Computing for Life Sciences (CS 59000)

Course Content: The accumulation of biological data, such as genome sequences, protein structures and sequences, metabolic pathways, has opened up a new way of research in biology – bioinformatics. In this course we will learn programming, bioinformatics databases, tools, and algorithms behind these tools. Covered topics will include methods for protein sequence comparison, protein structure comparison, protein structure prediction/modeling, protein docking prediction, protein function prediction, and protein network analysis.

This is a core course for CLS specialization of graduate degree.
See http://www.gradschool.purdue.edu/CLS/index.cfm

Grading: Grades in the course will be based on midterm & final exam (40%), homework assignments (30%), and projects (30%). The course will be concluded by presentations of group projects. Semester grades will be awarded based on the following minima of performance: 90% = A, 80% = B, 70% = C and 60% = D; < 60% = F. These thresholds will not be raised but may be lowered at the instructor's discretion.

Projects: Considering that this course is cross-listed between different departments, students will choose either of programming-type or mini-review-type (+ using existing software) project by the consent of the instructor.

Instructor: Daisuke Kihara (dkihara@purdue.edu), Robert Skeel (skeel@cs.purdue.edu)
Professor Skeel will teach in the first 3 weeks (Aug 22nd – Sep. 9th) on basics of unix commands and Python programming. Professor Kihara will teach for the rest of the semester on bioinformatics methods.

Time & Place: MWF, 11:30am - 12:20pm, LWSN 1106

TextBook: We don't use a specific textbook. But here are good books to refer to:

- Learning Python, 2nd Edition, by Mark Lutz, and David Ascher, O'Reilly, December 2003, ISBN: 0-596-00281-5,
- Biological Sequence Analysis, Durbin, Eddy, Krogh, Mitchison, ISBN 0-521-62971-3
- Protein Bioinformatics, by Ingvar Eidhammer, Inge Jonassen & William R. Taylor, Wiley, ISBN: 0-470-84839-1

Fall 2011: Provisional Schedule

Week	Lec. #	Date	Topic
1	1	Aug. 22	Introduction of the course, Unix file system (Skeel)
	2	Aug. 24	Unix commands
	3	Aug. 26	Computing and scripting in Python
2	4	Aug. 29	Strings and files
	5	Aug. 31	Regular expressions **Homework #1 (programming) out**
	6	Sep. 2	Control structures and variables
3	-	Sep. 5	(no class, Labor day)
	7	Sep. 7	containers
	8	Sep. 9	Functions **Homework #1 (programming) in**
4	9	Sep. 12	Numbers and graphics (Skeel)
	10	Sep. 14	Sequence alignment (Kihara)
	11	Sep. 16	Sequence alignment **Homework #2 (sequence alignment) out**
5	12	Sep. 19	Amino acid similarity matrices
	13	Sep. 21	Multiple sequence alignment
	14	Sep. 23	Sequence motifs **Homework #2 turn in**
6	15	Sep. 26	Profile motifs (Cont.) **Homework #3 (motif) out**
	16	Sep. 28	Database search, Statistical Analysis
	17	Sep. 30	Database search, Statistical Analysis (cont)
7	18	Oct. 3	several databases tools for motif, domain search **Homework #3 in**
	19	Oct. 9	Hidden Markov models
	20	Oct. 7	HMM (cont.)
8	-	Oct. 10	(no class, October break)
	21	Oct. 12	Protein Structure **Homework #4 (HMM) out**
	22	Oct. 14	Protein Structure (cont.)
9	23	Oct. 17	Protein structure comparison
	24	Oct. 19	neural network: Protein secondary structure prediction

	25	Oct. 21	Protein Secondary Str. prediction cont.
10	26	Oct. 24	Template based structure modeling **Homework #4 in**
	27	Oct. 26	Template based str. modeling (cont.)
	28	Oct. 28	Geometrical properties of protein structures **Homework #5 (protein structure) out**
11	29	Oct. 31	**Mid term exam**
	30	Nov. 2	Ab initio protein prediction
	31	Nov. 4	Protein-protein interaction network
12	32	Nov. 7	Protein-protein complex: Geometric hashing
	33	Nov. 9	Side-chain optimization **Homework # 5 in**
	34	Nov. 11	Function prediction from structure
13	35	Nov. 14	**Project out**
	36	Nov. 16	Comparison of chemical molecules (drugs)
	37	Nov. 18	Ligand (drug) protein interaction
14	38	Nov. 21	Comparative genomics
	-	Nov. 23	(no class, university holiday)
	-	Nov. 25	(no class, university holiday)
15	39	Nov. 28	Systems biology (gene expression, pathways etc.)
	40	Dec. 30	Systems biology
	41	Dec. 2	**Project presentation**
16	42	Dec. 5	**Project presentation**
	43	Dec. 7	Advanced research Topics
	44	Dec. 9	(Course evaluation)
13		Dec. 12	**Final Exam Week**
		Dec. 14	
		Dec. 16	