samples. Secondly, it allows the program to accommodate slow twists or bends in a crystalline region. Once the program has formed the three lattice vectors, it will search out lattice sites by performing a series of passes. In the first pass, the computer scans the neighborhoods of four atoms—the core atom and the three atoms used to form the three lattice vectors. Each neighborhood is scanned in turn, and the position vector **r** is formed with each of the neighboring atoms relative to the atom which "owns" the neighborhood. Each atom in the four neighborhoods is tested in the manner described in Section I, and those which satisfy the test then have their neighborhoods scanned, and so on. By doing this expanding search, gradual changes in the orientation of a particular crystalline region can be accomodated.

The neighbors array is formed by first sectioning the sample into a series of boxes. The boxes are constructed such that there are nearly as many boxes in the sample as there are atoms, so that there should

only be a few atoms, at most, in any one box. One can then form a list for each atom of atoms in nearby boxes.

Having done this, the program next searches through each of the atoms in the sample, trying to use each one as a "core" atom around which to form the three lattice vectors. To find the first lattice vector, the program scans the neighbors of the core atom, and forms a candidate lattice vector with each of the neighboring atoms (see Figure 1).

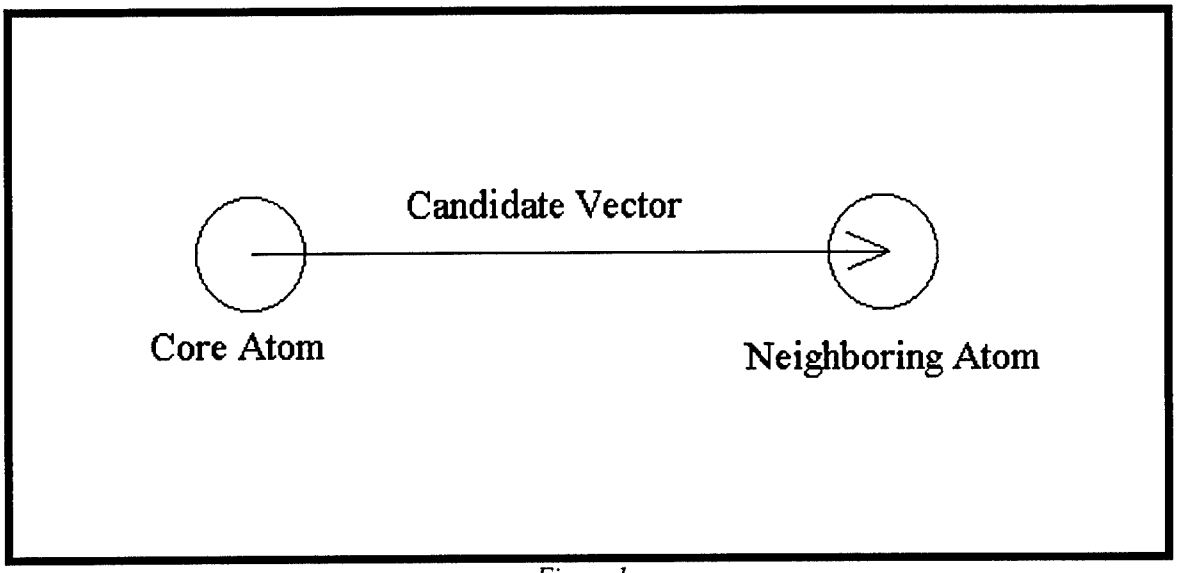



*Figure 1*

Having found a candidate vector, $\mathbf{a}_1$, in this manner, the program then scans the neighborhood of the core atom for an atom located at $-\mathbf{a}_1$, so that the three atoms form a straight line. This is done to insure that the atoms used to form the lattice vector are really sitting at lattice sites, and are not just secondary basis atoms, such as those found in the diamond structure, or some sort of defect or interstitial atom. If such an atom can be found, the program will consider this a valid first lattice vector, and store it while it continues to scan the rest of the core atoms neighborhood. The candidate vector which had been previously stored will only be replaced if a shorter vector which satisfies all of the above criteria can be formed.

The search for the second lattice vector proceeds in much the same way as the search for the first, except that an additional condition must be imposed. It is important that the second lattice vector not be parallel (or anti-parallel, for that matter) to the first, so that the two form a plane and span two-dimensional space. If two vectors, $\mathbf{a}_1$ and $\mathbf{a}_2$, are parallel, then

$$\mathbf{a}_1 \cdot \mathbf{a}_2 = a_1 a_2 ,$$

while, if they are anti-parallel,

$$\mathbf{a}_1 \cdot \mathbf{a}_2 = -a_1 a_2 .$$

So, in order to reject any candidate vectors which are collinear with the first vector, the program introduces the following conditions on any candidate vector $\mathbf{a}_2$:

$$\mathbf{a}_1 \cdot \mathbf{a}_2 - a_1 a_2 \neq 0$$

and

$$\mathbf{a}_1 \cdot \mathbf{a}_2 + a_1 a_2 \neq 0$$

In practice, it is not sufficient simply to insure that the two previous expressions are non-zero, and one must actually exclude a small range of angles around the 0 and 180 degree regions to insure that the two lattice vectors actually span a plane. In the developed software, the excluded regions were set to 5 degrees around the parallel and anti-parallel directions, which seems to work well (see Figure 2).

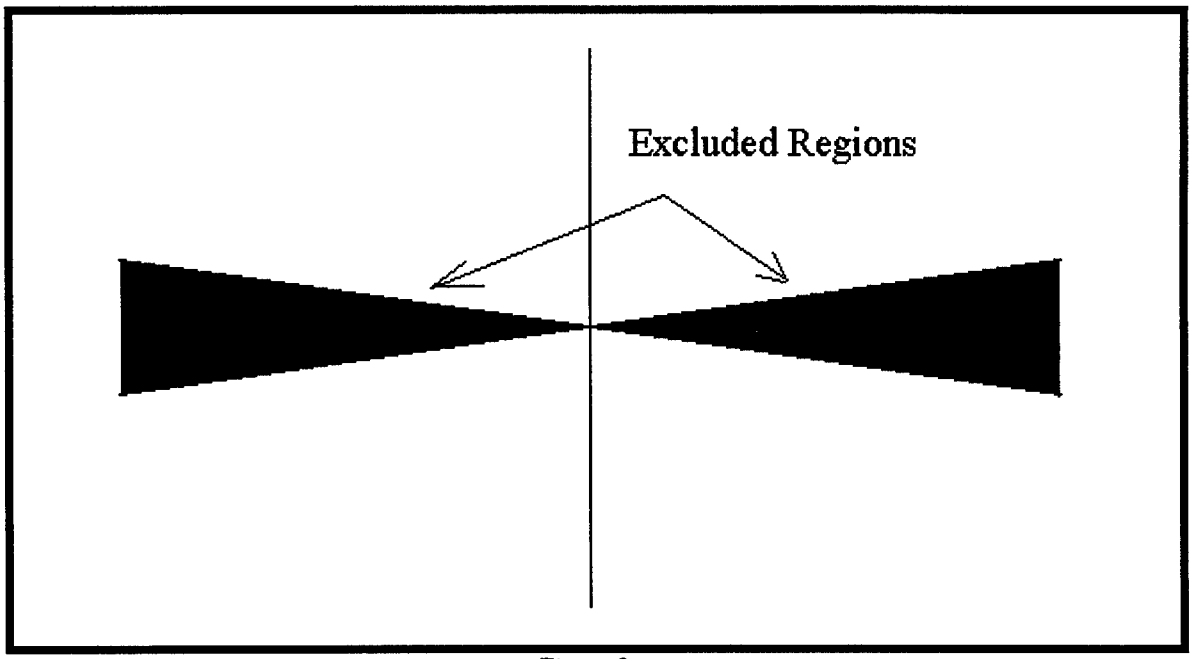



*Figure 2*

The search routine for the third lattice vector is again very similar, except that now a candidate for the third lattice vector must have a least some component which is not in the plane of the first two vectors, so that the set of three lattice vectors span three dimensional space. This is accomplished by insuring that

$$\mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2)$$

be non-zero. The cross product of the first two lattice vectors will yield a vector which is perpendicular to the plane they form (see Figure 3).

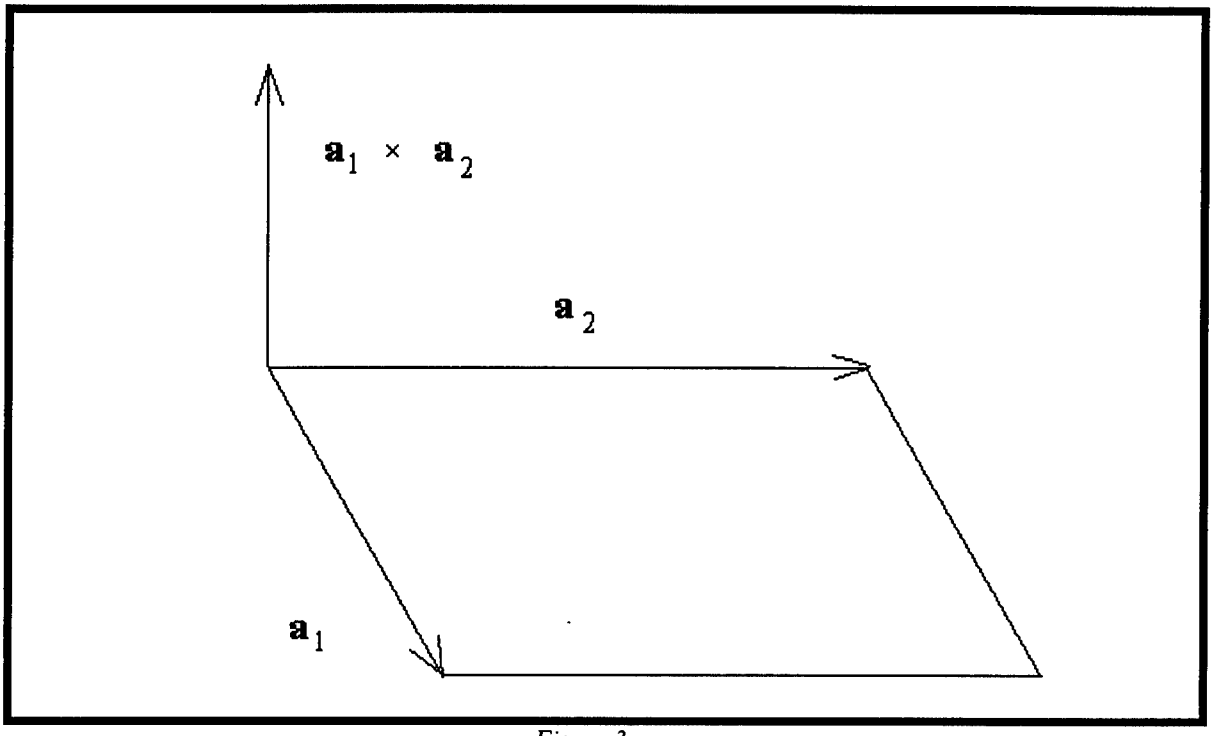



*Figure 3*

If we insure that any candidate for the third lattice vector has at least some component parallel or anti-parallel to this direction, then the three lattice vectors will in fact span all three dimensions. This can be guaranteed by requiring the triple product to be non-zero. As in the previous case, however, this proves inadequate in practice. In the developed software, this condition is modified somewhat so that the third lattice vector must form an angle of at least five degrees with the plane of the first two vectors.

This condition can also be viewed in another way. The absolute value of the triple product of three vectors can be taken to be the volume of the parallelepiped formed by the three vectors. If that volume were to go to zero, it would indicate that the parallelepiped had been squashed into a figure whose dimensionality was less than three, and thus the three lattice vectors used to form it would not span three-dimensional space.

It is conceivable that the computer will not be able to satisfy all of the criteria needed to form the three lattice vectors for a given core atom. If this is the case, then the program will move on to the next atom in the sample and try to use it as a core.

Once the computer has found a core around which it can form lattice vectors, it will attempt to make all three of the vectors as short as possible by adding or subtracting them. Due to imperfections in the simulated solids, the lattice vectors chosen by the computer are not always primitive, and this process tries to remedy that.

Having now formed the shortest possible set of three lattice vectors, the program then forms the reciprocal lattice vectors and looks for a second basis atom. Such an atom should be located within the cell formed by the three lattice vectors, and its position, relative to the core, would be given by

$$\mathbf{r} = f_1\mathbf{a}_1 + f_2\mathbf{a}_2 + f_3\mathbf{a}_3$$

where the f's are fractions between 0 and 1. The dot products between the three reciprocal lattice vectors and this position vector would yield three fractions:

$$\mathbf{r} \cdot \mathbf{b}_1 = f_1$$

$$\mathbf{r} \cdot \mathbf{b}_2 = f_2$$

$$\mathbf{r} \cdot \mathbf{b}_3 = f_3$$

So, in order to find such an atom, the program scans the neighborhood of the core and looks for an atom whose dot products are all fractions. If such an atom is found, a flag is tripped and the vector from the core atom to the second basis atom is stored.

There are many more complicated crystal forms which have a basis of more than two atoms. Since the software was designed primarily to analyze metals, it was not necessary to allow for these more complicated forms. However, it should not be difficult to modify the code to handle these forms if that should ever become necessary.

At this point, the program is ready to begin searching the sample for atoms which can be considered as part of this particular crystalline region. It does so by scanning the neighborhood of the core atom, forming the position vector of each of the neighbor atoms relative to the core atom, and then taking the dot products with the reciprocal lattice vectors, as described in the Section II. Ideally, if a neighbor atom were to be considered as part of the crystalline region, the three dot products would yield perfect integers. Usually, the dot products are not perfect integers, even for atoms that one would be inclined to say are part of a crystalline region, and so a tolerance parameter must be introduced. In the developed software, this parameter is an indication of how close each of the dot products must be to a perfect integer to be accepted, and thus can range from 0 (requiring a perfect match) to one-half, which would accept anything as a match , since if a result were more than .5 above a particular integer, it would be accepted as being less that .5 below the next highest integer, and so on. In the developed software, this parameter is

stored in a text file which is read by the program, and it may be easily changed by the user. This procedure is also performed on the three atoms used to form the lattice vectors.

Once this first pass is complete, the same procedure is performed with each of the newly selected atoms, and again with each of the newly selected atoms from that pass, and so on until there is a pass in which no new atoms are selected as part of the region. If it had been determined that there was a second basis atom, the program would then look in the neighborhood of each selected atom and try to find the second basis atom.

In developing the software, it was decided that it would be useful to allow the user to instruct the computer to ignore regions which were smaller than a specified size. This was done so that the user, when examining the reports generated by the program at the conclusion of its run, would not have to spend time sorting through information about small regions in which he or she was not really interested. At this point, the computer will determine how many atoms it has located at lattice sites and second basis atoms sites (if it has previously determined that there is a second basis atom). If the number of atoms in this region is found to be of significant size, the atoms making up that region are tagged with an identifying number. If the region is too small, nothing is done.

The program then repeat this search routine, attempting to use every atom in the sample as a core, and skipping over those which have already been selected as part of another region. When the program has eventually tried to use every atom in the sample as a core, it will end and send information regarding the crystalline regions it has located to a variety of output files. The developed software generates a set of four files based on the information it has generated. Two of these are data files which can be used by a visualization program. The third is a text report containing information about the direct and reciprocal lattice vectors and sizes of each region. The last file is a text file which is meant to be read by a second program and contains information about the direct and reciprocal lattice vectors. This second program examines the magnitudes of the three lattice vectors and the angles between them, and compares them to a library of common lattice types. This program generates a second text report which indicates which, if any, of the common lattice types each region appears to match.

## IV. Results

As has been mentioned, software has been developed which utilizes the preceding principles and procedures. This software has been applied to samples ranging in size from a few hundred atoms to about seventy thousand, and there is no reason that larger samples could not be analyzed in this manner. In the samples that have been analyzed so far, which have mostly been generated using the Lennard-Jones potential, most of the crystalline regions that have been recognized as some common type have been Face-Centered Cubic, with a few Hexagonal Close-Packed structures as well, which is not surprising, since the two forms are geometrically and energetically similar.

The following table gives a series of typical execution times for this program when analyzing files of different sizes on two different computers.

**Table 1: Processing Time vs Sample Size for Two Computers**

| Sample Size | 200 MHz Pentium PC | 250 MHz Ultra-Sparc Workstation | Ratio |
|---|---|---|---|
| 2217 | 5 sec | 3 sec | 0.60 |
| 9758 | 65 sec | 39 sec | 0.60 |
| 12709 | 95 sec | 54 sec | 0.57 |

The first two samples contained substantial regions of crystallinity, while the third was essentially an amorphous sample. It has been observed that the degree to which a sample has crystallized can have a substantial impact on processing time, even for samples of similar size.

## V. Conclusion

In this paper, a procedure for analyzing the results of molecular dynamics simulations has been presented. Software based upon this procedure has been used to analyze simulated solids, and appears to do so successfully, yielding information about the nature and extent to which a solid may crystallize under different conditions, such as temperature or annealing time.

The source code for the program was written in C. It has been successfully compiled on PC's and workstations operating under UNIX and Windows 95 (although it runs as a DOS application). Those wishing to obtain the source code should contact one of the following people:

Mr. C M Palmer (the author)

mattpalm@gwis2.circ.gwu.edu

MattPalmer@aol.com

Dr. H F Helbig

helbigh@rl.af.mil

Mr. J V Beasock

beasockj@rl.af.mil

The program currently runs as a command-line application. A true Windows 95 version will probably be developed over the next few months.

# System-Level Hardware/Software Partitioning of Heterogeneous Embedded Systems

D'Sunte L. Wilson
Graduate Student
Laboratory for Engineering Man/Machine Systems(LEMS)

Brown University
Box D
Providence, RI 02912

Final Report for:
Graduate Student Research Program
Rome Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, DC

and

Rome Laboratory

August 1997

# System-Level Hardware/Software Partitioning of Heterogeneous Embedded Systems

**D'Sunte L. Wilson**
**Graduate Student**
**Laboratory for Engineering Man/Machine Systems(LEMS)**
**Brown University**

## ABSTRACT

Industry and academia today are looking for ways to design hardware and software together without the rigors of manual trial and error. Hardware/software codesign is a method in which the best solution of the hardware and software are combined to find the maximum performance of a system. Hardware/software codesign needs to meet the complexity of the design of electronic systems today with its many different heterogeneous components. In this technical paper presentation, the focus is on hardware/software partitioning. Hardware/software partitioning is a crucial aspect in any codesign method, which finds the maximum performance of an embedded system. Embedded systems are specific systems which contain hardware and software components that are particular to a certain task. The performance requirements that are evaluated are power, area, and time for an embedded system. A CAD(Computer-Aided Design) tool, named COSH, will be introduced to accomplish this mission of finding an efficient combination of hardware and software units. The results show that the theory developed through COSH finds an efficient combination of hardware and software units given various integrated circuits.

# System-Level Hardware/Software Partitioning of Heterogeneous Embedded Systems

**D'Sunte L. Wilson**

## INTRODUCTION

### A. Motivation

Embedded systems are becoming an increasingly new area for the field of computer design. As more computer design designs are developed, whether it is a portable computer or an airplane systems used to fly from New York to Los Angeles, embedded systems are a key aspect in those designs. My motivation for this research stems for the search of how power dissipation can be conserved in embedded systems. The users of embedded systems can find more constructive uses of their systems by decreasing power consumption. [12] Embedded system CAD tools are not fully prepared to answer to the needs of the users, especially in the field of power consumption. Hardware/software codesign can be a viable solution to address the need to lower energy dissipation within embedded systems, along with the constraints of area and time. Figure 1 shows the current design process of hardware and software taken separately.

Arguably, the most important aspect of the codesign process is hardware/software partitioning. Hardware/software partitioning looks at the units of software and hardware and actually develops ways to combat the problems of area, timing, and communication. I am proposing to use this same hardware/software approach to find a lower power consumption within an embedded device. Hardware/software partitioning with respect to power consumption is an new area of research for academia and industry. Power

consumption within any electronic device must be addressed because it effects even the average person using a portable CD player.

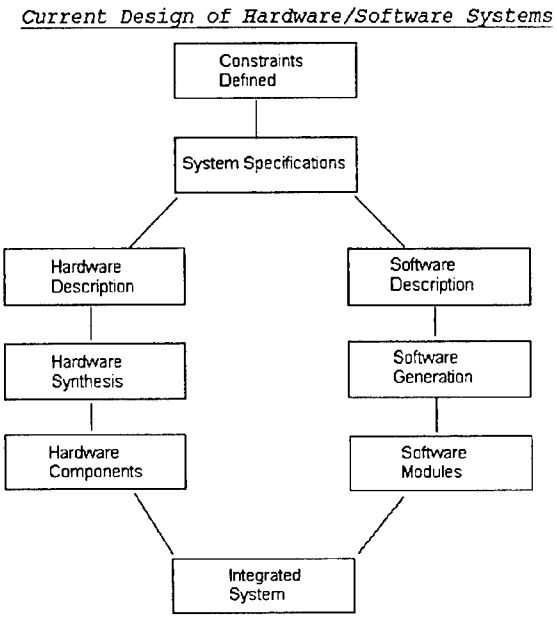



**Figure 1. Current Design of Hardware/Software Systems**

## B. Applications

Embedded systems have many applications to our everyday lives. They range from pagers to fly-by-wire fighter aircraft. Without the use of embedded system technology, many of the everyday computer items we use would be obsolete. These embedded systems, however, for personal use are often portable. When you have mobile computing involved, there are a range of issues that must be addressed. Arguably the most important

aspect of any portable device is power consumption. A mobile computer system, such as a wearable computer has many issues involved in its design with many heterogeneous parts. For any wearable computer, battery life is a definite issue that must be addressed. Wearable computers operators need to know that their device will be operational in a certain amount of time to properly deal with their activity. For any application where power consumption is an issue, the problem can be broken down into two main areas. These areas that hardware and software must deal with are time and cost. Time is a constraint when looking at the amount of performance that will be needed for a system to properly function. Cost refers to the area and the production costs of the system. These two key constraints affect the power dissipation of any embedded system.

## C. Objective

The investigation of the codesign problem will be approach through hardware/software partitioning in this paper. Partitioning is used to decompose a hardware/software system into several smaller elements to find an more efficient overall system by interchanging hardware and software elements. Partitioning is a key aspect in the codesign process. Without partitioning taking place, the system cannot achieve its maximum performance within the constraints. The existing research work done in the codesign area is limited because of the infancy of the research area. There are no set of instructions or theory when codesign is approached. Therefore, many different approaches are taken within the codesign research area.

Many take a biased approach toward the codesign process. If a software or hardware dominating approach is taken, it gives out two different results. My goal in this

paper is to look upon a non-biased approach towards hardware/software partitioning which will lead to a better overall solution. I will always look at any system as a whole instead of optimizing hardware and software components separately. This will be achieved by using COSH-DD. COSH-DD, standing for Codesign of Software and Hardware—Decision Diagrams, uses an approach where decision trees are used to optimize the embedded computer design. With the growing complexity of computer systems today, this partitioning technique can be made to be a great practical tool.

### D. Outline of paper

This paper will give the reader a chance to see the background , formulation, and results of this computer design problem. Section 2 addresses the background of the hardware/software codesign area with an emphasis on hardware/software partitioning. Section 3 addresses the problem formulation and the concept of decision diagrams, named COSH-DD to address the problem of hardware/software partitioning. Section 4 uses the results found to verify the method of using the COSH-DD algorithm. Section 5 concludes the paper and suggests new areas of further research.

## BACKGROUND

### A. Design of Computing Systems

Many computer developers have been designing software and hardware separately for many years. They have encountered problems of interfacing the two entities after they have already fully developed the hardware and the software. Kalavade points out that when designing computer systems many elements such as severe cost, performance, and design-time constraints have to be taken into account.[1] Many designers also overlook the essential nature of power in any design of a computing system. Computer designers are people who want to design and build real systems. They want to grasp the fundamentals, but also need to understand how things work in the real world. Thomas and Adams define a system-level approach towards computer design.[2]. Often times, the lower level is extremely well defined but for a developer who needs the top-level design and concerned with time-to-market a CAD tool must be able to provide for a system-level approach. The current approach towards designing a system requires endless testing and many man-hours to determine the "right fit" for a computer system.

Hardware/software codesign is an extremely diverse area. Many different approaches have been taken in the codesign field in dealing with computer systems. The theories range from the Princeton model of Malik and Wolf to Kalavade's Design Assistant. The reader should be aware that these theories show the complexity and the diversity in the hardware/software codesign research area.

The Princeton model looks at a general model using conservative estimates in its partitioning.[3] These conservative estimates give a great estimation, but they have serious flaws because the model is inflexible. Without flexibility given to the user, the

emphasis of a true design tool is lost. The quality that can be taken away from the Princeton model is the performance analysis on embedded systems. Its detailed information gives a clear understanding of how embedded systems should be evaluated.

The Princeton model describes new methods for the performance analysis of an embedded application execution on multiple processors. The algorithm can simultaneously design the hardware engine which consists of a network of heterogeneous processing elements, either CPUs or ASICs, and the application software architecture, which consists of allocating functions to processing elements in the hardware engine and scheduling their execution. The hardware engine and application software architectures together can be refer together as the embedded system architecture. The model uses this embedded system synthesis concept to codesign the hardware and software together to meet performance and cost goals. This method is efficient on the surface, but it lacks depth when dealing with specific issues of power, communication time, and delay.

Rajesh Gupta's model focuses on heterogeneous partitioning.[4] This approach also gives a more generalized approach from a hardware point of view. He uses HardwareC, a HDL(Hardware Description Language), to simulate hardware executions. The point of view taken by Gupta seems to be the most beneficial to the approach presented in this paper. Gupta explains that the heterogeneous computer chips and components are left without any clear-cut developing. We need to move towards a new way to design a computer systems without trial and error being used by the designer. However, we must still leave enough freedom for the designer to actually design his own custom computer system.

Gupta looks into the hardware/software partitioning aspects differently than Wolf. He views that the assignment of operations to a processor and to one or more application-specific hardware circuits involves additional delays due to communication overheads. Any good partitioning scheme must attempt to minimize communication overhead. Gupta points out that as the operations in software are implemented on a single processor, increasing the number of operations in software increases the degree of utilization of the processor. Consequently, the overall system performance is determined by the effect of the hardware-software partition on performance parameters that go beyond area and delay attributes.

The Lycos model, presented by Knudsen and Madsen, depends on too much human interaction.[5] This takes away from a hands off approach towards partitioning. There seems to be no clear cut distinction as to which is software and which is hardware when designing. This may lead to many problems in the design. Their approach, however, has great qualities. They use a library to define the hardware and the software used in evaluating the system. This method seems to be suitable for user interaction. It allows the designer to input what he will be using. The designer can then pick and choose amongst the library which systems are the best for his needs.

The Lycos system displays hardware/software partitioning in a different aspect than the previous two model. In the Lycos model, the user is given the choices to make the changes to the communication overhead, the software complexity, and specify a certain target architecture of hardware and software components.[6] This approach can be describe as functional partitioning. The problem with functional partitioning is that it relies heavily on user interaction. It takes away from the whole aspect of a codesign process.

The Lycos' CAD tool gives too much chance for error. The Lycos method while looks effective is basically the same as designing hardware and software together without any computer aid. The Lycos model does very little to effectively exploit the great potential for energy reduction through modification of software.

Kalavade's Design Assistant gives a more universal approach with a focus on the actual system level hardware units.[7] This approach gives the user a more hands on development on the design of the system from a system-level perspective. Integer linear programming is used to find the right combination of hardware and software. Kalavade's Design Assistant takes another look at hardware/software partitioning. She views hardware/software partitioning in two parts. One aspects is binary partitioning. Binary partitioning is the problem of determining, for each node, a hardware or a software mapping and a schedule. The second aspect is extending partitioning which describes the problem of selecting an appropriate implementation, over and above binary partitioning. Kalavade's theory of developing each task together probably is the most efficient out of all the previous models.

Jian Li's approach towards hardware/software partitioning is an approach where complex mathematical models are not presented.[8] He uses the TDT(Time Decision Tables) with Don't Care sets to minimize his solution. This approach alleviates the problem of mathematical models when using hardware/software partitioning. Even though this method may have its flaws, it is still easily understandable by any designer.

Tiwari's paper uses the actual instruction sets to get an approximation of power. The power consumption has not been deeply developed in hardware/software embedded systems when dealing with predesigned computer chips and circuits.[9] His approach of

seeing how each software instruction dissipates a certain amount of power opens up another area of research in the codesign field. This approach gives some motivation for considering power when using hardware/software partitioning.

In summary, two key aspects hold over all these theories. One aspect is that hardware is composed of a set of functional resources which can be accessed concurrently. The second is that hardware functional resources are already bound to software i/o operations. None of theories presented adequately use software to truly codesign their processes. Researchers often take different approaches to the theory for practical applications. Kalavade addresses many concerns in the codesign area, however leaves out the power consumption. However, Kalavade's insight is invaluable. Gupta and De Micheli look at the hardware/software codesign of all the issues involved also. Their definition of the codesign area really made the problem of hardware/software codesign somewhat easier to figure out .[10] De Micheli himself address the huge areas of research in the codesign area in his article.[11]

However, arguably the most important aspect of the codesign process is hardware/software partitioning. Hardware/software partitioning looks at the units of software and hardware and actually "codesign" them within each other . There are many concerns in the codesign process. These concerns are communication between hardware and software, area, timing issues, and power consumption.

Many researchers have touched upon this codesign theory. However, their approach is either from a software dominated approach or a hardware dominated approach. In a software dominated approach, the user tailors his needs with the view that changing the software executions or instructions can greatly reduce the overall cost of the

system. This theory, although correct on the surface, lacks depth because it rarely deals with the physical limitations that hardware can impose on any system. In a hardware dominated approach, the system's hardware is optimize through current techniques of technology mapping ,Binary Decision Diagrams(BDDs), graph models, or some other hardware approach of evaluating circuits. These approaches while valid, take into little consideration the complexity of a software algorithm and its instructions in terms of the overall cost. The goal is to have a "same" design approach when dealing with embedded systems. This will greatly reduce any error in determining the best solution.

**PROBLEM FORMULATION**

A CAD tool developed called COSH(Codesign of Software and Hardware) will attempt to alleviate the problem of hardware/software partitioning. Partitioning is a key step in any codesign theory. It is the step where the interchange and comparison of hardware and software units must be made. COSH will attempt to show an unbiased approach toward the codesign theory for practical use in developing a portable device. The information obtained from COSH will allow the developer to see how the software and hardware they are using will directly effect the energy consumption and daily operation of the device. This will allow the developer to make necessary changes without expensive last minute changes.

This paper will present a universal approach to hardware/software partitioning. Hardware/software codesign is a method where the best solution of the hardware and software of a system are combined to find the maximum performance of a system. This

method can have great applications when used on embedded systems. The method I propose is to look into hardware-software partitioning. Hardware-software partitioning is a subset of hardware-software codesign, but it is a crucial aspect of the process. Without efficient partitioning, the system cannot truly be minimized for maximum performance. My approach views hardware and software as the same "unit" or entity. This allows for more efficient partitioning because it alleviates the problem of software or hardware dependency. The CAD tool give a unique solution to a hardware-software system for maximum performance with minimal use of hardware and software components. There will be a library for hardware implementation functions and specified software implementation functions. There are many difficult aspects that must are determined through the use of this CAD tool. There are as follows:

1. Determining which qualities of software that will be evaluated.
2. Determining how the hardware should play a role.
3. The actual construction, i.e. definition, of the computer code.
4. The problem of treating the software and hardware as the same entities may cause issues in the overall evaluation.

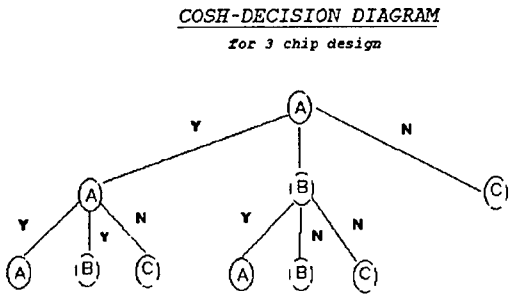



**Figure 2. COSH- Decision Tree Diagram**

## Description of the System Model

At the system start, the user is asked for the time, area, and power constraints along with the number of software instructions of the system. Figure 2 gives you a schematic view of the approach taken. The CAD tool checks whether or not the constraints are realistic with the given library The library is defined by with each chip having the characteristics of area, time delay, power, average clock cycles per instruction(CPI), and clock rate. This defined library gives the user the option to interchange its hardware and software requirements for a system design without too much interaction with the CAD tool. It allows the user to have hands-on development of a system, but allows the speed of the CAD tool to determine the right combination of chips.

After the constraints are determined, the first initial partition takes place with system specifications of the library and the entered constraints by the user. After the first split is made, the hardware and software components' maximum performance is decided. Each chip is checked to determine if they overstep the constraints. The time, area, and power constraints use Equation 1 to determine its feasibility:

$$\begin{aligned} &\textbf{Chip time } (t_{chip}) \leq \textbf{Time constraint } (T_c) \\ &\textbf{Chip area } (a_{chip}) \leq \textbf{Area constraint } (A_c) \\ &\textbf{Chip power } (p_{chip}) \leq \textbf{Power constraint } (P_c) \end{aligned} \qquad (1)$$

The power constraint, $P_c$ , has a safety factor of 10%. When the user enters the power constraint, the CAD tool takes only 90% of that value for a safety factor, K, given the nature of power supplies and systems. If the constraints are met for each chip defined in your library for time, area ,and power, we must then find the constraints met for the number of software instructions defined by the user. Equation 2 gives the relationship between time and software instructions given the elements in each library:

$$\textbf{Software Time } (t_{soft}) = \textbf{(\# of Instructions)(CPI)( 1/ Clock Rate)} \qquad (2)$$

The software time, $t_{soft}$ , is compared to the time constraint just as the chip time is compared to the same time constraint in Equation 1.

Once each individual chip meets all the constraints, the partitioning can continue by taken different combinations of chip designs and determining if these combinations meet system specifications. Equation 3 takes the area summations of each chip design and sees if they meet the area constraint:

$$\textbf{}\sum \textbf{Individual chip areas} \leq \textbf{Area constraint } (A_c) \qquad (3)$$

The power constraint with the 10% safety factor uses Equation 4 to meet its design requirements:

$$\sum \text{Individual chip power} \leq \text{Power constraint } (P_c) \qquad (4)$$

The time considerations differ from those present in Equations 3 and 4. Another library is defined where the communication time between each chip is properly defined. This library gives the user the ability to change the flexibility of the communication time between certain chips in the first library. The equation used for the hardware time consideration is given by Equation 5:

$$\sum \text{Individual chip times} + \text{communication time between chips} \leq \text{Time constraint}(T_c) \qquad (5)$$

For software time considerations , we take the software time defined in Equation 2 and uses it as:

$$\sum \text{Individual software times} \leq \text{Time constraint}(T_c) \qquad (6)$$

Equation 6 does not differ from Equation 3 and 4. The only equation that differs is Equation 5 which gives the communication times of each chip for an overall time consideration.

Each solution is determined through the use of each equation above. This method allows the user to make decisions based on the system-level requirements. It also gives the user different combinations of the determined systems in order for the user to make the determination of which system is the best. This partitioning allows for the user to have the final say in design. It cuts down time for pointless combinations in the design. The Decision Diagram method, shown back in Figure 2, is a system where the CAD tool helps user design the system, while the user helps the CAD tool also design the system. This

approach leads to an unbiased approach. The CAD tool does not allow the user to make uneducated decisions that can waste money, but it allows the user to manipulate a system for any specifications needed.

## RESULTS

The CAD tool COSH evaluated chip designs containing 6, 20, 50, and 100 chips. This chip designs were based on already existing chip information. The 6-chip system always was easily attainable. It met the qualifications of area, time, and power when it was designed based upon the information processed by COSH. The power safety, K, proved to be important in the 6-chip design as power fluctuated up to 3% based upon the values entered. The CPU time in determining a 6 chip system was not a factor.

The 20 chip system also used the same construction techniques as in the 6 chip system. It easily met the constraints of area and power, however in some combinations the timing in the actual construction of the 20 chip system was off up to 3%. This problem can be contributed to unpredictable nature of software and software timing. The power safety factor in this design was crucial as power fluctuated up to 5% based upon the values entered. In a 20 chip system, the CPU time in its construction was also nonexistent.

The 50 chip system used a predetermined chip construction of a mini-wearable computer device to determine its efficiency. Every evaluation given by COSH gave the exact same construction of the device, which took the original designers 8 months to determine. Just as the original designers still have problems with timing, COSH displayed

problems with the timing of the system-level chip design. The power requirement was easily met under the existing systems. The safety factor for the power constraints proved to be a great asset to this already existing design.

For the 100 chip system, an experimental cellular phone was use to determine COSH's effectiveness Again, the original took 13 months to construct, while COSH did the same construction in less than 30 seconds. The problem of timing showed up again in this evaluation. Upon further review, it seems that the software instructions vary the timing of system as it is processed from chip to chip. In some cases, the timing given in COSH and those in the actual construction were off by 20%. This seems to be an interesting results as COSH gives the same actual construction of the system evaluated.

The results given in this paper are in its infancy. Many more tests have to be taken to find out the problems facing the timing issues. However, the results given for small systems of 6 and 20 chips are accurate. The evaluation process yielded that the area and power requirements were always met without any error. The timing constraints showed differences as the system became bigger. This result can be seen as software instructions having a direct effect on the actual system design. The power safety factor accounted for software instruction power dissipation. The results given gave great expectations in developing a more efficient CAD tool that finds a unique, easily understandable, and efficient process.

## CONCLUSION

Hardware/software codesign is a concept which takes on many different approaches. The approach taken is a universal approach with limited human interaction. The partitioning aspect of the codesign concept is probably the most crucial aspect. If an efficient system can be found to reduce the time and cost of producing hardware/software systems, the computer industry and academia can put their efforts towards other concerns. At this point, I am in the infancy of my research in this new and developing area. I am confident that with research and proper resources that my universal theory will be a viable theory.

The drawbacks towards a universal approach can be overcome through continuous testing and finding new ways to explore different options. Once the drawbacks are overcome, such as timing, many embedded systems whether large or small can gain great strides in their development. My proposed theory will find the maximum performance between hardware and software. With the complexity of digital mixed hardware/software systems growing, a short-cut or time-saver must be found in their development. Once the problem of codesign is properly defined and solved, the computer field will grow from it.

## REFERENCES


[1] A. Kalavade and E.A. Lee, "A Hardware-Software Codesign Methodology for DSP Applications", *IEEE Design & Test of Computers*, vol. 10, no. 3, September 1993, pp. 16-28.

[2] J. Adams and D. Thomas, "The Design of Mixed Hardware/Software Systems", *IEEE/ACM Proc. Of 33rd Design Automation Conference(DAC)*, pp. 515- 520, 1996.

[3] S. Malik, W. Wolf, A. Wolfe, Y-T. Li, and T-Y. Yen, "Performance Analysis of Embedded Systems", *Hardware/Software Codesign,* McGraw Hill, pp. 45-74, 1996.

[4] R. K. Gupta, "Analysis of Operation Delay and Execution Rate Constraints", *IEEE/ ACM Proc. Of 33rd Design Automation Conference(DAC)*, pp. 601-604, 1996.

[5] P. V. Knudsen and J. Madsen, "Aspects of System Modeling in Hardware/Software Partitioning", Department of Information Technology, Technical University of Denmark, 1995.

[6] P. V. Knudsen and J. Madsen, "PACE: A Dynamic Programming Algorithm for Hardware/ Software Partitioning", Department of Information Technology, Technical University of Denmark, 1996.

[7] A. P. Kalavade, "System-Level Codesign of Mixed Hardware/Software Systems", Ph. D. Thesis, University of California at Berkeley, 1995.

[8] J. Li and R. Gupta, "Limited Exception Modeling and Its Use in Presynthesis Optimizations", *IEEE/ ACM Proc. Of 34th Design Automation Conference(DAC)*, pp. 341-346, 1997.

[9] V. Tiwari, S. Malik, and A. Wolfe, "Instruction Level Power Analysis and Optimization of Software", *Journal of VLSI Signal Processing*, Kluwer, pp.1-18, 1996.

[10] R. K. Gupta and G. De Micheli, "Hardware-Software Cosynthesis for Digital Systems", *IEEE Design & Test of Computers*, vol. 10, no. 3, September 1993, pp. 29-40.

[11] G. De Micheli, "Hardware/Software Co-Design: Application Domains and Design Technologies", *Hardware/Software Codesign,* McGraw Hill, pp. 1-28, 1996.

[12] G. Borriello, P. Chou, and R. Ortega, "Embedded System Co-Design: Towards Portability and Rapid Integration", *Hardware/Software Codesign,* McGraw Hill, pp. 243-264, 1996.

# SOFTWARE VERIFICATION GUIDE USING PVS

Luke J. Olszewski
Associate Professor
Department of Mathematics and Computer Science

Georgia Southern University
Post Office Box 8093
Statesboro, GA 30460

Final Report for:
Summer Research Program
Rome Laboratory

Sponsored by:
Air Force Office of Scientific Research
Bolling Air Force Base, Washington, DC

And

Rome Laboratory

August 1997

12-1

# Software Verification Guide
# Using PVS

**Luke J. Olszewski**
**Instructor of Mathematics**
**Department of Mathematics and Computer Science**
**Georgia Southern University**

## Abstract

This technical report will detail some of the syntax and semantics of the Prototype Verification System (PVS) used on several examples. It will include the specification and verification of the existence of a linear limit and a simple traffic light. A survey of several of the PVS prover commands and their applications will be taken to act as a guide to introduce Software Specification with PVS. Observations of employing PVS to aid in the co-design of complex systems will also be reviewed.

# Software Verification Guide
# Using PVS

**Luke J. Olszewski**

## 1. Introduction

The Prototype Verification System contains a specification language and a proof checker. It is a system that was developed at SRI International which verifies formal specifications and checks formal proofs [1]. As a whole, it is an interactive tool used as a theorem prover/proof checker where user-defined proof strategies enhance the automation of the proof checker. PVS is implemented in common Lisp with ancillary functions provided in C, Tcl/TK, and LaTeX, and uses GNU EMACS for its interface [1]. The basic deductive steps in PVS include atomic commands for induction, quantifier reasoning, automatic conditional rewriting, simplification using arithmetic and equality decision procedures, and propositional simplification using binary decision diagrams. The PVS proof checker prompts the user for appropriate commands at certain subgoals. The execution of such commands may produce further subgoals or complete the subgoal and move on to the next step. This allows the user to actively engage in the Formal Methods of specifying and proof checking a design.

There has been a push in this country to use Formal Methods (FM) in the design of complex computer systems throughout the 1990's. It is clear that there are many advantages of applying FM to both hardware and software applications at the beginning of a project [3], yet often there is little communication between the hardware and software developers at any stage. This gap in the process of designing systems can be branched if the project as a whole can be abstracted to a higher level where functions are specified independent of the apparatus that will perform specific operations. PVS can help in bridging the gap because of its dynamic employment of algebraic and arithmetic specification properties. Such implementations go beyond the scope

of this report, but will be discussed briefly.

This report will act as a guide that details some of the basic prover commands used in PVS. It is not intended to replace any of the PVS manuals, but may make using PVS less painful to the beginner. It assumes that the reader is familiar with FM and mathematical logic concepts from predicate calculus and is simultaneously using PVS manuals provided by the NASA Langley Research Center. Information and manuals can be obtained from their web site at **http://atb-www.larc.nasa.gov/fm.html**.

## 2. Syntax and Semantics

The manuals that can be obtained from the web site mentioned above are good in that they offer the user excellent examples, but it lacks the simplicity that is required for first time users of such tools. This section will detail and consolidate some of the information that this author believes was left out in those manuals. This survey is not meant to replace PVS documentation, but rather to shed light on some of the assumed knowledge by the authors of [1] and [2].

### 2.1 Some Basic Commands

We now assume that the reader has the appropriate manuals and has PVS up and running. Once the PVS proof checker is invoked, hit the Esc key ("ESC -" is displayed in the buffer at the bottom of the EMACS window) and the x key ("M-x" is displayed in the buffer), then type ff (find file) and hit the return key to find a file. PVS will prompt you to type the name of the PVS file, leaving off the .pvs extension. It will then look for the file *.pvs. If no such file exists, the user must hit the enter key twice to get back to the PVS Welcome buffer and type M-x nf (ESC, x, and nf for new file). Type the name of the new file without the extension and press enter, and PVS will create the *.pvs file for you. Now the formal specification may be typed in, or open a specification that came with the PVS package. When the formal specification is completed, typecheck the file to determine if any Type Check Conditions (TCC's) are generated and let PVS attempt to prove them with M-x tcp (typecheck and prove). PVS will parse (check for correct syntax) and typecheck (check, for example, that real functions don't map to booleans, etc.) the specification. When completed without errors, the cursor will go back to the PVS file. Move the cursor to the lemma or theorem to be proved and hit Ctrl+c keys ("C-c -" will be displayed in the bottom buffer), then hit the p key. The PVS proof checker

will automatically be invoked and the theorem or lemma that the cursor was on will be displayed with "Rule?" beneath it. (If the user is in the X-Windows environment, all of these commands can be found in the menus at the top of the window.)

Now enter the first prover command enclosed in ()'s (e.g. (skosimp*)). The following commands must be within parenthesis since the proof checker uses Lisp. Also, IMPLIES and => are used interchangeably, along with & and AND. The specification commands are not case sensitive.

Some basic commands for the Theorem Prover are listed below [2]. The syntax as it would appear during the proof checking is: Rule? (COMMAND).

(**expand** "NAME") Below is an example where abs, though not highlighted when typed in, is a defined function in the PVS prelude file.

```
{-1} abs(x!1 - a) < delta
  |-------
{1}  abs(f(x!1) - L) < eps
```

Applying (expand "abs") yields

```
{-1}    (IF x!1 - a < 0 THEN -(x!1 - a) ELSE x!1 - a ENDIF < delta)
  |-------
{1}     (IF f(x!1) - L < 0 THEN -(f(x!1) - L) ELSE f(x!1) - L ENDIF < eps)
```

Performing a rewrite instead of an expand allows the user to control which abs will be expanded. In other words, applying rewrite instead of expand would result in the expansion of only the precondition, {-1}.

(**flatten**) simplifies disjunctive (OR)

(**grind**) Performs the following as specified in the training manual and when executed: skolemizes, instantiates, and lift-if's. It also rewrites prior to skolemization and asserts when completely simplified. See the limit example below.

(**ground**) Invokes built in descision procedures, splits, flattens, and appears to assert when done. For example:

```
[1] IF a THEN b
```

(ground) yields:

```
{-1} IF a
  |-------
{1} THEN b
```

(**help** NAME) e.g. (help split) displays text for command "split".

(**inst?**) Tries to find an appropriate variable substitution

```
{-1} (FORALL (x: arb): P!1(x) => Q!1(x))
  |-------
{1} Q!1(x!1)
```

Applying (inst?) results in *x*!1 substituted in for *x*.

```
{-1} P!1(x!1) => Q!1(x!1)
  |-------
{1} Q!1(x!1)
```

(**lift-if**) Applies a function to a conditional statement.
e.g. *f*(If *a* THEN *b* ELSE *c* ENDIF)
(lift-if) yields: IF *a* THEN *f*(*b*) ELSE *f*(*c*) ENDIF

(**lemma** "LEMMA's NAME") Used to call a proven or imported theorem, lemma, or axiom.

(**postpone**) Postpones the proof of a subgoal. Helpful if some subgoals can be proved and others cannot. A comparison of the proven with the unproven may help in determining any holes in the specification.

(**prop**) Invokes several lower level rules to carry out a proof without showing the intermediate steps. It simplifies propositional statements into its axioms.

(**quit**) Leaves the proof checker. It cannot be used during the proof check.

(**rewrite** "NAME") Similar to the expand command.

(**skolem!**) Skolemization is used to eliminate the universal forall and the existential there exists quantifiers. Quantifiers are either changed into either skolem constants (universal quantification) or skolem terms (existential quantification). Existential quantification may be

instantiated with user selected terms (see example below).

```
(FORALL (p:bool, q: bool): ((p => q) AND p) => q)
```

(skolem!) yields

```
{1} ((p!1 => q!1) AND p!1) => q!1
```

(**skosimp***) Flattens and skolemizes repeatedly until simplified. Often skolemizing is done as the first step.

(**split**) Generates subgoals from IF THEN statements. If there are several, the (split) command may be used several times in succession.

```
{1} ((p!1 & q!1) & r!1) => (p!1 & (q!1 & r!1))
```

(flatten) yields:

```
{-1} p!1
{-2} q!1
{-3} r!1
  |-------
{1} (p!1 & (q!1 & r!1))
```

(split) yields 3 subgoals:

```
1.1
{-1} p!1
{-2} q!1
{-3} r!1
  |-------
{1} p!1

1.2
{-1} p!1
{-2} q!1
{-3} r!1
  |-------
{1} q!1

1.3
{-1} p!1
```

```
{-2} q!1
{-3} r!1
  |-------
{1} r!1
```

(**typepred** "NAME") Causes type constraints in the sequent. For example, let delta be defined as a positive real constant: delta: posreal, then

```
{-1} |x-a| < delta
```

(typepred "delta") yields:

```
{-1} delta >= 0
{-2} delta > 0
{-3} |x-a| < delta
```

(**undo**) Undoes the previous command

## 2.2 Function Declarations

An example of a real function specificaion with $x$ and $f(x)$ of type real is

$$f(x\text{: real})\text{: real} = 2x + 1.$$

The beauty of this specification language is that it is inherently easy to specify constants, variables, and functions of different types. The above example could have been much more complex and of different types.

It is with this in mind where the capacity of PVS to be utilized at a more abstract level can aid in the co-design of systems. As can be seen in Figure 1, PVS can be used to specify the functions $F_{1..n}$, and its composites a,b,… The declarations of the sorts domain and range can then be analyzed to determine if a hardware, software, or hardware/software design is most suitable given the parameters and constraints of the problem. It is obvious from this paradigm that in co-designing a system, the functionality relations between software and hardware must be integrated to achieve the optimal design with minimal simulation. The reader is referred to the PVS Training Manual [2] for more examples of functions.

### 2.3 A Little more on TCC's

Type Check Conditions (TCC's) occur when there may be a discrepancy in types during the typecheck. For example, if *x* is a real variable and *f* is an integer function, a TCC may occur. PVS, however, may prove this to be legal and at the end of the typecheck, the bottom buffer will display

1 TCC  1 proved  0 subsumed  0 unproved

which does not imply that your specification has been verified. The prover must still be invoked and the theorems proved as previously discussed. Note that if *f* and *x* are both declared as real types, that no TCC's will occur. Similarly, TCC's may be generated and unproved by the "M-x tcp" command.

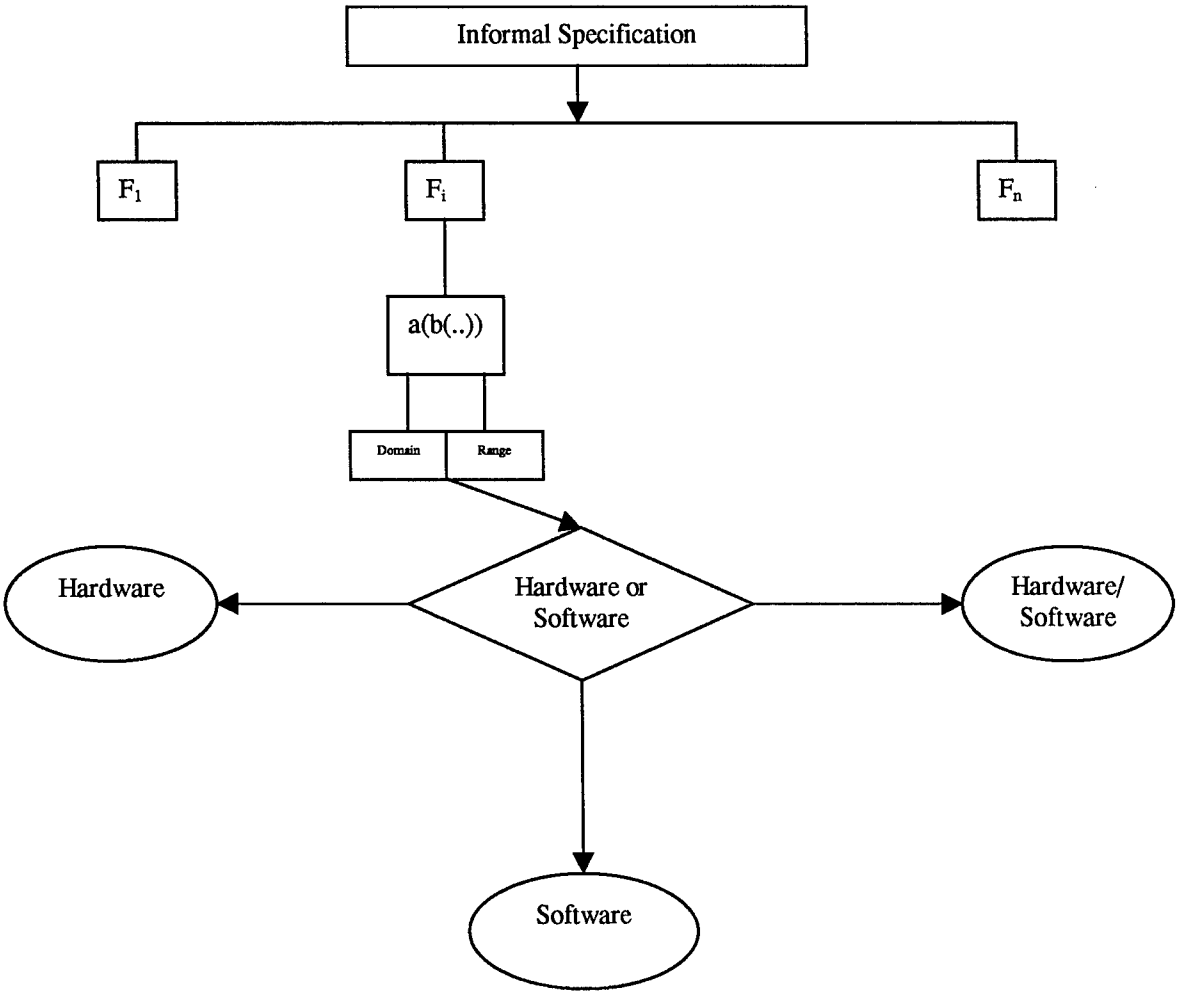



**Figure 1**

## 3. Limit Problem

Informal specification of the problem:

Prove that the limit of $f(x) = 2.1x + 1.2$ is 3.51 as $x$ approaches 1.1.

Definition of a limit: $f(x) = mx + b$ is $L$ as $x$ approaches $a$ means that $\forall\ \varepsilon > 0$, $\exists\ \delta > 0$ such that if $|x - a| < \delta$, then $|f(x) - L| < \varepsilon$. In PVS, it was found that if slope $m$ is defined in decimal form, there is a type problem and the limit could not be proven. By rewriting $m$, $a$, and $L$ as rationals (e.g. 2.1 = 21/10), the existence of the linear limit became provable in PVS.

Formal Specification:

```
limit: THEORY

BEGIN

x: VAR real
m,b: real
eps: posreal

f(x: real): real = 21/10*x + 12/10
delta: posreal = eps/(21/10)
L: real = 351/100
a: real = 11/10

Lim: LEMMA abs(x - a) < delta
                IMPLIES abs(f(x) - L) < eps

END limit
```

Several proofs can be generated which lead to the same conclusion. PVS performs this operation automatically and saves the proof in a *.prf file. Let us now evaluate two of them and determine the more meaningful one. Note that in the proofs below, each command can be executed in its natural sequence when the prover is invoked. In other words, in Generated Proof 1, (skosimp*) would be the first prover command, (expand "abs") would be the second, and so on.

Generated Proof 1:

```
(|limit|
 (|Lim| "" (SKOSIMP*)
  (("" (EXPAND "abs")
    (("" (LIFT-IF)
      (("" (GROUND)
        (("1" (TYPEPRED "delta") (("1" (GRIND) NIL)))
         ("2" (TYPEPRED "delta") (("2" (GRIND) NIL)))
         ("3" (TYPEPRED "delta") (("3" (GRIND) NIL)))
         ("4" (TYPEPRED "delta") (("4" (GRIND) NIL)))))))))))
```

Without the "LIFT-IF", the subgoal will not be generated. The "TYPEPRED" command, on the other hand, is unnecessary. Also, one might ask "What does ground and grind do?" Generated Proof 1 hides the intricacies of this proof. A more lucid proof is given in Generated Proof 2 below.

Generated Proof 2:

```
(|limit|
 (|Lim| "" (SKOSIMP*)
  (("" (EXPAND "abs")
    (("" (LIFT-IF)
      (("" (SPLIT)
        (("1" (SPLIT)
          (("1" (FLATTEN)
            (("1" (REWRITE "L")
              (("1" (REWRITE "a")
                (("1" (REWRITE "delta ")
                  (("1" (REWRITE "f") (("1" (ASSERT) NIL)))))))))))
           ("2" (GRIND) NIL)))
         ("2" (SPLIT)
          (("1" (FLATTEN)
            (("1" (REWRITE "f")
              (("1" (REWRITE "L")
                (("1" (REWRITE "a")
                  (("1" (REWRITE "delta") (("1" (ASSERT) NIL)))))))))))
           ("2" (PROP)
            (("2" (REWRITE "f")
              (("2" (REWRITE "a")
                (("2" (REWRITE "L")
                  (("2" (REWRITE "delta")
                    (("2" (ASSERT) NIL)))))))))))))))))))))
```

A comparison of the two proofs above illustrates some of the hidden functions of grind as previously mentioned in section 2.1.

Errors in the attempt to proof the specification of the existence of a linear limit for the general line $f(x)=mx+b$ were not so easily determined. From the first proof tree, the grind command caused PVS to halt. Utilizing the structure of the second proof allowed the discovery of some of the discrepancies in the specification. The verification of the general case is not yet complete, but at least half of the subgoals have been proven. Thus, it is suggested to the user to grind with discretion.

## 4. More Examples

The following examples illustrate the basic commands given above when applied in a simple example. Only the commands which changed the preconditions or the postconditions are listed. It can be assumed that the other seemingly appropriate commands listed in section 2.1 had no apparent affect on the condition of the proof. DeMorgan's Law states that for all p and for all q, not (p and q) if and only if not p or not q.

Truth Table for NOT(p & q) => (NOT p or NOT q)

| p | q | NOT(p & q) | IMPLIES | NOT p | NOT q | NOT p OR NOT q |
|---|---|---|---|---|---|---|
| T | T | F | T | F | F | F |
| F | T | T | T | T | F | T |
| T | F | T | T | F | T | T |
| F | F | T | T | T | T | T |

**Figure 2**

Figure 2 shows that this is a tautology for any predicates p and q. A similar argument can be used to show (NOT p OR NOT q) => NOT(p & q). What follows is the formal specification as it would appear in the PVS file de_morgan.pvs:

```
de_morgans_law: THEORY
BEGIN

p,q: VAR boolean
```

```
De_Morgans_Law_conj: THEOREM NOT(p & q) IFF (NOT p OR NOT q)
De_Morgans_Law_disj: THEOREM NOT(p OR q) IFF (NOT p & NOT q)

END de_morgans_law
```

Proof of DeMorgan's Law as it appears in the de_morgan.prf file:

```
(|de_morgans_law|
(|De_Morgans_Law_conj| "" (SKOSIMP*)
 (("" (SPLIT) (("1" (PROP) NIL) ("2" (FLATTEN) (("2" (GRIND) NIL)))))))
(|De_Morgans_Law_disj| "" (SKOSIMP*)
 (("" (SPLIT)
   (("1" (FLATTEN) (("1" (SPLIT) (("1" (PROPAX) NIL) ("2" (PROPAX) NIL)))))
    ("2" (FLATTEN)
     (("2" (SPLIT) (("1" (PROPAX) NIL) ("2" (PROPAX) NIL)))))))))))
```

After typechecking, the prover is invoked and the verification is done and displays the following:

```
DeMorgans_Law: {1} FORALL (p: boolean, q: boolean):
                     NOT((p & q)) IFF (NOT p OR NOT q)

        (skolem!) yields:
                |-------
              {1} NOT((p!1 & q!1)) IFF (NOT p!1 OR NOT q!1)

        (skosimp*) yields:
                 |-------
              {1} NOT((p!1 & q!1)) IFF (NOT p!1 OR NOT q!1)

        (grind) QED completes the proof
```

Result from (skolem!) on DeMorgan's Law:

```
{1} NOT((p!1 & q!1)) IFF (NOT p!1 OR NOT q!1)

(split) Splitting the conjunctive yields two subgoals
        DeMorgans_Law_conj.1:
           |-------
         {1} NOT((p!1 & q!1)) IMPLIES (NOT p!1 OR NOT q!1)
    (postpone) yields the second conjunctive from the split:
```

```
        DeMorgans_Law_conj.2:
           |-------
          {1} (NOT p!1 OR NOT q!1) IMPLIES NOT((p!1 & q!1))

    (prop) QED completes the proof
    (ground) QED completes the proof
    (grind) QED completes the proof

Result from the (split) above which yielded DeMorgans_Law_conj.1:
    {1} NOT((p!1 & q!1)) IMPLIES (NOT p!1 OR NOT q!1)

    (flatten) yields:
          {-1} p!1
          {-2} q!1
           |-------
          {1} (p!1 & q!1)
    (prop) completes the proof of DeMorgans_Law_conj.1 and displays
          DeMorgans_Law_conj.2
    (ground) QED completes the proof
    (grind) completes the proof of DeMorgans_Law_conj.1 and displays
          DeMorgans_Law_conj.2
    (skosimp*) yields:
          {-1} p!1
          {-2} q!1
           |-------
          {1} (p!1 & q!1)
```

At this juncture, applying (assert) will complete the proof of DeMorgans_Law_conj.1 and display DeMorgans_Law_conj.2 which can be similarly proof checked.

Several other examples (i.e. Modus Ponens, Modus Tollens, etc.) were easily specified and checked using PVS. There was an attempt at this point to write specifications using ORA Larch/VHDL, however, due to the lack of documentation, this author ceased his efforts and returned to writing specifications in PVS.

## 5. The Traffic Light Problem

The informal specification is as follows:

> Design a traffic light that will be placed at the intersection of two roads with traffic in both directions.

Formal Specification:

```
traffic_lt  % [ parameters ]
            : THEORY
% NOTE: the percent symbol "%" is used to comment lines out
% in a specification file
BEGIN

colour: TYPE = {red,yellow,green}

ns(x: colour): bool
ew(x: colour): bool

init_state_ns: AXIOM ns(green) = TRUE XOR ns(yellow) = TRUE XOR
ns(green) = TRUE
init_state_ew: AXIOM ew(green) = TRUE XOR ew(yellow) = TRUE XOR
ew(green) = TRUE

state1: AXIOM ns(green) = NOT ew(green) & NOT ew(yellow)
state2: AXIOM ew(green) = NOT ns(green) & NOT ns(yellow)

ew_green: LEMMA ew(green) OR ew(yellow) => ns(red)
ns_green: LEMMA ns(green) OR ns(yellow) => ew(red)

END traffic_lt
```

PVS Proof:

```
(|traffic_lt|
(|ew_green| "" (FLATTEN)
 (("" (PROP)
        (("1" (LEMMA "init_state ns")
```

```
    (("1" (LEMMA "init_state_ew")
      (("1" (GRIND) (("1" (LEMMA "state2") (("1" (GRIND) NIL)))))))))
  ("2" (LEMMA "init_state_ew")
   (("2" (LEMMA "init_state_ns")
     (("2" (GRIND) (("2" (LEMMA "state1") (("2" (GRIND) NIL))))))))))))))
```

The initial states for each street utilize the exclusive or, XOR, and allow each side to have exactly one light on at any single instance. The axioms state1 and state2 specify that if one street has a green or yellow light, the other has a red light. The proof of the lemmas ensure that this will hold true for any case.

## 6. Conclusion

PVS is a powerful, interactive tool that is comparatively well documented. However, it was discovered that the tools used in applying Formal Methods are at times ambiguous and cumbersome, including PVS. This is not due to the lack of elegance of the package. The difficulty is found in using the documentation. This report has simplified examples that enable the user to get started writing meaningful specifications. I would like to stress the fact that the importance of formally specifying and verifying large, complex hardware and software systems, is yet unrealized by the majority of educators. This becomes a burden to industry that clearly has a requirement for reliable systems, with increasing levels of sophistication and integration of hardware and software designs. As mentioned in section 3, PVS would appear to be well equipped to handle a formal specification at the highest level of abstraction and enable a more harmonious co-design environment. Future research in Formal Methods must endeavor to integrate co-designed projects into the development process to ensure the production of reliable, high quality, easy to upgrade, and low maintenance systems.

## 7. Acknowledgements

I would like to thank the staff at RDL, AFOSR Program Office, and the Rome Laboratory/ERDD for their consistent and expeditious help with my administrative needs. I would also like to extend my gratitude to Mike Nassif, my focal point, whose guidance allowed my research to be productive and my summer at Rome Labs an enjoyable one.

## References

[1] Crow, Owre, Rushby, Shanker, Srivas, "A Tutorial Introduction to PVS," Presented at WIFT '95: Workshop on Industrial-Strength Formal Specification Techniques, Boca Raton, Florida, April 1995.

[2] Butler, DiVito, Miner, Carreno, Caldwell, "PVS Training Manual," NASA Langley Research Center Workshop, 25th - 28th February, 1997.

[3] Ford, Ford, "Introducing Formal Methods", Ellis Horwood Limited, 1993