The MathWorks MATLAB&SIMULINK

Using MATLAB for Life Sciences, Stanford University 10/8/04

The MathWorks MATLAB&SIMULINK

Today's Agenda

12:15 PM	**Welcome and Introductions**
12:30 PM	**The MATLAB Development & Deployment Platform**
1:30 PM	**Break**
1:45 PM	**MathWorks Tools for Bioinformatics**
2:30 PM	**MathWorks Tools for Biomedical Signal Analysis and Simulation**
3:15 PM	**Questions/Answers and Wrap up**
3:45 PM	**Seminar closes**

The MathWorks MATLAB&SIMULINK

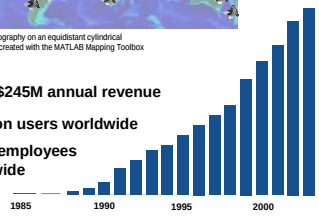
2004 - The MathWorks 20th year

Earth's topography on an equidistant cylindrical projection, created with the MATLAB Mapping Toolbox

$245M annual revenue

1 million users worldwide

1000+ employees worldwide

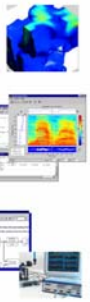


- Headquarters: **Natick, Massachusetts, USA**
- European offices: **UK, France, Germany, Spain, Benelux, Italy, Switzerland, Sweden**
- North American offices: **CA, MI, Washington DC, TX**
- Asia/Pacific office: **Korea**
- Worldwide consulting
- Distributors in 21 countries

The MathWorks MATLAB&SIMULINK

MathWorks Mission and Vision

Accelerate innovation and discovery in engineering and science



MATLAB

- a powerful, high-level language to develop algorithms, collect and analyze data, and visualize information

Simulink

- a graphical system to model and simulate complex systems, and implement real-time and embedded systems

The MathWorks MATLAB&SIMULINK

MathWorks Products are Used in Various Industries

- Aerospace and Defense
- Automotive
- **Biotech, Pharmaceutical and Medical**
- Communications, Semiconductor
- Education
- Financial Services
- Industrial Equipment and Machinery
- Instrumentation
- **Medical Devices and Instrumentation**

The MathWorks MATLAB&SIMULINK

Thousands of universities teach students using MathWorks products.

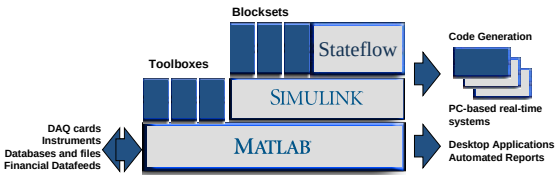
More than 700 textbooks for education and professional use, in 19 languages

- Biosciences
- Controls
- Signal Processing
- Image Processing
- Mechanical Engineering
- Mathematics
- Natural Sciences
- Environmental Sciences

The MathWorks Product Family

Integrated for:

- technical computing, data analysis and visualization
- system modeling and simulation
- implementation of real-time embedded software

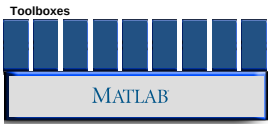


The MathWorks Product Family

Toolboxes for Modeling, Analysis, and Computation

Specific functionality for data analysis, modeling, design, and other capabilities

•Bioinformatics	•Neural Networks	•Symbolic Math	•Wavelet
•Statistics	•Optimization	•Curve Fitting	•Signal Processing
•Image Processing	•Compiler	•Filter Design	•Fuzzy Logic



Today's Agenda

12:15 PM	Welcome and Introductions
12:30 PM	**The MATLAB Development & Deployment Platform**
1:30 PM	Break
1:45 PM	MathWorks Tools for Bioinformatics
2:30 PM	MathWorks Tools for Biomedical Signal Analysis and Simulation
3:15 PM	Questions/Answers and Wrap up
3:45 PM	Seminar closes

Key Features of MATLAB

- Interactive environment to explore data & algorithms
  - Optimized for technical computing
  - Simplified common tasks (import, plot, etc)
  - Intuitive Desktop interface
- Platform to develop tools & user friendly applications
  - Advanced visualization & data analysis capabilities
  - Access Java & other programming languages (API)
- Highly extensible with toolboxes…

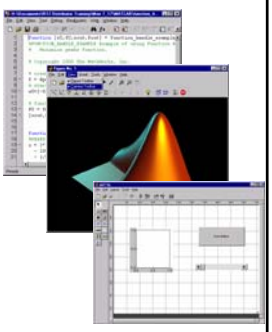
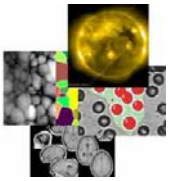


Image Processing Toolbox ~~3.2~~ 4.0

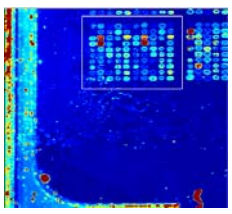
- **Advanced image processing and analysis tools for MATLAB**
- **Numerous image and scientific file formats supported, including HDF-EOS and DICOM**
- **Many person years of development effort**
- **Over 160 state-of-the art functions**



- **Deblurring & enhancement**
- **Image Registration**
- **Transforms (spatial, frequency)**
- **Morphological analysis**
- **Segmentation**
- **Region properties**

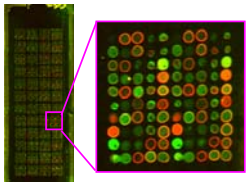
High-throughput experimental techniques require automated image analysis

- Automate image and statistical analysis
- Try out different algorithms
- Build software applications
- Gather quality control measures
- Normalize



Analyzing DNA with Microarray Imaging

Fluorescently tagged mRNA from different cells are hybridized to a microscopic array of hundreds of thousands of cDNA spots that correspond to different genes. Illuminated spots emit different color light, indicating which genes are expressed (e.g., green=control, red=sample, yellow=both).



Through image analysis, the fluorescence at the site of each immobilized cDNA can be quantified. For example, the log ratio of red-to-green intensity gives a measure of gene expression.

Application Challenges

- Clean up images with noise
- Correct for rotation, skew => regular spot spacing (rows, cols)
- Isolate sub-image array of colored spots
- Separate red and green planes
- Remove non-uniform local background
- Identify regular grid pattern of spots on slide
- Address individual spots by region of interest
- Integrate red and green intensity values
- Detect poor spot quality and flag as bad data points
- Determine gene expression from intensities
- Develop robust algorithm to automate process
- Deploy application to implement algorithm.

What did this case study show?

1. **MATLAB** environment was great for developing an algorithm (environment + language + graphics)
2. **Image Processing Toolbox** provided a rich set of functions for segmentation, region properties and background removal
3. **Signal Processing Toolbox** provided autocorrelation function to determine spot periodicity.

What else could you do?

1. Make your program user user friendly for coworkers not familiar with MATLAB (`guide`)
2. Turn your M-code into standalone C program for distribution (**MATLAB Compiler**)
3. Use multivariate analysis of spot clusters to determine which genes matter (**Statistics Toolbox**)
4. Learn to differentiate good and bad spots based on shape features (**Neural Network Toolbox**)
5. Link results from microarray analysis with other data systems and analyses (**Database Toolbox**),

Today's Agenda

12:15 PM	**Welcome and Introductions**
12:30 PM	**The MATLAB Development & Deployment Platform**
1:30 PM	**Break**
1:45 PM	**MathWorks Tools for Bioinformatics**
2:30 PM	**MathWorks Tools for Biomedical Signal Analysis and Simulation**
3:15 PM	**Questions/Answers and Wrap up**
3:45 PM	**Seminar closes**

MATLAB and the Bioinformatics Industry

- Biotech companies were using MathWorks tools
  - Image Processing, Statistics, etc.
- MATLAB used in Undergraduate and Graduate Bioinformatics courses
- Additional functionality specific to biotech needed
  - File I/O, visualization, etc.

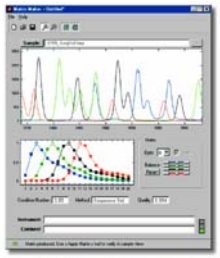
Presentation Layout

- How does MATLAB fit in to Bioinformatics?
- Bioinformatics Toolbox:
  - Sequence Alignment
  - Microarray visualization
  - Phylogenetic analysis.
- Integrating and Deploying Bioinformatics Tools with MATLAB
- Summary & Questions

How does MATLAB fit in to Bioinformatics?

- Sequence Analysis
  - Algorithm design, pattern matching
- Gene Expression Analysis
  - Classification, normalization, clustering
- Computational proteomics
  - Mass Spectrometry data analysis, Monte Carlo simulation
- Systems Biology
  - In silico simulation, pathway inference

User example: Genetic Sequence Base Calling



A portion of the DNA dye-label spectral profile, which allows the researcher to read the sequence of bases in a selected strand of DNA.

Complete draft of the human genome, accelerated by Applied Biosystems — using MATLAB algorithms.

"Having one integrated package is a big advantage. Using MATLAB and the MATLAB Compiler reduced my development time by a factor of 4 or 5."

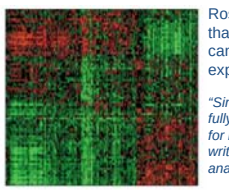
"MATLAB has always been ideal as an algorithm prototyping tool," Labrenz concludes, "but the MATLAB Compiler and C/C++ Math and Graphics Libraries add a whole new dimension, allowing rapid delivery of sophisticated solutions."

Jim Labrenz, Applied Biosystems

How does MATLAB fit in to Bioinformatics?

- Sequence Analysis
  - Algorithm design, pattern matching
- Gene Expression Analysis
  - Classification, normalization, clustering
- Computational proteomics
  - Mass Spectrometry data analysis, Monte Carlo simulation
- Systems Biology
  - In silico simulation, pathway inference

User example: Breast Cancer Prognosis



Two-dimensional clustering of DNA microarray gene expression data from breast cancer samples. Red indicates over-expression; green indicates under-expression. Patients with the expression pattern similar to the top of the plot generally have poor outcome.

Rosetta Inpharmatics recently developed a tool that enables clinicians to determine a breast cancer patient's prognosis based on the gene expression profile of the primary tumor.

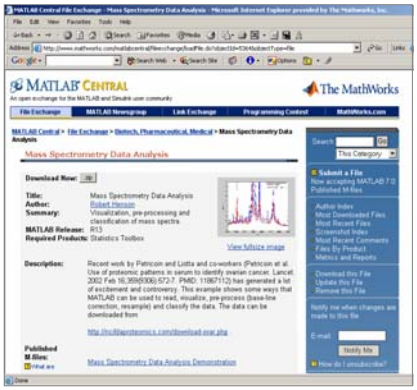
"Since MATLAB and the Image Processing Toolbox are fully integrated and the MATLAB platform is very good for matrix calculation, we did not have to spend time writing the low level image processing and the basic data analysis routines like vector and matrix calculations"

"Our research scientists are happy with the quick feedback," Dr. Dai says. "Using MathWorks tools, we can respond to their requests very fast, and it's easy for the scientists to use these tools. Using the GUIs that we develop in MATLAB, they can access functions without having to remember the underlying code."

Dr. Hongyue Dai,
Rosetta Inpharmatics/Merck & Company

How does MATLAB fit in to Bioinformatics?

- Sequence Analysis
  - Algorithm design, pattern matching
- Gene Expression Analysis
  - Classification, normalization, clustering
- Computational proteomics
  - Mass Spectrometry data analysis, Monte Carlo simulation
- Systems Biology
  - In silico simulation, pathway inference

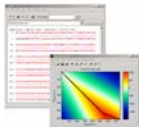


How does MATLAB fit in to Bioinformatics?

- Sequence Analysis
  - Algorithm design, pattern matching
- Gene Expression Analysis
  - Classification, normalization, clustering
- Computational proteomics
  - Mass Spectrometry data analysis, Monte Carlo simulation
- Systems Biology
  - In silico simulation, pathway inference

Data Analysis Toolboxes

Bioinformatics



Signal Processing

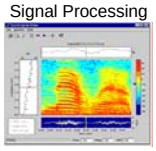
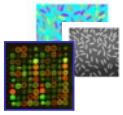
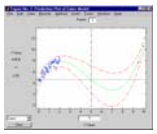


Image Processing

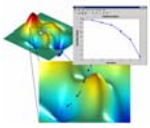


Statistics



Neural Networks

Optimization



Presentation Layout

- How does MATLAB fit in to Bioinformatics?
- Bioinformatics Toolbox:
  - Sequence Alignment
  - Microarray visualization
  - Phylogenetic analysis.
- Integrating and Deploying Bioinformatics Tools with MATLAB
- Summary & Questions

Bioinformatics Toolbox 1.1
Function Overview

- **File I/O**
  - Read FASTA, PDB, GenePix, Affymetrix and many more format files
- **Web connectivity**
  - Directly access GenBank, PDB, EMBL, PIR, BLAST,…
- **Sequence analysis**
  - Base density, codon counts, ORF finding,…
- **Sequence alignment**
  - Local, global and profile HMM based alignment
- **Phylogenetic Analysis**
  - Tree building, interactive tree viewer
- **Microarray normalization & visualization**
  - Normalization tools, Gene filters, expression profile cluster analysis, two-way clustering,…
- **Protein visualization**
  - Hydrophobicity plots, Ramachandran plots,…
- **Tutorials**
  - Examples of using toolbox functions and of connecting to standard tools such as BioPerl, BioJava, Web Services etc.

Sequence Alignment Tutorial Example

- Get human and mouse genes from GenBank
- Look for open reading frames (ORFs)
- Convert DNA sequences to amino acid sequences
- Create a dotplot of the two sequences
- Perform global alignment
- Perform local alignment

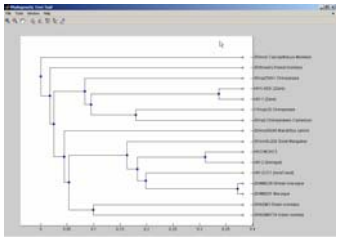
Microarray Data Analysis Tutorial Example

- Plot expression profiles for genes
- Filter genes based on information content of profile
- Perform hierarchical clustering
- Perform K-means clustering
- Perform Principal Component Analysis

Reference:
DeRisi, JL, Iyer, VR, Brown, PO. "Exploring the metabolic and genetic control of gene expression on a genomic scale." Science. 1997 Oct 24;278(5338):680-6.

Phylogenetic Tree Analysis

- Interactive tree viewer
  - Reorder, prune, and rename branches
  - Explore distances
  - Read or write Newick-formatted files
- Similarity metrics
  - Jukes-Cantor
  - p-distance
  - alignment-score
  - user-defined

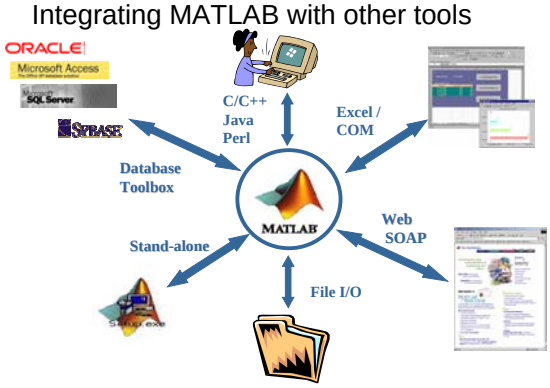


```
>>tr = phytreeread('pf00002.tree')
>>view(tr)
```

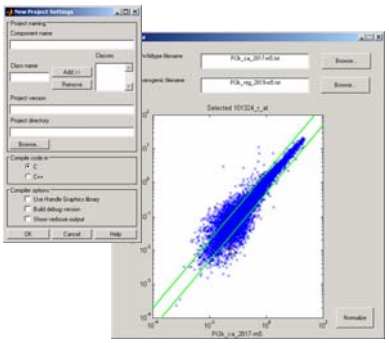

Presentation Layout

- How does MATLAB fit in to Bioinformatics?
- Bioinformatics Toolbox:
  - Sequence Alignment
  - Microarray visualization
  - Phylogenetic analysis.
- Integrating and Deploying Bioinformatics Tools with MATLAB
- Summary & Questions

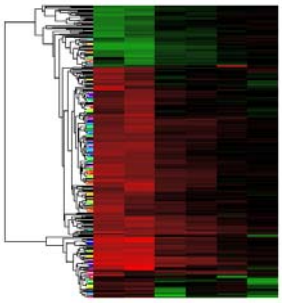
Integrating MATLAB with other tools



Creating a Standalone Application



Visualization for Gene Expression Analysis: Excel Link & Excel Builder

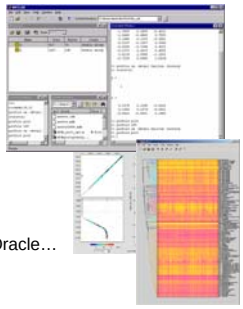


Data from DeRisi JL, Iyer VR, Brown PO. "Exploring the metabolic and genetic control of gene expression on a genomic scale". Science. 1997 Oct 24;278(5338):680-6. PMID: 9381177

Presentation Layout

- How does MATLAB fit in to Bioinformatics?
- Bioinformatics Toolbox:
  - Sequence Alignment
  - Microarray visualization
  - Phylogenetic analysis.
- Integrating and Deploying Bioinformatics Tools with MATLAB
- Summary & Questions

Summary: Why use MATLAB for Bioinformatics?



- Powerful Language
  - Fast & Flexible
- Analyzing Data
  - 1000's Math & Graphics functions
- Quick Prototyping
  - Less programming
- Work with existing data / programs
  - C/C++, Excel, COM, Java, Perl, Oracle…

Questions?

Today's Agenda

12:15 PM	**Welcome and Introductions**
12:30 PM	**The MATLAB Development & Deployment Platform**
1:30 PM	**Break**
1:45 PM	**MathWorks Tools for Bioinformatics**
2:30 PM	**MathWorks Tools for Biomedical Signal Analysis and Simulation**
3:15 PM	**Questions/Answers and Wrap up**
3:45 PM	**Seminar closes**

MATLAB Connects to your hardware devices

Data Acquisition Toolbox
Plug-in data acquisition boards

Instrument Control Toolbox
Electronic and scientific instrumentation

New

Image Acquisition Toolbox
Cameras and Video Equipment

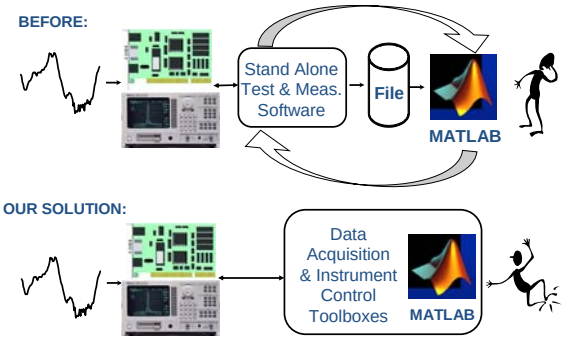
MATLAB
Low-level interfaces to other hardware

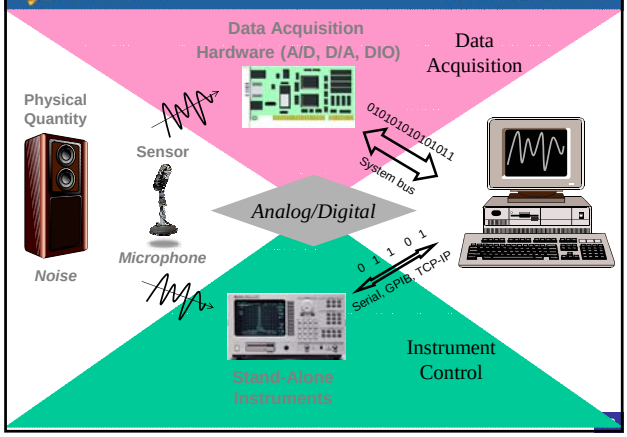
Image Acquisition Toolbox: Supported Hardware

- Matrox:
  - Meteor-II
  - Meteor-II/Digital
  - Corona-II
  - Meteor-II/Multi-Channel
  - Orion
  - Genesis-LC
- Data Translation:
  - 3120
  - 3152-LS
  - 3155
  - 3130 Series
  - 3153
  - 3157
  - 3152
  - 3154
- Windows Video Devices
  - USB/FireWire cameras, DV camcorders, image capture boards, TV tuner cards, etc.
  - Requires:
    - Windows Driver Model (WDM) or Video For Windows (VFW) driver
    - Microsoft DirectX 9.0 or later

Visit http://www.mathworks.com/products/imaq/ for more information

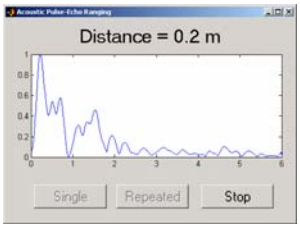
Working with live data in MATLAB



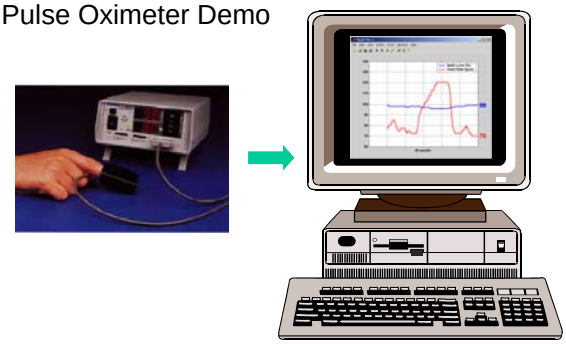


Data Acquisition Demonstration

- Principle of ultrasonics – time resolved distance
- Transmit chirp waveform from speaker
- Receive microphone waveform simultaneously
- Cross correlation
  - determine propagation delay
- Divide by wave speed to get distance.



Analyzing Measured Signals – Pulse Oximeter Demo

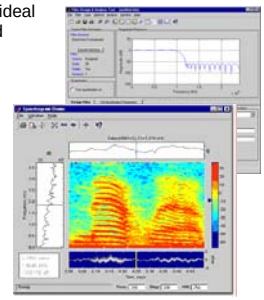


Analyzing Measured Signals

The Signal Processing Toolbox is the ideal environment for signal analysis and DSP algorithm development.

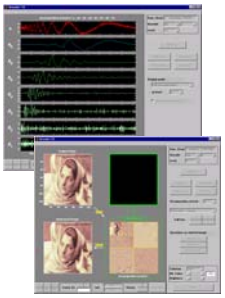
- Waveform generation
- Statistical signal processing
- Spectral analysis
- Parametric modeling
- Transforms
- Windows
- Linear system transforms
- Basic filter design and analysis

```
>> specgramdemo
>> sptool
```



Wavelet Toolbox

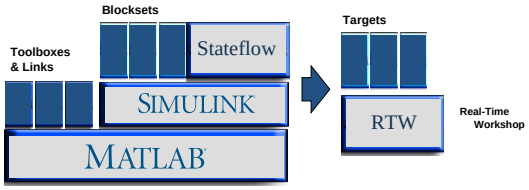
- Continuous wavelet transform
- Discrete wavelet transform
- Wide selection of wavelet basis functions
- Wavelet packet transform
- De-noising and compression



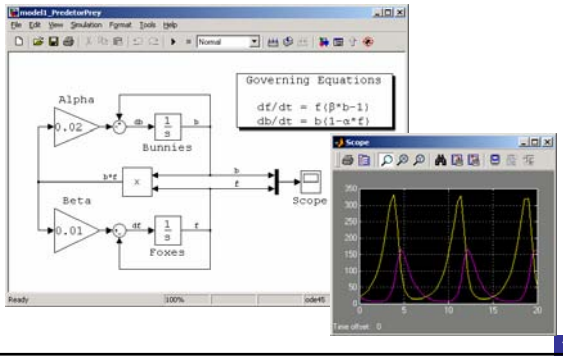
```
>> wavemenu
```

Simulink - The leading environment for modeling and simulating dynamic systems

- System Level Modeling - for simulating biophysical systems
- Automatic Code Generation - for rapid prototyping, hardware-in-the-loop, and production embedded software
- Real-Time Systems-PC-based turnkey systems for rapid prototyping, HIL, and deployment embedded DSP targets for medical devices



Simulink Example – Predator/Prey Model



Deploy your software on an embedded system.

- Automatically generate production code from the specification defined by the model
  - Accurate representation of the model
  - Traceable to model
    - Symbol names, function boundaries
    - Comments/tags link code segments to model
- Code is easy to test and validate
  - Use existing methods
- Integrate with existing or legacy code and runtime environments

Today's Agenda

12:15 PM	**Welcome and Introductions**
12:30 PM	**The MATLAB Development & Deployment Platform**
1:30 PM	**Break**
1:45 PM	**MathWorks Tools for Bioinformatics**
2:30 PM	**MathWorks Tools for Biomedical Signal Analysis and Simulation**
3:15 PM	**Questions/Answers and Wrap up**
3:45 PM	**Seminar closes**

Technical Support

- Technical Support
  - 90% of problems solved in 24 hours
  - 60+ Application Engineers on staff, ½ with Masters Degrees
- World Wide Web (www.mathworks.com)
  - 24x7 self-service technical support
  - over 9,000 technical solutions
  - software archive (ftp.mathworks.com)
  - MATLAB Digest – electronic newsletter
- Newsgroup (comp.soft-sys.matlab)

Training

Invest in your Success

- Expert Trainers Provide
  - Hands-on experience at solving real-world problems
  - Individualized attention
  - Over 30 courses offered in Public, Onsite, and Web-based settings
  - Customized courses to suit your needs
- Application-specific courses
  - MATLAB for Signal Processing
  - MATLAB for Image Processing
  - Statistical Methods in MATLAB
  - Test and Measurement using MATLAB
  - ...

www.mathworks.com/training

The MathWorks Consulting

- Our Goal
  - To partner with the clients and help them succeed in
    - modeling, designing and implementing sophisticated MATLAB-based applications
    - expediting and planning the large scale adoption of The MathWorks toolset within your organization
- Our approach
  - Joint team effort
  - Rapid deployment
  - Several Milestones, less than 3 months apart, with deliverables
- Experts in the following areas
  - MATLAB, SIMULINK, Stateflow and related tools
  - Software Engineering
    - Java, SQL, C/C++, VB, GUI and Database
  - Control engineering

MATLAB Connections

Over 300 add-on products and services from partners that complement and extend MathWorks products

- Specialized third-party toolboxes for MATLAB
- Interfaces to partners' software and hardware products
- Specialized training courses and consulting services
- Turnkey systems providers that incorporate MathWorks products

- To join please see, http://www.mathworks.com/products/connections/join.html

Further Information

- Product Information and Demos

  Trials and technical literature are available through the MathWorks. www.mathworks.com

- MATLAB Central
  - File exchange and newsgroup access for MATLAB and Simulink users
  - www.mathworks.com/matlabcentral
  - Access to comp.soft-sys.matlab

Contact Information

Jamie Winter , Senior Account Manager for
Stanford University and University of California at Berkeley
508-647-7463
508-647-4598, FAX
winterj@mathworks.com

- Licensing / Pricing
- Training

Pat Box and Jane Tansuwan, IT Systems & Services
650-723-4428
software@stanford.edu

- Training Interests