Computational Modeling in Cognitive and Social Science

Prof. Paulo Blikstein
General Information

School of Education EDUC 390X
Computer Science CS 424M
Winter 2011
Thursdays, 2:15 - 5:05pm
Room TBD

Professor

Paulo Blikstein
520 Galvez Mall - CERAS 232
Tel. (650) 736-0966
paulob@stanford.edu
Office hours: Tuesday, 2-4pm

Teaching assistant

TA: TBD
Office hours: TBD

Course Web Site & email lists

http://computationalmodeling.stanford.edu [not yet online]
Email list for course instructors: cmcss-faculty@lists.stanford.edu
Email list for all course members (students and faculty): cmcss2011@lists.stanford.edu
Course Abstract

Computational modeling and data-mining are dramatically changing the natural sciences, and more recently also the social and behavioral sciences. Traditional analytical techniques are insufficient to investigate complex dynamic social phenomena such as social networks, online gaming, diffusion of innovation, opinion dynamics, classroom behavior, and other complex adaptive systems. In this course, we will learn about how modeling, network theory, and basic data-mining can support research in cognitive, and social sciences, in particular about issues of learning, cognitive development, and educational policy.

Course Description

Over the last two decades, the use of computer modeling in the natural sciences has exploded, and recently this trend also reached the social and behavioral sciences. The exponential growth of data sources and data modalities has deeply transformed the work of these scientists, and traditional data collection and analysis techniques have become insufficient for a wide range of research questions – especially when they require the understanding of thousands of elements interacting in a multidimensional space. As social interaction and learning start to take place on several different social settings and media, these new methodological tools become a crucial part of the researcher’s toolbox.

For example, without computational tools, collecting, modeling, and analyzing data from social networks or online games would be unfeasible. Similarly, much of the contemporary research on mobile device usage, market dynamics, diffusion of innovation, human and business networks, consumer behavior, and opinion dynamics, to mention a few, would have been impossible without the use of computational methods and tools.

The goal of this course is to introduce students to these new methods. We will explore computational methods to support research in the learning, cognitive, and social sciences, with a particular focus on issues of learning, cognitive development, and educational policy. Our main tool will be multi-agent computer modeling. This modeling technique (also called agent-based modeling) helps explain systems where a large number of independent entities simultaneously interact with each other without central
control generating non-linear behaviors, adaptation, and learning, and resulting in emergent macroscopic patterns for the larger system.

Methods such as agent-based modeling are part of the emergent field of the “complexity sciences.” If you have heard of a social phenomenon that exhibits a “biological behavior,” an idea that is “spreading virally,” a stock market in panic, or just watched flocking birds or an ant nest, you already know a lot about complexity sciences.

The complexity sciences use novel investigation methods and frameworks to understand natural and social systems which cannot be described well with conventional models and formulas. Those systems are better understood starting from the simple behaviors of its micro elements (people, schools, businesses, or birds), instead of just examining the overall macroscopic system (a society, a classroom, a market, or a flock). Typical of complex phenomena is that the aggregate behaviors at the macroscopic level are not premeditated by any of the micro behaviors. For example, flocking birds do not intend to construct an arrow-shaped structure, and stock brokers do not want to crash the market. Rather, each element (or “agent”) follows its local rules (following another bird, for example, or trying to get the better price for a stock), and the overall pattern emerges from these multiple local behaviors.

Apart from this main methodological strand, we will examine several other topics, such as network theory, data-mining, text-mining, natural language processing, and sensor-based data collection. These techniques will allow students to bring real-world data into their models and get useful results, which could be immediately applied to their current research projects.

The course will have five main activities:

- **Lectures and class discussions**: here we will get in touch with the theoretical foundations in the complexity sciences, network theory, and relevant literature in the cognitive and learning sciences.
- **Tool review**: in class, we will review existing tools for computer modeling, data-mining, network analysis, and data visualization, and learn the cutting-edge of what’s “out there”.


Invited speakers from universities and industry will tell students about the state-of-the-art of the field.

Mini-tutorials: since this course is for both novices and experts, we will have mini-tutorials to get everyone up to speed on the software packages/programming languages we will use.

Project work: students will have an immersive, hands-on experience in creating a computer model, preferably by bringing their own data or research questions to the class.

Goals

Students will explore, construct, and analyze multi-agent models, using state of the art modeling tools, and discuss methodologies for rigorous replication, verification and validation of models. Additionally, we will introduce students to related modeling approaches, such as systems dynamics, and useful data-analysis techniques, such as text-mining and machine learning. In addition to the methods, the class will discuss questions general questions related to complex systems, for example, emergence and self-organization.

Students taking this course will develop basic mastery, and critical understanding, of agent-based modeling, and become familiar with fundamental tenets of complexity theory and its epistemological, cognitive, and methodological challenges. Participants in this course will become part of a small but growing nucleus of researchers beginning to appreciate the pertinence of complexity studies to understanding individual action in interpersonal milieus.

I can’t program. Should I enroll? (“YES!”)

No previous programming or technology background is assumed. Even if you have never programmed in your life, you can still enroll! This class is a good fit for people both with and without a technical background. In previous courses at Stanford and in other universities, students who had never written a single line of code created wonderful projects.
### Summary Schedule

<table>
<thead>
<tr>
<th>Class</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>Introduction, complex systems, emergence, agent-based modeling</td>
</tr>
<tr>
<td>Class 2</td>
<td>Agent-based modeling and system dynamics</td>
</tr>
<tr>
<td>Class 3</td>
<td>Networks and network analysis</td>
</tr>
<tr>
<td>Class 4</td>
<td>Agent-based modeling and participatory simulations</td>
</tr>
<tr>
<td>Class 5</td>
<td>Verification, replication, and validation, real-world data, sensors</td>
</tr>
<tr>
<td>Class 6</td>
<td>Text mining and machine learning</td>
</tr>
<tr>
<td>Class 7</td>
<td>Case studies: cognition and learning</td>
</tr>
<tr>
<td>Class 8</td>
<td>Case studies: social behavior and policy</td>
</tr>
<tr>
<td>Class 9</td>
<td>Ontological correspondence, advanced applications and new frontiers</td>
</tr>
<tr>
<td>Class 10</td>
<td>Final presentations</td>
</tr>
</tbody>
</table>

### Detailed schedule and readings

There will be no printed course reader. All materials will be online as PDFs. As a result, we will save one third of a tree over the ten weeks, and a lot of electrical energy. This will make us a bit less guilty for using so much energy to run lots of simulations for many hours.

**Class 1 - Intro to Computational Modeling in Cognitive and Social Science**

- Overview of the course & logistics
- What is a complex system? What is emergence? What is a model? Why are models important?
- Chaos, emergence, order, self-organization.
Modeling and simulation in the social and cognitive science: examples of current research.

Class activity: Introduction to NetLogo: creating your first agent-based model.

Readings:


Class 2 – Agent-based modeling, introduction to system dynamics

- Chaos, emergence, order, self-organization (II)
- Comparing agent-based models and system dynamics: when to use what?
- Examining your own research questions through computational modeling: how to think like a modeler, and why you should care – a new lingua franca for research.
- Stationary and movable agents; interactions between agents; agent topologies, artificial life, comparison with systems dynamics models.

Class activity: creating your first system dynamics model.

Readings:

Class 3 – Networks and Network analysis theory

- Networks, properties of networks, network metrics, agents in networks.
- The small-world problem
- **Class activity:** playing with social network data, adding social network data to your agent-based models.
- **Readings:**

Class 4 – Agent-Based Models (II) and Participatory simulations

- Using human agents in computational models
- Participatory simulations
- Combining human and virtual agents
- **Class activity:** creating your first participatory simulation (in groups).
- **Readings:**

**Class 5 – Verification, replication, and validation, real-world data, sensors.**

- Methods for verification and validation of models
- Model replication
- Systematic exploration of model parameter space, sensitivity analysis
- Sensors, real-world data feeds, Bifocal Modeling

**Lab activity (in the CERAS lab):** playing with real world sensors, data feeds, GoGo Boards and Arduinos.

**Readings:**


Class 6 – Text mining and machine learning

- Qualitative interview analysis and coding
- Text mining & parsing in Python
- Basic concepts in machine learning
- Class activity: text mining using Python
- Readings:

Classes 7 – Case studies: cognition and learning

- Case studies from psychology, and learning.
- Readings:
- Readings (one of the following):


**Classes 8 – Case studies: social behavior and policy**

- Case studies from policy, and social systems.

- **Readings:**


- **Readings** (one of the following):


---

**Class 9 - Ontological correspondence, advanced applications and new frontiers**

- Ontological correspondence, heterogeneous agents, interactions, interactions with the environment, bounded rationality, learning.

- Abstract models, middle range models, facsimile models.

- Integration of Machine Learning

- Evolutionary computation

- Geographical Information Systems (GIS)

- **Readings:**


**Class 10 – Final presentations**

**Assignments**

- **Due on Week 2**: Your first agent-based model
- **Due on Week 3**: Your first system dynamics model, and a 2-page comparison with the agent-based model
- **Due on Week 4**: Playing with social network data
- **Due on Week 5**: Participatory simulations (in groups)
- **Due on Week 6**: Bifocal model (in groups)
- **Due on Week 7**: Text mining
- **Due on Week 8**: Short paper on case studies (in groups)
- **Due on Week 9**: Software review (in groups)

**Final projects**

The final project is a big part of the course, and consists of a model, a set of results, and a paper. You can choose a theme close to your own research as a final project topic, and thus use data and literature that you already have. There are five main possibilities for final projects:

(a) Build models of school content for student learning, e.g. chemistry, material sciences, physics. These models would be mostly used as materials within a curriculum to teach a
particular content area.

(b) Build agent-based models that emulate observed behaviors of learners, whether individuals or cohorts, e.g., to model patterns of classroom collaboration, teamwork, individual cognition, self-efficacy, motivation, “peer-pressure,” etc.

(c) Create simulations that foreground and problematize central tenets of education-research theoretical models, e.g., pitting Piagetian and Vygotskian perspectives, or issues in other fields, such as social sciences, linguistics, political science, etc.

(d) Create simulations that foreground and problematize educational policy research questions, such as school choice, vouchers, achievement gap, school funding, etc.

(e) Develop creative combinations thereof.

The timeline for the final project is:

- **Due on Week 4**: One-page idea
- **Due on Week 5**: Revised one-page idea
- **Due on Week 6**: Individual meetings with instructors
- **Due on Week 7**: Final project proposal
- **Due on Week 10**: Presentation
- **Due on Week 11**: Final paper

**Software/hardware platforms used**

- **NetLogo**: a multi-agent version of Logo, this language is tuned for constructing models of dynamic systems. It is useful for creating models of ecological systems, chemical systems, economic trade, social behavior, etc.
- **GoGo Board** (interface for physical sensors).
- **Natural Language Toolkit** (text-mining, text-parsing)
Summary of Requirements

This course is designed to be somewhere between a class and a working group. We will work together to make sense of readings, and, for most of the class projects, you will be working in small groups. The requirements for everyone are:

✓ Set up your personal blog for the course.

✓ Keep up with the readings and participate in class, both in person and virtually. You will be expected to post a comment on each week’s reading by Wednesday at 12pm.

✓ Complete and present (mostly group) assignments.

✓ Review one modeling, analysis, or visualization software package and present your review in class.

✓ Design and implement your final project, and give a presentation during the last week of the course.

In addition, due to the group project nature of the class, you are asked to email us in advance if you cannot make a particular class meeting, and also responsible for communicating with your project mates.

Grading

All assignments will be graded as either incomplete, complete or outstanding. If an assignment is judged incomplete, you will have an opportunity to complete it or redo it the following week. You will also be assessed on your class participation both in class and on-line.

Overall, the final project will be worth 50% of the final grade, assignments will be worth 25%, and 25%
will be given for class participation. A specific rubric and the evaluation standards for the final project will be distributed during the course.

Additional foundational (& fun) literature (optional)


