Intro to Computer Science, Biology and Biomedical Informatics (CoSBBI)
CoSBBI Summer Academy - 5607 Baum Blvd and Shadyside Hospital
Monday, June 17th, 2013

Michael J Becich, MD PhD - CoSBBI Co-Director
Joyeeta Dutta-Moscato, PhD Candidate - CoSBBI Co-Director

CoSBBI team: Amie Draper, Arielle Fisher, Adam Handen & Andy King (PhD candidates)

Michael J. Becich, MD PhD - Chairman and Professor, becich@pitt.edu
Department of Biomedical Informatics - http://www.dbmi.pitt.edu
Associate Director Cancer Informatics, University of Pittsburgh Cancer Institute
Former Chairman of Pathology, UPMC-Shadyside & Hillman Cancer Center
Professor of Pathology, Info Sciences/Telecomm, Clinical/Translational Sciences
Co-Host, Pathology Informatics 2013 (see http://www.pathinformatics.pitt.edu)
Biomedical informatics: The broad discipline concerned with the study and application of computer science, information science, informatics, cognitive science and human-computer interaction in the practice of biological research, biomedical science, medicine and healthcare. Bioinformatics, imaging informatics, clinical informatics and public health informatics can be considered as sub-domains within biomedical informatics.
Background: Edward Shortliffe
Biomedical Informatics Applications

Basic Research

Applied Research

Biomedical Informatics Methods, Techniques, and Theories

Bioinformatics
Imaging Informatics
Clinical Informatics
Public Health Informatics

Molecular and Cellular Processes
Tissues and Organs
Individuals (Patients)
Populations And Society

Key Components of Knowledge Engineering

Knowledge Engineering is a cyclical process fundamental to the management of data to knowledge. This process is key to many if not most Biomedical Informatics research applications yielding software to enhance discovery.
Two Connected Worlds...

Biomedical Informatics

Patient Records

Clinical Trials

Disease Database

.name
.synonyms
.related/similar diseases
.subtypes
.etiology
.predisposing causes
.pathogenesis
.molecular basis
.population genetics
.clincal findings
.system(s) involved
.lesions
.diagnosis
.prognosis
.treatment
.clinical trials......

Disease World

OMIM Clinical Synopsis

Pubmed

Regulome

Transcriptome

miRNAome

Proteome

Interactome

Metabolome

Physiome

Pathome

Genome

Variome

Pharmacogenome

382 “omes” so far........

and there is “UNKNOME” too -
genomes with no function known

http://omics.org/index.php/Alphabetically_ordered_list_of_omics

Two Connected Worlds...

Biomedical Informatics

Patient Records

Clinical Trials

Disease Database

.name
.synonyms
.related/similar diseases
.subtypes
.etiology
.predisposing causes
.pathogenesis
.molecular basis
.population genetics
.clincal findings
.system(s) involved
.lesions
.diagnosis
.prognosis
.treatment
.clinical trials......

Disease World

OMIM Clinical Synopsis

Pubmed

Regulome

Transcriptome

miRNAome

Proteome

Interactome

Metabolome

Physiome

Pathome

Genome

Variome

Pharmacogenome

382 “omes” so far........

and there is “UNKNOME” too -
genomes with no function known

http://omics.org/index.php/Alphabetically_ordered_list_of_omics
Mining Bio-Medical Mountains & YOU
How Computer Science can help Biomedical Research and Health Sciences
The demise of expert-based practice is inevitable

Human Cognitive Capacity

Facts per Decision

1990 2000 2010 2020

1000

100

10

5

Proteomics and other effector molecules

Functional Genetics: Gene expression profiles

Structural Genetics: e.g. SNPs, haplotypes

Decisions by Clinical Phenotype

William Stead:

http://courses.mbl.edu/mi/2009/presentations_fall/SteadV1.ppt
Bioinformatics Opportunities

Bioinformatics
Biochemistry
Biology
Computer Science
Computer Engineering
Mathematics
Physics
Linguistics
Education, Sociology, Philosophy, Psychology, Community Studies

A research degree in any of these majors will take you far!

B.S. (B.A.)

M.S. (M.A.)

Ph.D.

Entry-Level - Company
National Laboratory
Teaching - Private Schools

Director/Professor - University
Company (Pharmaceutical)
National Laboratory
Research Foundation

Research Staff - Company/University
National Laboratory
Research Foundation
Teaching - Community College
Public Schools
Informatics is one of ten “ahead of the curve” careers

Careers that are “relatively new, already viable, and promise further growth...” (Nemko, 2008)

Health Informatics Specialist/Manager. "The job market for health informatics people is absolutely out of sight," exclaims Merida Johns, founding director of the graduate program in Health Informatics at the University of Alabama-Birmingham. And it’s no surprise: Hospitals, insurers, and regional collaboratives are switching to electronic medical records. Nurses and doctors, urged to do more evidence-based medicine, are using computerized expert systems to guide their diagnoses and treatment recommendations. Healthcare providers are also collecting more data to evaluate quality of care.
The number 1 FAQ:

How much biology should I know??

No simple or straight-forward answer... unfortunately!

But the mantra is:

Take the classes and
Interact routinely with biologists

High School Senior Summer Internships
Summer Undergraduate Research Fellowship
But I want to start with some basics..

2. http://www.ebi.ac.uk/2can/

Books

1. Introduction to Bioinformatics by Teresa Attwood, David Parry-Smith
2. A Primer of Genome Science by Gibson G and Muse SV
4. Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology by Dan Gusfield
5. Bioinformatics: Sequence and Genome Analysis by David W. Mount
6. Discovering Genomics, Proteomics, and Bioinformatics by A. Malcolm Campbell and Laurie J. Heyer
Biological Challenges - Computer Engineers

• Post-genomic Era and the goal of bio-medicine
  - to develop a quantitative understanding of how living things are built from the genome that encodes them.

• Deciphering the genome code
  - Identifying unknown genes and assigning function by computational analysis of genomic sequence
  - Identifying the regulatory mechanisms
  - Identifying their role in normal development/states vs disease states
Biological Challenges - Computer Engineers

• Data Deluge: exponential growth of data silos and different data types
  - Human-computer interaction specialists need to work closely with academic and clinical biomedical researchers to not only manage the data deluge but to convert information into knowledge.

• Biological data is very complex and interlinked!
  - Creating information systems that allow biologists to seamlessly follow these links without getting lost in a sea of information - a huge opportunity for computer scientists.
A major goal in molecular biology is **Functional Genomics** - Study of the relationships among genes in DNA & their function - in normal and disease states.
Informatics – Biologists’ Expectations

• Representation, Organization, Manipulation, Distribution, Maintenance, and Use of information, particularly in digital form.

• Functional aspect of bioinformatics: Representation, Storage, and Distribution of data.
  - Intelligent design of data formats and databases
  - Creation of tools to query those databases
  - Development of user interfaces or visualizations that bring together different tools to allow the user to ask complex questions or put forth testable hypotheses.
DBMI Academic Degree Options

Biomedical Informatics (SOM)
- Certificate
- MS
- PhD
- Post-Doc Program

Intelligent Systems (A&S)
- MS and PhD Biomedical Informatics Track
  - General BMI Track
  - Specialized Concentrations
    - Clinical Informatics
    - Bioinformatics
    - Dental Informatics
    - Health Services Research
    - Infectious Disease Informatics (Biosurveillance)
    - Clinical and Translational Science Informatics
  - On Line Certificate Program
  - Clinical Informatics Boards (2013)

Twenty-fifth year as a formal training program in biomedical informatics and successfully renewed training program grant this year for five more years (to 2017).
Biomedical Informatics Trainees

35 trainees for 2013-14 year

**Status**
- New: 7
- Continuing: 15
- “Finishing Up”: 13

**Gender**
- Male: 25
- Female: 10

**Concentrations**
- General: 23
- Bioinformatics: 4
- Dental Informatics: 4
- Health Services Research: 1
- Global Health Informatics: 3

**Biomedical Informatics**
- PhD: 22
- Masters: 7
- Certificate: 3

**Intelligent Systems**
- PhD: 3

**Prospective Students**
We offer a range of training experiences to accommodate diverse backgrounds and aspirations. Active participation in research and development is a key element of the educational experience at DBMI.

**Training Program**
- Masters Degree (MS) in Biomedical Informatics
- Doctoral Degree (PhD) in Biomedical Informatics
- MS or PhD in Intelligent Systems - BMI Track
- MD/PhD in Biomedical Informatics
- Non-degree Postdoctoral Training
- Certificate in Biomedical Informatics

**Apply to the Program**

**Questions?**
dbmi@pitt.edu

Why choose Pittsburgh?
Why choose DBMI?
Summary of Faculty Research

- Translational Bioinformatics (5)
- Clinical/translational informatics (5)
- Biosurveillance/public health informatics (4)
- Machine learning and biostatistics (4)
- Pathology informatics (3)
- Pharmacogenomics (1)
- Human computer interaction (1)
- Vocabularies and ontologies (1)
- Knowledge Representation (1)
- Imaging informatics (1)
- Tutoring systems (1)
Figure 3. Practical model for the design and execution of translational informatics projects, illustrating major phases and exemplary input or output resources and data sets.

http://www.ploscompbio.org/article/info:doi/10.1371/journal.pcbi.1002826
Corporate/Industry Successes

- De-ID Data Corp – De-identification Software - Cooper and Saul

- Omnyx – GE and UPMC Joint Venture in Pathology Imaging - Becich

De-ID Data Corp = http://www.de-i-data.com/

Omnyx = http://www.omnyx.com
• Gerry Douglas, PhD - Director, Center for Health Informatics for the Underserved
  - Founder of Baobab Health (see http://www.baobabhealth.org)
  - Won prestigious T.E.D. (Tech, Enter, Design) Award for his work


Slide 22
Baobab Health - 10 years of work in Malawi
Based on “open EMR” and CDA

16 SITES DEPLOYED AND 7 SITES IN DEPLOYMENT

KEY RESULTS ACHIEVED

VISION & VALUES

~1,300,000 patients registered

40,000 lab tests across 3 sites

185,000 X-Ray/Ultrasound investigations

50,000 patients on ART therapy across 9 sites

160,000 HIV testing and counseling sessions across 5 sites

Baobab vision is to use technological innovation to improve the delivery of healthcare in the developing world

Central Region

- KCH
- Lighthouse Trust
- Martin Preuss
- M’bangombe
- Dedza DH
- Salima DH
- Kasungu
- N’goni HC
- Ntcheu DH
- St. Gabriel MH
- Mangochi DH
- Matawale HC
- Bwaila Hospital

Queen Elizabeth
- Machinga DH
- Zomba CH

Founded in 2000
46 employees of whom 43 are Malawian
43 Colleges offer Bioinformatics to Undergrads

- Baylor University
- Gannon University
- Iowa State University
- Loyola University
- Ramapo College of New Jersey
- Rensselaer Polytechnic
- Rochester Institute of Technology
- Saint Bonaventure University
- St. Vincent College
- Stevens Institute of Technology
- SUNY
- University of Buffalo
- University of Maryland: Baltimore County
- University of Nebraska- Omaha
- University of Pennsylvania
- University of Pittsburgh
- University of Toronto
- UC – Irvine
- UC – San Diego
- UC – Santa Cruz
- University of Waterloo

Modified from Peter Randall and Andrew King
Top 21 Bioinformatics Degree Programs

http://colleges.findthebest.com/d/o/Bioinformatics

[Diagram showing the top 21 bioinformatics degree programs with universities listed and their rankings and enrollment numbers]

NOT CONFIRMED
Informatics / Computer Science Jobs

Did this in 2012 = 19,526 jobs
Currently 19,816 (2014)

1 - 25 of 22,083 jobs found
BIOINFORMATICS  http://www.cs.pitt.edu/undergrad/bioinformatics/

ABOUT THE DEGREE

Bioinformatics is the theory, application and development of computing tools to solve problems and create hypotheses in all areas of biological sciences. Biology in the post-genome world has been and continues to be transformed from a largely laboratory-based science to one that integrates experimental and information science. Bioinformatics has contributed to advances in biology by providing tools that handle datasets too large and/or complex for manual analysis. Examples of some of these tools include assembling the DNA sequence of entire genomes, gene finding algorithms, microarray expression analysis, molecular systems modeling, and biomarker discovery from mass spectra. Computational tools are central to the organization, analysis and harvesting of biological data at the level of macromolecules, cells and systems. Consequently, there is a growing need for trained professionals who understand the languages of biology and computer science. Biologists trained in more traditional programs may not have a working knowledge of statistics and algorithms, whereas computer scientists trained in more traditional programs may not have a working knowledge of the chemistry and biology required in the field.

The Undergraduate Bioinformatics Degree at the University of Pittsburgh, which is operated jointly by the Departments of Biological Sciences and Computer Science, program offers training that builds a solid foundation in chemistry, biology, computer science, mathematics and statistics. This training will enable students to communicate fluently with experts across these disciplines, and to have the skills necessary to apply computing tools to address contemporary problems in biology and medicine. The training will enhance the professional opportunities for undergraduates to pursue careers in pure or applied research in academia, government, pharmaceutical, medical, or biotechnology sectors.
U Pitt's Bioinformatics Course Curriculum

- Curriculum Loaded with Biology and Computer Science
- Focused learning in:
  - Chemistry
  - Math
  - Statistics
- Includes a “real” research experience with “Capstone” mentored research project.
- 2 yrs electives (CoSB & BMI)
- Work your summers in a research laboratory with the Department of Biomedical Informatics
  - Now a job guarantee!!!
  - i-STEM (Informatics for Science Technology Engineering and Math)
  - 501c3 (not for profit) now proposed to fund summer research for high school and college undergrads
Bioinformatics at Pitt
by Andy King (Grad Student in DBMI)

- Biology: Bio 1 & 2, Genetics, Biochem, Computational Bio
- Computer Science: The four core courses of a CS degree
- Chemistry: Chem 1 & 2, Ochem 1 & 2
- Statistics: Stat 1000 and Applied regression
- Math: Calculus 1 and Discrete Mathematics
- 12 credits of upper level electives
- Capstone & Research
- Labs not required
What does informatics mean to biologists?

The ultimate goal of analytical bioinformaticians is to develop predictive methods that allow biomedical researchers and scientists to model the function and phenotype of an organism based only on its genomic sequence. This is a grand goal, and one that will be approached only in small steps, by many scientists from different but allied disciplines working cohesively.
the Ultimate Goal

Disease World

Biomedical Informatics

Biology & Bioinformatics

PubMed

Disease Database

- Name
- Synonyms
- Related/Similar Diseases
- Subtypes
- Etiology
- Predisposing Causes
- Pathogenesis
- Molecular Basis
- Population Genetics
- Clinical findings
- System(s) involved
- Lesions
- Diagnosis
- Prognosis
- Treatment
- Clinical Trials

- Personalized Medicine
- Decision Support System
- Outcome Predictor
- Course Predictor
- Diagnostic Test Selector
- Clinical Trials Design
- Hypothesis Generator

Computer Scientists

Patient Records

Clinical Trials

Genome

Regulome

Transcriptome

Variome

Proteome

Interactome

Metabolome

Physiome

Pharmacogenome

Pathome

OMIM

► Personalized Medicine
► Decision Support System
► Outcome Predictor
► Course Predictor
► Diagnostic Test Selector
► Clinical Trials Design
► Hypothesis Generator

Regulome

Transcriptome

Variome

Proteome

Interactome

Metabolome

Physiome

Pharmacogenome

Pathome

OMIM

► Personalized Medicine
► Decision Support System
► Outcome Predictor
► Course Predictor
► Diagnostic Test Selector
► Clinical Trials Design
► Hypothesis Generator

Regulome

Transcriptome

Variome

Proteome

Interactome

Metabolome

Physiome

Pharmacogenome

Pathome

OMIM

► Personalized Medicine
► Decision Support System
► Outcome Predictor
► Course Predictor
► Diagnostic Test Selector
► Clinical Trials Design
► Hypothesis Generator
“To him who devotes his life to science, nothing can give more happiness than increasing the number of discoveries, but his cup of joy is full when the results of his studies immediately find practical applications”

— Louis Pasteur

"And that’s why we need a computer." & YOU

Thank You!
References Courtesy of Bill Hersh:

Biomedical and Health Informatics: Improving Health, Healthcare, and Biomedical Research with Information Technology

William Hersh, MD
Professor and Chair
Department of Medical Informatics & Clinical Epidemiology
Oregon Health & Science University
Portland, OR, USA
Email: hersh@ohsu.edu
Web: www.billiherhsh.info
Blog: informaticsprofessor.blogspot.com


Chapter 3

BIOINFORMATICS CHALLENGES AND OPPORTUNITIES

Peter Tarczy-Hornoch\textsuperscript{1} and Mark Minie\textsuperscript{2}

\textsuperscript{1}Department of Medical Education and Biomedical Informatics; \textsuperscript{2}Health Sciences Libraries, University of Washington, Seattle, WA 98195
E-mail me if you want a copy of this book and I will have it shipped to an address of your choosing (send it to me in your e-mail)
The deadline for submission has now passed. We are not accepting any additional applications. Thank you!

The Association for Pathology Informatics Summit 2014 Travel Awards will provide stipends for documented trainees to attend the Pathology Informatics Summit 2014 Conference (PI 2014) at the Wyndham Grand Pittsburgh Downtown, Pittsburgh, PA on May 13-16, 2014.

About the Award
The travel award is intended to offset the costs of attending the Pathology Informatics 2014 conference for trainees to increase involvement in pathology informatics within the broader pathology community. The Pathology Informatics 2014 conference will provide a limited number of travel awards with a maximum reimbursement of $1,200 each.

Eligibility
Applicants must be trainees in an official training program. The trainee's supervisor must be willing to document the trainee's status in writing in order to be eligible for the award (see application guidelines below). All applicants will be considered regardless of age, race, gender, national origin, sexual orientation, or religion. Trainees may include:

- Residents or fellows in a training program accredited by the Accreditation Council for Graduate Medical Education (ACGME)
- Masters, pre-doctoral and post-doctoral students

http://www.pathologyinformatics.com
Association for Pathology Informatics (API)

http://www.pathologyinformatics.org

“...to advance the field of pathology informatics as an academic and a clinical subspecialty of pathology...”

Supporting Sponsors

Associate Membership

Students, residents, and fellows in training.
- Access to API “members only” Listserv.
- One publication if accepted, at no charge, for Journal of Pathology Informatics ($200 value).
- Eligible to serve on API committees.

$50 membership cost supported by either the Peter J. Becich Grant* or Teaching Institutional membership programs.

*The Peter J. Becich Educational Grant was established in 2003 by API’s first President, Michael J. Becich, MD, PhD, in honor of his father, Peter J. Becich. The purpose of the grant is to offer a full membership rebate to Associate members (individuals in training, including students, residents and fellows). Since its inception, over 200 Annual Associate memberships have been funded by the Becich Grant.
JPI and Assn for Pathology Informatics are the Home for Digital Pathology - Great Academic and Strategic Partnership with Multiple Benefits!!! YOU have the opportunity to publish in JPI!!!