COMPUTING/SOFTWARE SKILLS AND EXPERIENCE

Mark Yashar
http://www.linkedin.com/in/markyashar
https://github.com/markyashar

OVERVIEW

- Operating Systems: **Windows**, **Linux** (**Red Hat, Centos, Ubuntu**), **Unix**, **Mac OS**.
- Programming Languages and Data Analysis Packages: **Python** (including **numpy, matplotlib, scipy, pandas, seaborn** and **scikit-learn** libraries), **C/C++** (including object-oriented programming and associated use of **gdb** and **ddd** debuggers and **Eclipse**), **MATLAB**, **Fortran**, **Perl**, **R**, Unix shell scripting, **IDL**, **Mathematica**, **HTML**, **Java**, **Berkeley DB XML**, **MySQL,** Common Astronomy Software Applications (**CASA**), Image Reduction and Analysis Facility (**IRAF**), **Supermongo**, **Meqtrees**, Weather Research and Forcasting Model (**WRF**), NCAR Command Language (**NCL**), NetCDF Command Line Operators (**NCO**).
- Other Software Applications: **LaTex**, **EXCEL** (including use of formulas, functions, and plotting features and capabilities), Concurrent Versions System (**CVS**), **VMware Workstation**, **Liferay Enterprise**, **Git.**
- Supercomputers (e.g., **Cray** XE6) and high-performance computing (**Hadoop**)

OPERATING SYSTEMS

- I've made extensive use of and have extensive experience with **Linux** operating systems, including Red Hat, Centos, and Ubuntu, including installing, building/compiling, configuring (and updating), and running various software packages and associated dependencies on different Linux systems. I have also carried out some associated system administration tasks.
- I have extensive experience using Windows and Mac OS X operating systems, including installation, configuration, and usage of various software packages (e.g., Anti-Virus and Anti-Malware software, X-Windows and SSH software such as SSH Secure Shell, NX Client for Windows, Exceed for connecting with Linux machines from a Windows Desktop, NX Player for connecting with Linux machines from a Mac OS X Desktop, etc.).
- I assisted in supporting and maintaining Windows OS (including Windows NT support, maintenance, and Desktop administration), software, software licenses, and printer systems in my job with the Particle Astrophysics group at SLAC.

UNIX SHELL SCRIPTING

- Utilized and modified UNIX shell scripts to create Measurement Sets for radio astronomy interferometry simulations for the work I did at NCSA/UIUC.

- Wrote and used UNIX shell scripts with 'awk' and 'sed' in various projects to find files with certain creation dates and words or dates in their file names and then re-

directed output to a text file.

- Wrote and used UNIX shell scripts for various projects to search through text in files in a subdirectory and then print out those file names which contain that particular text or print the output of a particular piece of text (e.g., error messages) for particular files.

- Wrote and modified UNIX shell scripts that involved submitting (batch) jobs to run MCMC code and other code on a Linux Cluster and supercomputers for work at UCD and UCB (see resume and Linkedin profile for more details).

PYTHON EXPERIENCE AND SKILLS

- I wrote python scripts to generate plots of floating point operations per second (computational costs) vs. antenna diameter and other antenna parameters for work I did at NCSA/UIUC.

- I wrote Python scripts to carry out Monte Carlo simulations for calculating means and standard deviations of image source flux densities related to different analytical expressions for dynamic range that were being investigated and tested in my work at NCSA/UIUC.

- I calculated image statistics (means, standard deviations, root mean square, etc. of pixel brightness values) for FITS file images in my work at NCSA/UIUC.

- I utilized, modified, and wrote python scripts involving functions for various projects.

- I've written python scripts for computing radio antenna array configuration layouts (Earth-Centered Coordinate Systems, Longitude and Latitude positions, (x,y) locations, etc.) for a random distributions of antennas within a circular area, including Gaussian distribution, uniform circular random distribution, log-spiral distribution, etc., for my work at NCSA/UIUC.

- I've written python scripts for generating plots and histograms from data read from files for various projects.

- I wrote python scripts to generate time-series plots and mean time series plots (e.g., CO_2, pressure, or temperature vs. time) from data in a file for my work at UCB.

- Additional extensive independent study work and training in Python involving file I/O, IPython, Jupyter notebooks, numpy, scipy, pandas, matplotlib, seaborn, Python SQL drivers , scikit-learn.

C/C++ EXPERIENCE AND SKILLS

- I installed, built/compiled, configured, and utilized a development version of a software package (mostly written in C++) called the Common Astronomy Software Applications (CASA) package to carry out the SKA observation and imaging simulation work (at NCSA/UIUC) that I describe in my resume/CV and in my Linkedin profile. This involved installing and building/configuring all of the dependencies for CASA, as well as configuring makefiles and makedefs as necessary for setting up a local workspace. I also used gdb, ddd, and Eclipse C++ debuggers (i.e., setting breakpoints, stepping through the code line by line, using the back-trace feature to look at the call stack of the dumped core to found out where among the routines segmentation faults were occurring, etc.) to carry out a static and dynamic analysis of the CASA C++ code (along with some Python and Fortran code as well). I also modified the CASA C++ code to test and experiment with the imaging simulations and the resulting dynamic ranges and fidelities of the resulting simulated images. The goal of the work was to be able to modify and test the CASA C++ code to allow the SKA beam model to be imported into the beam library, corrupt the simulated visibilities with that beam model, the remove or reverse the effects of the beam model and pointing errors using a particular algorithm, and then compare the images with images that were not corrupted with the primary beam model. I gained some knowledge and developed some skills with C++ object-oriented programming during the course of this project.

- I studied coursework and attended some lectures on C++ and Object-Oriented programming in a graduate computational physics course I took at UC Davis.

- I sat in on a 1-week C++ course (including Object-Oriented Programming) taught by Dr. Paul Kunz at SLAC shortly before working there.

- I assisted with the use of Concurrent Versions Systems (CVS) and compiling and building Paul Kunz's Hippodraw C++ code (including the associated C++ library) when I was working at SLAC.

- I've engaged in self-directed study/learning and tutorials (online and via books) of C++ to try to learn more about and get up to speed with C++ (including study of object-oriented programming).

- I wrote and modified programs in C for a graduate computational physics course at UC Davis to time for loops and some basic math operations such as addition, subtraction, multiplication, division, taking square roots, etc.

- I utilized Numerical Recipes in C code (e.g., in connection with Kolmogorov-Smirnov Tests) and made some small modifications to the code as necessary for work on the MACHO project at LLNL that I described in my resume and Linkedin profile.

- I have corresponding experience with various C/C++ GNU tools such as GNU make, GCC, GDB, GNU build system (autotools), etc.

MATLAB EXPERIENCE AND SKILLS

- I have general experience with MATLAB in opening a text file for writing and reading, plotting (with labels) data from text file, and generating different kinds of plots (e.g., line plots, contour plots, log-log plots, 3-D plots, etc).

- Utilized and made modifications to MATLAB code for the MCMC analysis of a dark energy model as part of my thesis work at UCD (supervisor: Professor Andreas Albrecht). This involved extensive use of MATLAB in working with vectors and arrays, solving for posterior probability distributions (i.e., the probability of the parameter given the data), calculations of means, averages, standard deviations, root mean squares, χ^2 values, calculations of covariance matrices and other linear algebra operations, and generating error contour plots and line plots of various cosmological parameters vs. time and scale factor.

- For another dark energy project at UCD (supervisor: Professor Andreas Albrecht), I assisted a graduate student in writing MATLAB code to grid out in parameter space what χ^2 was for each point (i.e., a χ^2 grid search). This code used four nested for loops, with one for loop for each parameter varied. I generated 3-dimensional χ^2 plots as a function of the varied parameters to help visualize the change in χ^2 (as it moved towards its minimum value) with different parameter values.

- For a data analysis in astrophysics graduate course at UCD taught by Professor Chris Fassnacht, I wrote MATLAB code to calculate correlation coefficients, for fitting data to a straight line, and for general least-squares fitting.

- I sat in on a graduate cosmology course at UCD (taught by Professor Andreas Albrecht) in which I wrote and utilized MATLAB code that used data structures to pass parameters to functions (i.e., object-oriented MATLAB), defined and set up vectors, generated plots of dark energy density and other cosmological parameters as a function of scale factor, read data into an array, called functions, opened a text file for writing and reading, and plotting that data from that text file. I also used and wrote MATLAB code that utilized an ordinary differential equation (ODE) solver for solving differential equations and which carried out integration operations.

- For my work at UCB, I utilized and modified MATLAB code to input CO_2 emission data into the WRF-Chem model.

- I've used MATLAB/Octave to carry out exercises for an online Coursera Machine Learning class (Andrew Ng).

FORTRAN EXPERIENCE AND SKILLS

- I utilized and configured the synthetic Color Magnitude Diagram (CMD) algorithm "StarFISH" (written in Fortran 77) in my graduate research work at LLNL to generate model CMDs.

- Wrote Fortran 77 programs to time for loops as part of a graduate computational physics course taught by Professor John Rundle (e-mail: rundle@physics.ucdavis.edu) at U.C. Davis.

- Wrote a Fortran 77 program to time some simple basic math operations such as addition, subtraction, multiplication, division, and taking square roots as part of a graduate computational physics course taught by Professor John Rundle at U.C. Davis.

- Studied and utilized random site percolation code written in Fortran 77 (involving the use of Monte Carlo Sweeps) as part of a graduate computational physics course taught by Professor John Rundle at U.C. Davis.

- For the final project of a graduate computational physics course taught by Professor John Rundle at U.C. Davis I wrote a Fortran 77 program that computed a closed orbital ellipse of an extrasolar planet orbiting a single star using data input by the user. The program queried the user to enter various orbital and physical parameters of the planet-star system and used the data to calculate the observed effective equilibrium blackbody temperature of the extrasolar planet for a given orbital phase. The program also calculated the planet-to-star flux ratios at given orbital phases. I have attached the Fortran 77 code ("orbits.f") that I wrote to this e-mail.

- I utilized Fortran on the data processing and data extraction work (e.g., Fast Fourier Transforms) for the TEXES project at U.C. Davis described in my CV. The Fortran code read in in raw TEXES data, processed the data (e.g., flat-field calibration, correction of distortions, correlation and shifting, 2D FFT, store processed array, extraction and storing of spectrum, flag bad pixels, open raw and reduced header files, etc.) and wrote the processed data back to disk. (Supervisor: Dr. Matt Richter, U.C. Davis; e-mail: richter@physics.ucdavis.edu).

- I have general experience with for loops and file I/O in Fortran 77.

- I compiled, built, configured and utilized/ran Fortran 90/95 Weather Research and Forecasting model (WRF) code for carrying out mesoscale and regional atmospheric transport modeling of anthropogenic and biogenic carbon dioxide emissions for my work at UC Berkeley. I also made some necessary modifications to related pre-processing and other Fortran code (e.g., prep_chem_sources and WRF-Chem).

PERL EXPERIENCE AND SKILLS

- Wrote, modified, and used Perl scripts to automate archive maintenance and data

processing tasks for work as a Data Aide at SLAC.

- Modified and utilized Perl scripts for use as a 'wrapper' to run Monte Carlo simulations for Kolmogorov-Smirnov tests for the work I carried out as LLNL.

R EXPERIENCE AND SKILLS

- Installed, configured and utilized R software packages (including RStudio) and associated R libraries on Linux and Mac OS X operating systems.
- Modified and used R scripts (for my work at UCB) to input CO_2 emission fields data to create NETCDF-4 files containing arrays for CO_2 anthropogenic emissions for use in the WRF-Chem model runs, i.e., modified and utilized R scripts to edit NETCDF-4 so that they were in correct format and containing the necessary CO_2 emission fields and data to be input into WRF-Chem simulation runs.
- Installed, configured, modified, and utilized WRF Vegetation Photosynthesis Model (WRF-VPRM) biospheric model, mostly written in R, to simulate CO_2 biosphere fluxes and atmospheric CO_2 concentrations.

IDL EXPERIENCE AND SKILLS

- I have general experience with IDL in opening a text file for writing and reading, plotting data from a text file, and generating different kinds of plots (e.g., line plots, contour plots, log-log plots, 3D plots, etc.)

- Used IDL in an undergraduate astronomy lab course at SFSU taught by Prof. Geoff Marcy at SFSU to carry out orbital/celestial mechanics calculations (e.g., the two-body problem).

- I used IDL to display spectra data (flux vs. wave number) in the TEXES project at UCD described in my resume/CV.

- For the final project in the graduate computational physics course referred to previously (taught by Prof. John Rundle, UCD), I used IDL to generate plots of extra-solar planet orbit and property calculations (i.e., y-coordinate vs. x-coordinate in astronomical units, orbital speed vs. orbital phase, planet temperature vs. orbital phase, planet-star flux ratios vs. orbital phase).

- I also used IDL to generate random site percolation output plots for work on the random site percolation project in a computational physics graduate course (described above) at UCD, as described in my resume/CV.

MATHEMATICA EXPERIENCE AND SKILLS

- Used Mathematica to, for example:

- find the similarity transformation that diagonalizes a matrix [A],
- find eigenvalues and eigenvectors,
- solve with Gaussian elimination a linear system [A][x]=[b],
- give the LU decomposition of a matrix,
- compute the determinant of [A].

HTML EXPERIENCE AND SKILLS

- I've written and used HTML code to create web pages for courses and to post and create links to documents (as a student at SFSU and UCD).

MYSQL EXPERIENCE AND SKILLS

I have knowledge and experience using MySQL to:

- Installing, setting up, and configuring MySQL server or Mac OS X
- Creating at table
- Retrieving information from a table or multiple tables
- Selecting data from particular rows and columns
- Carrying out date calculations
- Pattern matching
- Aggregate Functions
- Join operations

SUPERCOMPUTING AND HIGH-PERFORMACE COMPUTING EXPERIENCE AND SKILLS

- I installed, built, compiled, modified, maintained, and utilized/ran WRF Fortran 90 and other code on a NERSC Cray XT CLE/Linux x86_64 system (called "Hopper"), PGI compiler with gcc in distributed memory parallel (Message Passing Interface – MPI) mode for multiple processors. I ran the WRF code in parallel on the compute nodes via MPI tasks.

- When working at NCSA/UIUC I studied the possibility and feasibility (but never actually carried out the project) of working on a radio imaging statistics analysis project involving Monte Carlo simulations at the National Center for Supercomputing Applications (NCSA) at UIUC in which Portable Batch System (PBS) batch job scripts would be submitted to a Dell Intel 64 Tesla Cluster consisting of 192 compute nodes, 1536 CPUs and 96 accelerator units with other technical specifications described at http://www.ncsa.illinois.edu/UserInfo/Resources/Hardware/Intel64TeslaCluster/TechSummary/ as well as to a Dell NVIDIA Cluster system consisting 36 Dell PowerEdge C6145 quad-socket nodes with dual 8-core AMD Magny-Cours 6136 processors and 64 GB of memory (with each node supporting 8 NVIDIA Fermi M2070 GPUs -- (http://www.ncsa.illinois.edu/UserInfo/Resources/Hardware/DellNVIDIACluster/

TechSummary/).
This project would have involved:

> * Converting a Load Sharing Facility (LSF) job script to a PBS jobs script.
> * Job scheduling
> * Eventually implementing and utilizing C++ MPI code

The project would have included generating a radio image separately with each computing node, corrupting each image with particular gain error models, distributing Monte Carlo or bootstrap samples across unique hosts, computing statistics for each image and final statistics for all images, with the eventual goal of calibrating the relation between how gain errors translate into root mean errors and to verify analytical dynamic range expressions derived in the memo listed in my CV. (Supervisor: Dr. Athol Kemball, UIUC; e-mail: akemball@illinois.edu)

- I wrote, modified, and utilized UNIX shell scripts to submit and run Matlab Markov Chain Monte Carlo (MCMC) code and to expedite the running of MCMC simulations (and the generation of MCMC output) on a Linux computing cluster for my PhD thesis work (dark energy research) as a graduate student at UC Davis. (Supervisor: Professor Andreas Albrecht; e-mail: ajalbrecht@ucdavis.edu).

- I submitted Load Sharing Facility (LSF) batch jobs to the SLAC Linux Batch farms (34 dual CPU 450 MHz machines) to convert raw data (*.dat) files to Flexible Image Transport System (FITS) format files For my work at SLAC. I worked on writing and executing Perl and UNIX shell scripts that created batch jobs to be submitted to Linux clusters at SLAC (.e.g., bsub -o $datfile.out -J $datfile ...). I examined and analyzed the corresponding output files, error files, log files, exit codes, error codes, etc., and looked for and reported error messages to try to ascertain whether the FITS files were successfully created. Unix 'cron' jobs and perl scripts were set up, written, and utilized to schedule when the batch jobs were submitted and to automate the process.

- Utilized and modified Hadoop and Python scripts to convert JSON-formatted data (by normalized it) to tab-delimited data.

VIRTUAL COMPUTING EXPERIENCE AND SKILLS

- In my job a NCSA/UIUC, I installed, configured, and maintained multiple 64-bit VMware (Workstation 7.1) virtual machines (VM) on a host Linux CentOS 5.x system using ISO disk images. Various guest operating systems, including CentOS 5.x and Ubuntu 9.x etc. were installed on the newly created VMs. This work also involved using the 'GParted' software package to expand the size of the VM root partition.

MISCELLANEOUS COMPUTING EXPERIENCE AND SKILLS

- In my work at NCSA/UIUC I assisted in installing, configuring, setting up, maintaining, and utilizing Java-based Liferay Portal (https://www.liferay.com/products/liferay-portal/overview) software with Apache Tomcat Server Bundle and MySQL Database on a Linux machine. This work also involved installing and configuring Java Standard Edition (SE) and Java Development Kit (JDK).