Intro to Software Engineering

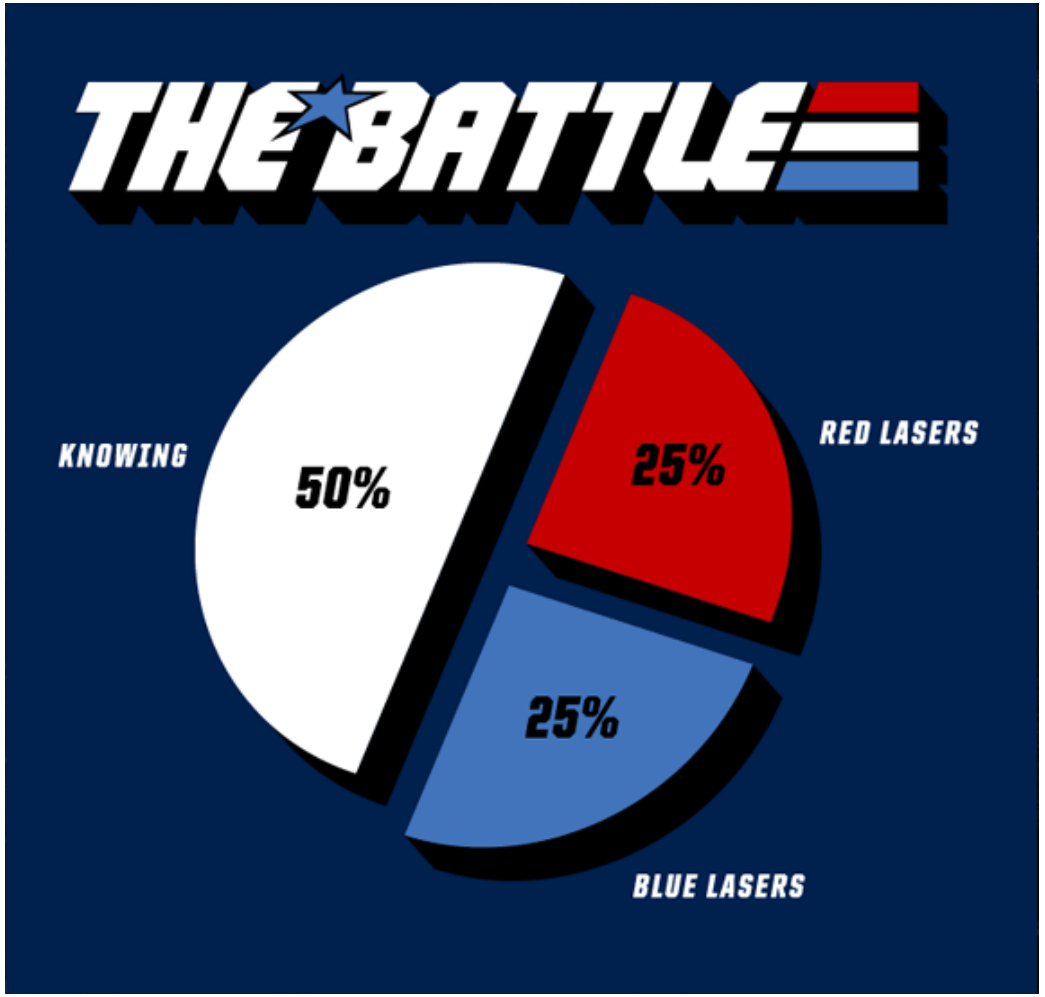
Andrew Hu

Administrivia

- Drylab Syllabus
- Git/SSH Lab through 4/8
  - Sign up on the Doodle (check Zulip #drylab)
  - Bring your laptop
  - Attendance mandatory, if you really know this stuff, talk to us and we may excuse you

Syllabus Overview

- Class is mandatory
- Set goals, and meet them reliably
- Documentation is half the battle



Class Philosophy

- A wrong answer is a good answer!
  - I would rather you give me a wrong answer than a right answer
- Talk to the people around you
- Don’t be afraid to ask questions
- Freire’s problem posing vs banking? A little bit of both

Today's Goals

- What do you think software engineering is?
- How do Software and Biology interact?
- Keeping the problem and solution together

What is Software Engineering?

Initial Ideas (Software Engineering)

My Suggestions

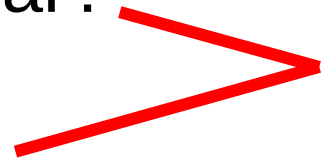
- Effective problem solving
- Dynamically adapting
- Limited scope (inch wide, mile deep)
- Secure

Is software: (high/low priority)

- Modular?
- Fast?
- Easy to use?

Is software: (high/low priority)

- Modular?
- Fast?



Maybe… Let's make it work first

- Easy to use?

Depends on the use case

Make it work, *then* make it good

Software & Biology: Why the heck are we doing this?

What does software have to do with biology?

Examples of software for biology

Computational Biology

- In many ways, biotech is the new frontier
- Many of its problems require new, interdisciplinary research
- Opportunity for creative solutions
- More than just “moving bits”

Case Study: Protein Folding

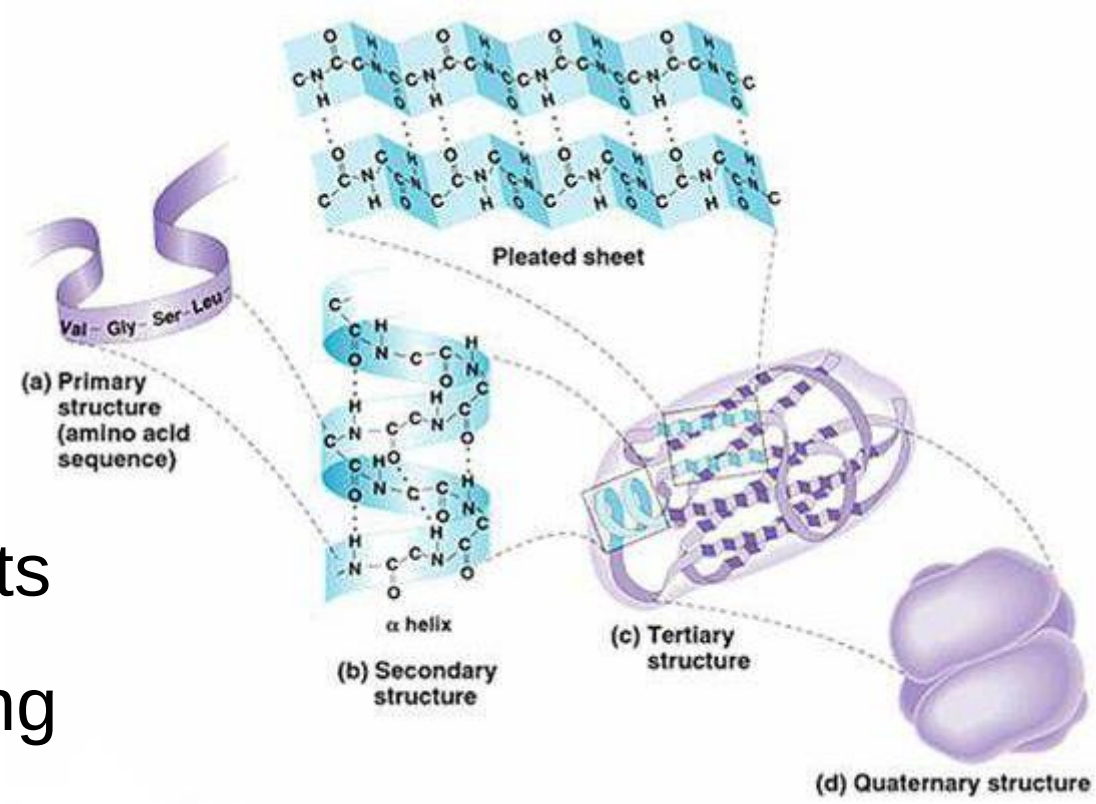
Feat. UW's very own Center for Protein Design and Center for Game Science

What is a protein?

- The machines of life
- They do everything interesting
- What your DNA actually codes for

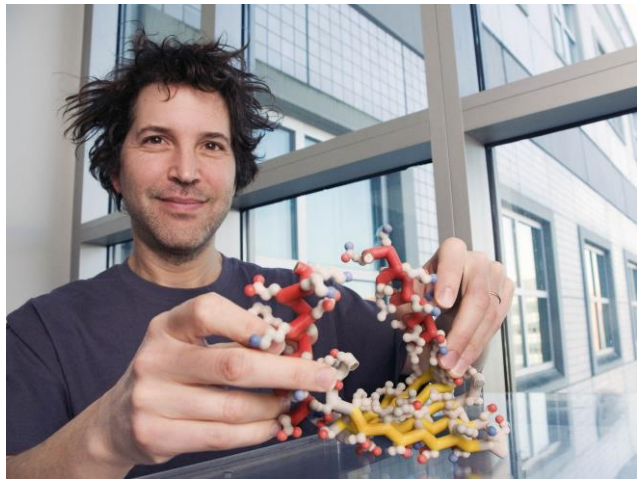
Protein Folding?

- Amino acids combine together to create a protein
- Creates a long amino acid chain
- However, each amino acid interacts with each other: This causes folding



Why do we care?

- How a protein folds determines its properties/abilities
- If we can predict how a protein will fold, we can intelligently design new proteins to perform specific tasks



<-David Baker

INSTITUTE FOR
Protein Design
UNIVERSITY of WASHINGTON

Problem: Protein Folding

- Very difficult to predict or model
- We understand the basic chemistry about why/how the amino acids are attracted to each other
- However, since each amino acid interacts with all the amino acids around it, the problem grows exponentially in difficulty as proteins get bigger

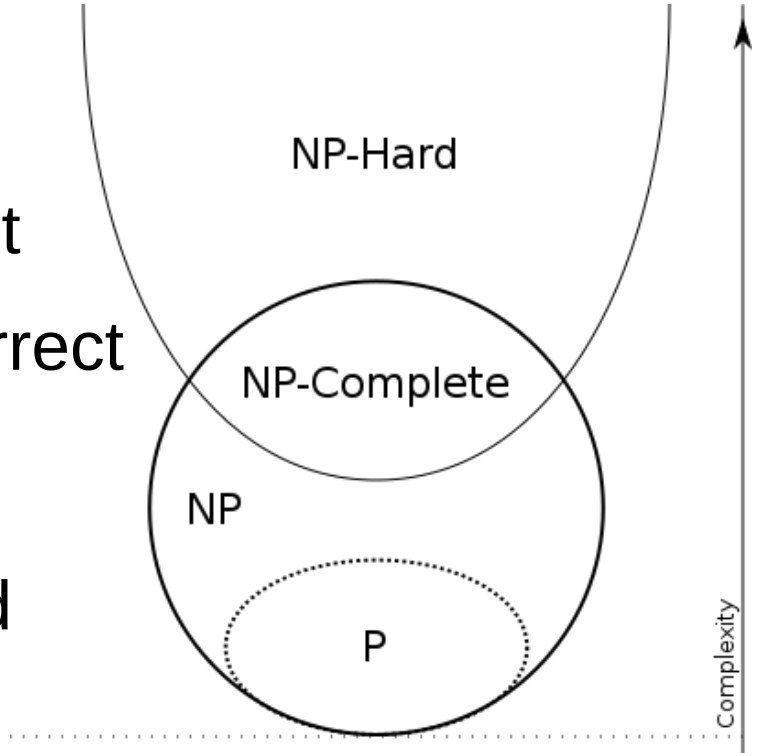
Enter: Computer Science

- Biologists stumped by this problem
- Interest of Computer Science researchers is piqued
- Theoretical computer scientists classify the problem as "NP-Hard" or "NP-Complete"

Gasp D:

Briefly, NP

- A classification of how hard a problem is to solve
- NP-Hard is the hardest type of problem
- Even with a magic genie that gives us a correct answer, we cannot verify that the answer is correct efficiently
- NP-Complete: hard, but answer can be verified efficiently

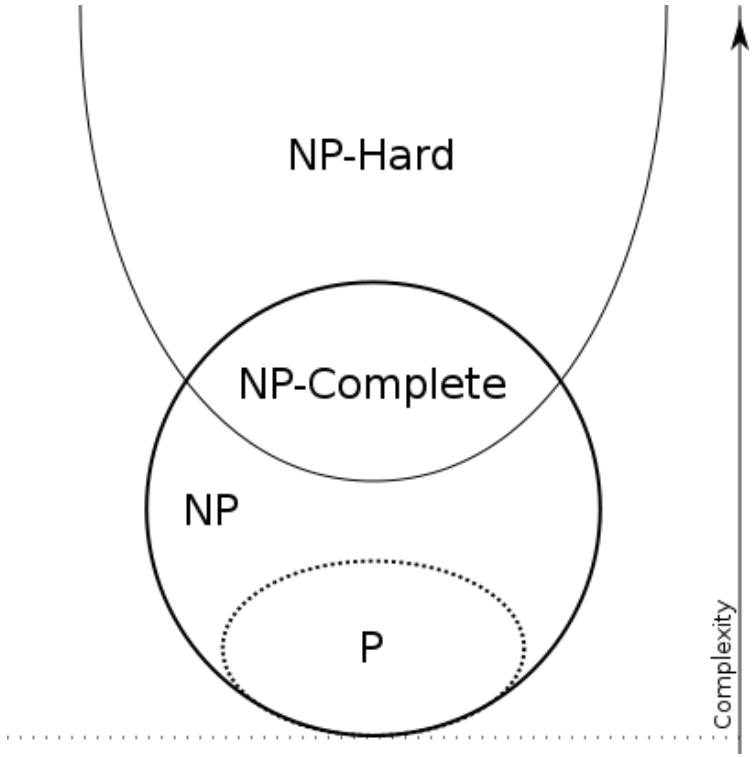


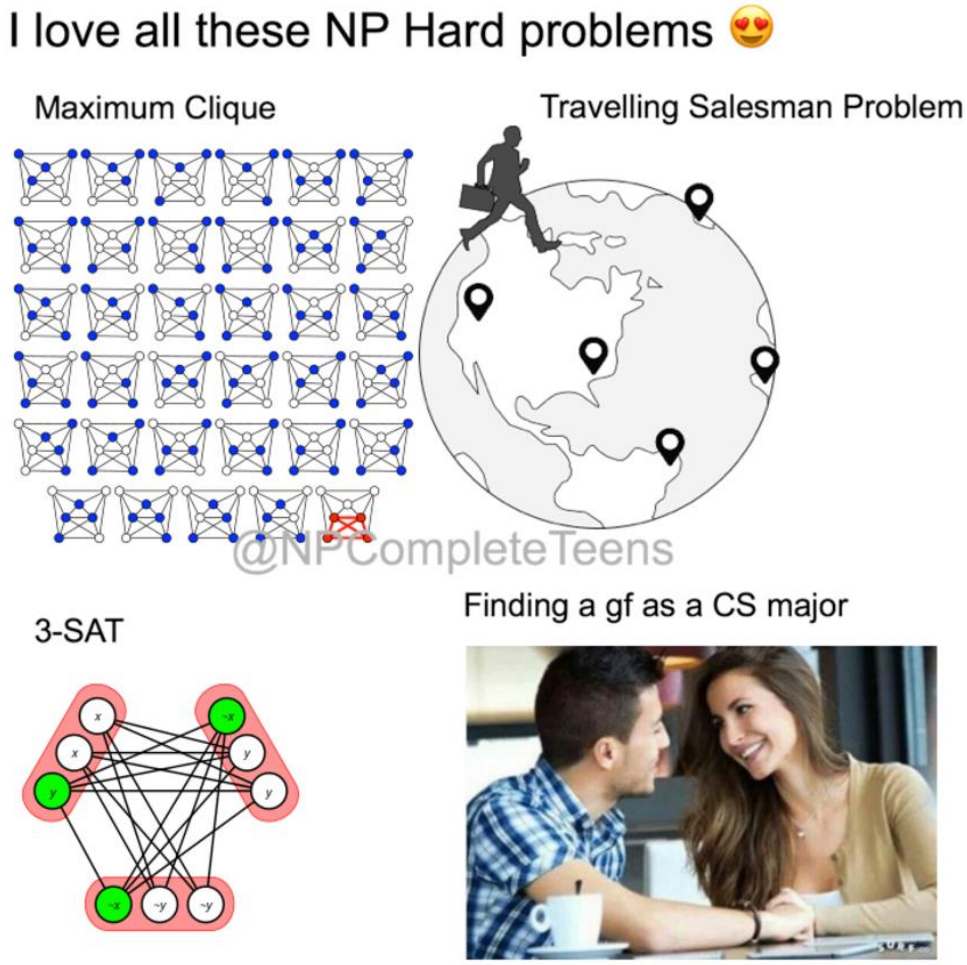
NP-Hard

Protein-Folding,
Best Chess Move

Password Cracking,
Travelling Salesman

Sorting,
Almost everything



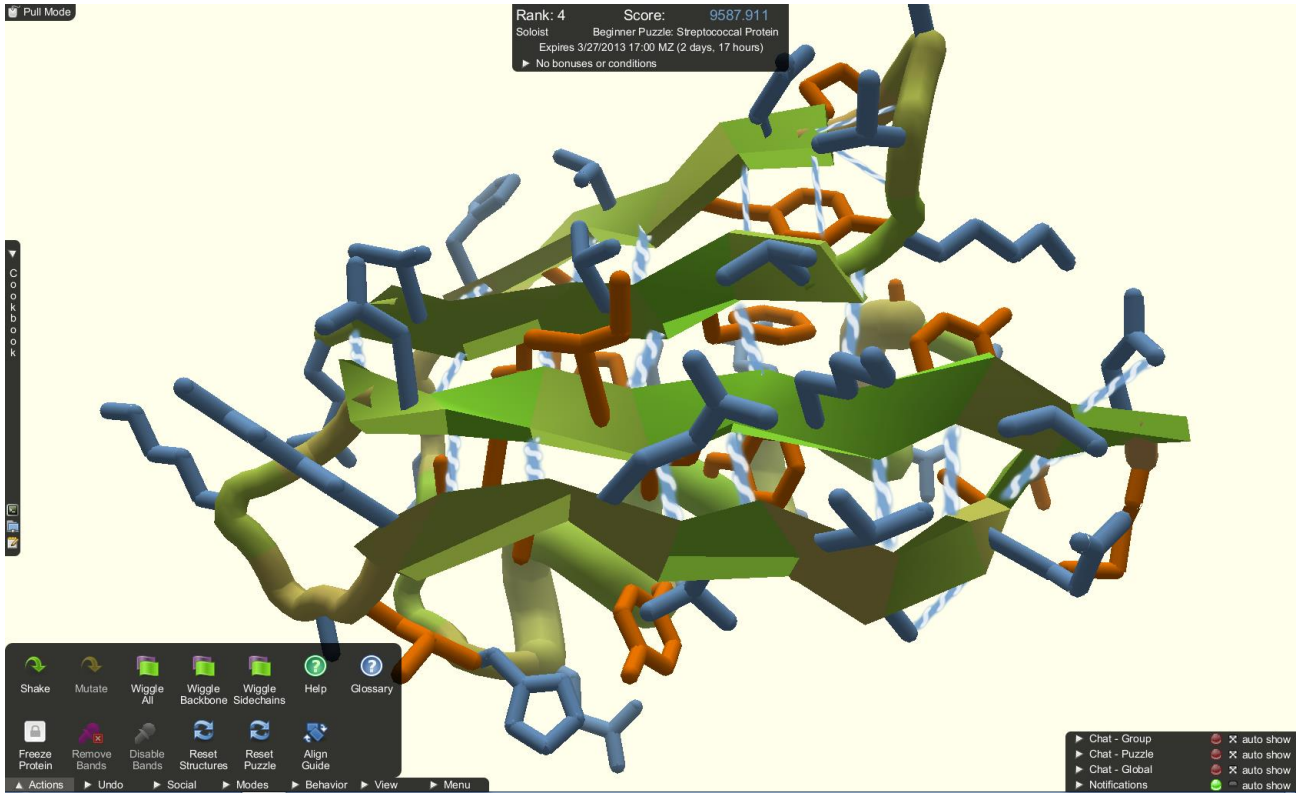


So Protein Folding is hard… Now what?

- Estimation methods
  - Statistical models & heuristics
- Intelligent Guessing
  - Using human ingenuity with Foldit
  - Not what is being solved, but how is it being solved
- A little bit of both: neural networks

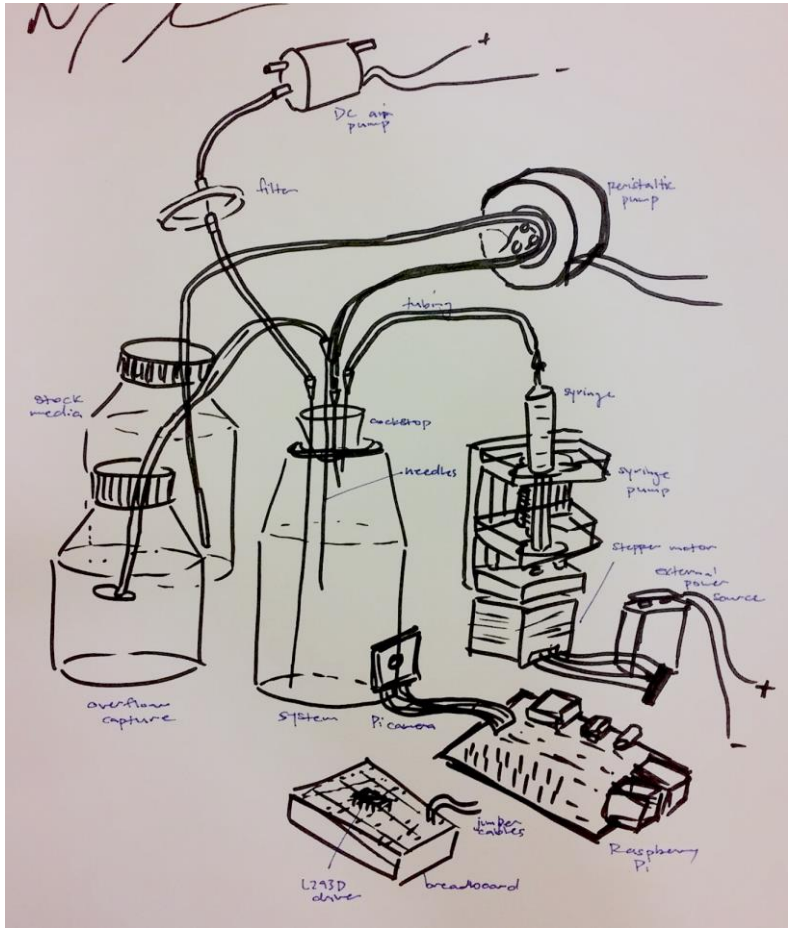
Foldit

- A puzzle game that is literally just protein folding
- Trying to figure out *how* the players solve protein folding



Case Study: Chromastat

What we were given…



It has to:

- Keep a cell culture alive
- Take input from sensors
- Inject chemicals to alter the culture state
- Treat the solution's color as an indicator of what to inject

What questions do we have about the biology?

Feedback Loop

Our Questions

What questions we didn’t come up with

- Population growth, what rate is it growing at?
  - Always in the log phase

Idea:

Do not divorce your solution from the problem

Keeping the Problem & Solution Together

- The problem is a biological problem
  - Do not pretend that it is magic or a black box
- You don’t have to know biology to know that we have questions
- Assume whoever gave you the problem has not explained it properly

Computer Scientists in Syn Bio