SJSU
SAN JOSÉ STATE UNIVERSITY

MSBI Orientation and Welcome Back
Computer Science Department

January 26, 2021
7:00PM PST
Welcome New Master's Students and Welcome Back Everyone!

• Getting started – including what’s new this year
• Professors: MSBI faculty members, student advisor, project advisors, and program coordinators
• Program requirement review & Technical Writing
• GPA and grad standing
• Academic honesty
• Writing project and research topics - project advisors introduce their topics: Oouverney, Andreopoulos, Heller, Lee, and Wesley.

Q&A
Getting Started

• We're glad you're here!
• Have you registered for classes?
  • Make sure to do this as soon as possible.
  • If a course you like starts full, remember people often drop after a few days, so be persistent.
• Have you found your class meeting info?
  • Because of Covid-19 classes are still online.
  • Faculty either are listing how to join class meetings (usually Zoom) on Canvas, via directly e-mailing you, on their syllabus, by giving the details to the CS Office, or a combination of these.
Getting Started

• If you are looking for a class you are not currently enrolled in, check if the faculty has left the meeting details generally visible on Canvas. If not, look at the online syllabus for the class. If not, use your phone/email-a-friend lifeline. If not, try directly emailing the faculty or CS office (Lauren.Elliott@sjsu.edu)

• Make sure to check your SJSU email and Canvas at least daily. You should probably keep a separate offline file with the meeting info for each of your classes. If you are in a different time zone, make sure to take account for the difference in time and double check for annoyances like daylight savings time.
Other Resources to Started

• Keep track of what's happening in the department on the Department Web Page.
• Master's program rules can be found at the MSBI website.
• Meet your advisor.
"Advisors": MSBI faculty members, student advisor, project advisors,

- **MSBI faculty members**: CS Professors Andreopoulos, Heller, Lee, Wesley, and Prof. Ouverney (Biology).
- **Before CS 297**: MSBI Student Advisor: Prof. Lee.
- **After CS 297**: Master Project (297/298) Advisors:
  - You need to meet with an advisor... now, and once every semester.
  - You can only register after your advisor removes the advising hold.
- **Things to talk about**:
  - What classes you plan to take this semester.
  - How you are currently doing (GPA-wise).
  - What's your plan for classes in future semester.
  - Any questions you have about the program.
... Beyond Advisors

- Program Coordinators: M. Moh (CS) and Ouverney (Biology)
- You can contact us to talk about getting signatures on MSBI Program Forms.
- For many typical situations, you can just fill out the form and send it to the Lauren Elliott of the CS office to set up a docuSign for one of us to sign it.
- CS Department Chair: M. Moh.
- College of Graduate Studies
Program Requirement Review

• Prerequisites: Listed on your admission letter.
• Program Requirement:
  • Core Courses: 19 units (including CS 200W)
  • Electives: 6 units
  • Culminating Experience: 297/298: 6 units.
• Total: 31 units.
• Refer to the MSBI website.
• The typical way to pass the Graduate Writing Assessment requirement is to take either CS200W.
• The Computer Science Department policy is that students should **complete the GWAR requirement in their first year** in the MS program and must complete it two semesters before graduation.
• The writing requirement is required by the Cal State system.
• The only substitutes are:
  • Passing the WST at waiver level
  • Getting special permission from Graduate Studies (e.g. if you wrote a Ph. D. thesis in another subject)
• To obtain a waiver you would also need to fill out the **GWAR Waiver Petition Form** and get the relevant signature (not easy).
• **If you get a waiver, you still need to take 31 total units** for our program, so you will need to choose an additional three-unit elective course to take in the place of CS200w, and get a course substitution form approved by the Program Coordinator.
GPA and Grad Standing

- If your **total GPA is less than 3.0**, you will end up on academic probation.
- You can score less than 3.0 (a B) on a course and not be on probation. However, you cannot use any course whose grade was less than or equal to a C- towards graduation.
- Being on probation means you cannot graduate until you're off probation.
- It also means you will be disqualified from the MSBI program if your GPA is not above 3.0 in the next semester.
- More information on this can be found on the [MSBI FAQ](#).
Academic Honest

• Academic Honesty is taken very seriously at SJSU.
• Make sure that any homework assignments, reports, etc. are written in your own words, use your own programming code.
• All CS298/CS299 reports will be checked against turnitin.com and it would be a pity to not complete your master's when you are very close to finished.
The Writing Project

• To finish your master's degree you need to complete one of two culminating experiences:
• A writing project course sequence such as CS297/CS298 and successful defense.
• A thesis course sequence such as CS297/CS299 and successful defense.
• The thesis (299) has stricter requirements and a faster timeline. The writing and formatting is checked by the Graduate Studies in addition to by your thesis advisor and committee.
• For both a writing project and a thesis, you need to find a project advisor (to take 297 and 298 or 299 from) and two committee members (this should be done by 298).
The Writing Project or Thesis

• You should try (this is a suggestion not a requirement) to take a course from the person who you want as your advisor before starting work under that person.
• You will probably have a better project experience if you choose an advisor that doesn't have too many students.
• Usually, it is the advisor that comes up with project suggestion in consultation with what the student is interested in.
• After Academic Honesty Policy, we will look at possible project advisors:
  • Prof. Ouverney (Dept. of Biology).
  • Prof. Andreopoulos, Heller, Lee, and Wesley (Dept. of Computer Science).
The Uncultured Majority

Cleber Ouverney

cleber.ouverney@sjsu.edu
We are 10% human, 90% microbes

$10^{13}$ Human cells and $10^{14}$ bacterial cells, hence 10% human and 90% bacteria

http://www.miller-mccune.com/science-environment/bacteria-r-us-23628/
Uncultured bacteria associated with human diseases

Access to human samples imposes a challenge.
Mixed Community

Cell shape, size, density, target molecules, autofluorescence

Single Cell

Pure Culture or Enrichment

Cell lysis

WGA

Novel Lactase

Deinococcus pimensis

E.coli express heterologous lactase

Oral sample
Acknowledgements

Funding: NIH (MBRS, MARC, RISE), NSF (RUI), HHMI

Adam Caldwell
Eamon Vandaei
Iana Elledge
Michael Abrams
Farsheed Ghadiri
Margarita Rangel
Stephanie Nystrom
Diana Romero
Becky Lee
Cecilia Brown
Ivan Cheng

Joseline Santos
Joshua Garcia
Ivan Cheng
Sharina Santos
Paige Grabski
Sam Smits

David Barton
Deepa Surendar
Kavitha Reddy
Jamsheed Ghadiri
Jorge Dinis
Sam Smits

David Barton
Deepa Surendar
Kavitha Reddy
Jamsheed Ghadiri
Jorge Dinis
Sam Smits
Build machine learning models of 90 carbohydrolytic-active enzymes that were published in Science (Hess et al., 2011)

Then use ML models to predict action of 28,000 putative enzymatically active genes

Learning outcomes: develop expertise on machine learning tools, how to productionize them and run on computing platforms.
Cyanobacteria metagenomic samples across multiple temperatures (JGI)
Find new strains through clustering polymorphisms and translocations between 16Ss
Eventual goal is to do ribosomal 16S assemblies
Conventional wastewater treatment without microbial aid are inefficient and expensive for extracting heavy metal contaminants
Microbes can also remove contaminants from water or waste streams by converting them into a less toxic form
Separation of thousands of plasmid sequences from soil microbes with deep learning
Applications to Cyanobacteria nitrogen-fixation bacteria
Learning outcomes: Use deep learning on Pacific Computing Platform for this analysis
RNA-Seq analysis of Phaeodactylum tricornutum: a diatom algae found in oceans

Find differentially expressed genes in transcriptomic samples from different conditions (oilfield water vs. normal)

Learning outcomes: Build expertise on RNA-Seq analysis tools
Phil Heller's Bioinformatics Research
philip.heller@sjsu.edu
Traditional Metagenomics
Marine Ecology, Climate Change, and Deep Learning
Bioinformatic Astronautics

GeneLab
Open Science for Life in Space
Antibiotic resistance in Wastewater / Biofilms

Neurodevelopmental Disorder

Gut-Brain Axis

Microbiota

THE LEE LAB
www.sjsu.edu/people/wendy.lee

SARS-CoV-2

Systematic NGS Errors

Genome in a Bottle

Thrombosis in Microgravity

APPLIED BIOINFORMATICS
Spring 2021
Leonard Wesley: Areas Of Bioinformatics R&D

• Three (3) Areas Of Interest/R&D via (Industry Funded Projects):

1. Identify Genetic Biomarkers Beyond Beta-Amyloid & Tau For Alzheimer’s Disease & Mild Cognitive Impairment:
   • The relatively new ATN (“A” for amyloid deposition, “T” for tau levels, and “N” for neurodegeneration) framework proposes describing Alzheimer's Disease (AD) and cognitive impairment (CI) based on a patient's biomarker profile. The biomarker profile consists of several metrics that include cerebrospinal fluid (CSF) total tau, F-fluorodeoxyglucose-positron emission tomography, accumulation of two key pathogenic proteins, amyloid beta (Aβ), and brain atrophy.
   • Most will now agree that CI and AD are complex multifactorial diseases, and Aβ and tau cannot account for all aspects of CI and AD.
   • There is a growing effort to identify single nucleotide polymorphisms (SNPs) associated with cognitive decline independent of amyloid β (Aβ) and tau pathology in CI and Alzheimer’s disease (AD).
   • Two objectives of the proposed project include:
     1. Conducting a similar analysis as that of the Hang-Rai Kim et al group using the ADNI-1 and ADNI-60/2 data set. Then assess the degree to which our findings confirm or refute the conclusions found by Hang-Rai Kim et al.
     2. Conduct a similar analysis as Hang-Rai Kim et al but consider AD related pathologies such as TDP-43 and demographic factors such as race and gender.
2. **Financial Capacity Index (FCI) Prediction:**

   - Loss of financial capacity (FC) is one of the earliest IADL changes to occur in patients with mild AD, which increases dependence of patients on their family members [1]. FC is the complex cognitive ability to carry out a wide range of daily activities, from the simple task of counting coins to more complex task like paying bills and making financial decisions. It can be measured with the financial capacity instrument (FCI) developed by Marson et al.

   - The objective of this research is to apply the NODDI algorithm to the MRI images to measure the white matter integrity of the brain and investigate how it relates to financial capacity of Alzheimer’s patients and healthy older adults, in order to determine which parts of the brain are necessary for the financial capacity of patients. Identified regions and cognitive tests will be used to build a machine learning model to predict FCI scores of individual subjects.
Leonard Wesley: Areas Of Bioinformatics R&D

3. Pancreatic Cancer (PC):
- 95% of patients diagnosed with PC die within 5 years from diagnosis.
- The sooner diagnosis can be made, the better the treatment options and outcomes.
- The current project involves identifying better PC biomarkers based on genetic profiles that are obtained from DNA that is extracted from liquid biopsies.
- Involves using Nanopore & Ion Torrent NSG technologies to sequence real pancreatic DNA/RNA, genome assembly of sequence results, and evidence-based analysis to predict PC sooner than currently achievable.
- Use evidence-based mathematical models & methods to predict PC.
Leonard Wesley: Areas Of Bioinformatics R&D

• Additional R&D Areas Of Interest:

  • Protein Structure Prediction & Protein Structure Scoring:

  • Small Chemical Molecule Affinity and Activity Prediction With Protein Targets:

  • Genotypic and phenotypic pattern analysis using statistical and evidence-based methods to identify novel drug targets.
Questions?

• Hopefully, this orientation was helpful.
• Good luck to everyone here at SJSU!
Thank you!