

INTRODUCTION TO COMPUTING FOR THE STUDY OF COMPLEX SYSTEMS

2011 ICPSR SUMMER PROGRAM IN
QUANTITATIVE METHODS OF SOCIAL RESEARCH

INSTRUCTOR: DANIEL MARTIN KATZ
CONTACT: dmartink@umich.edu

TIME: 5:00 – 6:00PM

LOCATION: MICHIGAN LAB OF THE HELEN NEWBERRY BUILDING

The Study of Complex Systems is an interdisciplinary enterprise. It includes concepts from a variety of academic fields including applied math, physics, political science, economics, information science, network science, sociology, biology, computer science, etc. Relevant concepts include: Bottom Up Approaches to Modeling (with particular emphasis on game theoretic and other classes of problems that cannot be solved analytically), Diffusion, Computation, Study of Information, Dynamics of Social and Physical Systems, Large N Data Analysis, Networks and other related structures, etc.

While I will present some of the relevant theoretical underpinnings, this computing course is designed to teach the mechanics of model implementation. In my experience, mechanics are a serious impediment that many scholars face in generating models capable of being published in academic journals. Implementation is the bridge between concept and scientific contribution.

Unlike many courses that relegate implementation to self-teaching, etc. in the ICPSR Summer Course in Complex Systems, we take implementation seriously. Therefore, over the 17 sessions you will be exposed to a variety of different models, techniques and software packages.

It is my hope that the foundation offered over the next few weeks - together with some hard work on your part - will allow for application of the concepts in complex systems to the substantive questions you find interesting.

These computing sessions assume **no prior knowledge** on the part of the student – just a willing to engage with the material.

The schedule is approximate and may be changed to best serve the pace of the class.

I will post the course slides and assignments to my blog:
<http://computationallegalstudies.com/icpsr-class/>

You will also find last year's slides on this page. While there is a decent amount of overlap with the 2010 version of the lab – there will be some important changes for 2011.

GOALS FOR THE COURSE

(1) I would like you to develop a good foundational understanding of Agent Based Modeling in the Netlogo framework. While Netlogo has important limitations related to issues of scalability, etc. it is good introductory environment useful for generating a wide range of models. We will also explore a new package called Nova (developed by Richard Salter @ Oberlin College).

(2) Agent Based Models v. System Dynamics Models v. Game Theoretic Models. We will discuss the distinctions between these approaches as well as the differences in implementation associated with these modeling paradigms.

(3) Develop a solid understanding of network theory as well as the relevant implementation for both theoretical and empirical network models. For theoretical models, we will primarily use Netlogo (although Python based implementation will be discussed). For those interested in empirical networks, Pajek will be briefly introduced. For those interested in large networks, Igraph (available for R and Python) will be also discussed.

(4) Introduction to various other tools that will assist you in exploring and analyzing complex systems. These include webscraping, cloud based computational social science, computational linguistics as well as an introduction to information theory.

THE SCHEDULE

Week 1

July 18

Intro to Agent Based Models and the Netlogo Modeling Environment

July 19

Introduction to the Netlogo Modeling Environment (Cont.) and Exploration of the Forest Fire Model

July 20

Forest Fire Model (Cont.) and Intro to Schelling's Segregation Model

HW Assignment #1 Distributed

July 21

Schelling's Segregation Model: Exploration of the Model's Code

July 22 Simple Birth Rates and Predator-Prey (Wolf Sheep)

Week 2

July 25 Hawk/Dove & The Flocking Model

HW Assignment #1 DUE

July 26 Parameter Sweeps and the Analysis of Model Runs – Part I

HW Assignment #2 Distributed

July 27 Parameter Sweeps and the Analysis of Model Runs – Part II
(Guest: Eric Provins)

July 28 Assessing Agent-Based Models: Stability, Sensitivity and
Robustness in Model Outcomes
(Guest: Charles Doriean)

July 29 Assessing Agent-Based Models:
Party Discipline Regimes in the U.S. House
(Guest: Charles Doriean)

Week 3

August 1 Tools for Computational Social Science: Natural Language
Processing, Automated Content Analysis, and Cloud Computing –
Part I (Guest: Abe Gong)

HW Assignment #2 DUE

August 2 Tools for Computational Social Science: Natural Language
Processing, Automated Content Analysis, and Cloud Computing –
Part II (Guest: Abe Gong)

August 3 **NO CLASS** (Please See Online Content)

August 4 Nova and System Dynamics Modeling
(Guest: Eric Provins)

- August 5** Introduction to the ‘New’ Science of Networks and Network Models and Methods -- Part I
- Week 4**
- August 8** Network Models and Methods -- Part II
Introduction to Empirical Network Analysis in Pajek – Part I
- August 9** Introduction to Empirical Network Analysis in Pajek – Part II
Diffusion on Networks
HW Assignment #3 Distributed
- August 10** Advanced Network Analysis Topics:
Exponential Random Graph (p*) Models
and Community Detection – (Guest: Michael Bommarito)
- August 11** **NO CLASS** Open Lab – Finish Final Projects
(Please See Online Content)
HW Assignment #3 DUE

Some Useful Background Reading:

Mark E. J. Newman, Complex Systems: A Survey (Submitted to Amer. J. Physics) available at <http://www-personal.umich.edu/~mejn/papers/cssurvey.pdf>

Steve Levy, *The AI Revolution Is On*, Wired Magazine (January 2011) available at http://www.wired.com/magazine/2010/12/ff_ai_essay_airevolution/

Clive Thompson, *What Is I.B.M.’s Watson?*, NY Times Magazine (June 14, 2010) available at <http://www.nytimes.com/2010/06/20/magazine/20Computer-t.html>

Big Data: The Next Frontier for Innovation, Competition and Productivity (McKinsey Global Institute http://www.mckinsey.com/mgi/publications/big_data/index.asp)

The Data Deluge, The Economist (February 25, 2010) available at http://www.economist.com/node/15579717?story_id=15579717 (Read the Full Special Report)

Mark Whitehouse, *Economists' Grail: A Post-Crash Model*, The Wall Street Journal, (Nov. 30, 2010) available at <http://computationallegalstudies.com/2010/12/02/economists-grail-a-post-crash-model-via-wsj-com/>

Riders on a Swarm — Might Mimicking the Behavior of Ants, Bees & Birds Be the Key to Artificial Intelligence?, The Economist (August 17, 2010) available at <http://computationallegalstudies.com/2010/08/17/riders-on-a-swarm-might-mimicking-the-behavior-of-ants-bees-birds-be-the-key-to-artificial-intelligence/>

SOME USEFUL WEBSITES:

<http://infosthetics.com/>

(Information Aesthetics)

<http://www.memetracker.org/>

(Meme Tracker)

<http://www.visualcomplexity.com/vc/>

(Visual Complexity)

<http://well-formed.eigenfactor.org/>

(Well-Formed Eigenfactor)

<http://computationallegalstudies.com/>

(Computational Legal Studies)

<http://www.drewconway.com/zia/>

(Zero Intelligence Agents)

<http://flowingdata.com/>

(Flowing Data)

<http://www.stat.columbia.edu/%7Eegelman/blog/>

(Statistical Modeling, Causal Inference, and Social Science)

<http://www.iq.harvard.edu/blog/sss/>

(Social Science Statistics Blog (Harvard))

<http://www.iq.harvard.edu/blog/netgov/>

(Complexity and Social Networks Blog (Harvard))

<http://haris-aziz.blogspot.com/>

(Algorithms, Game Theory, Complexity, Combinatorics, etc.)

<http://www.santafe.edu/>

(Santa Fe Institute)

<http://languagelog.ldc.upenn.edu/nll/>

(Language Log)

<http://www.ted.com/>

(TED)

<http://paul.kedrosky.com/>

(Paul Kedrosky)

<http://behind-the-enemy-lines.blogspot.com/>

(Computer Scientist in a B – School)

<http://agtb.wordpress.com/>

(Algorithmic Game Theory/Economics)

<http://anyall.org/blog/>

(Artificial Intelligence and Social Science)

<http://www.cscs.umich.edu/~crshalizi/weblog/>

(Cosma Shalizi)