Agent-Based Modeling Bootcamp for Health Researchers 2012

Dates: August 20-25, 2012
Location: Spinks Addition S320 Teaching Lab, University of Saskatchewan, Saskatoon, SK, Canada
Attendance: Limited to 35 participants
Fee: $900 per participant ($450 for students; *a limited number of student volunteer positions are also available – see the workshop information at the end of this document.*) Participants from previous years of the bootcamp who wish to attend the bootcamp again may do so at no charge, space permitting.
Website and Registration: http://tinyurl.com/HealthABMBootcamp2012
Contact: modelingtutorials@cs.usask.ca

The Agent-Based Modeling Bootcamp for Health Researchers is an intensive, hands-on tutorial that seeks to provide health science researchers with a systematic, practical and accessible introduction to agent-based modeling for health, using familiar language and concepts. The tutorial will cover a broad set of topics in agent-based modeling, offer optional material for those seeking additional depth of coverage, and emphasize the complementary insights when agent-based modeling is combined with traditional health sciences approaches and with other systems science methods. Although the course material will be provided in an intensive five-day "bootcamp", video of all sessions will be recorded. These videos, course presentations, example models, and other material will be made available for post-tutorial review. A post-course meeting will allow participants to brainstorm with the instructor and TAs concerning ideas for modeling projects.

Instructor

Dr. Nathaniel Osgood, University of Saskatchewan, Department of Computer Science & Associate Faculty, Department of Community Health and Epidemiology.

About the Instructor
Nathaniel Osgood is an Associate Professor in the Department of Computer Science and Associate Faculty in the Department of Community Health & Epidemiology, the School of Public Health and Division of Bioengineering at the University of Saskatchewan. His research – which has resulted in dozens of papers in peer-reviewed journals and conferences – is focused on providing tools to inform understanding of population health trends and health policy tradeoffs. Dr. Osgood has been applying Agent-Based modeling to understand human health and behaviour for over 20 years, including work conducted in the communicable, chronic and zoonotic disease areas. Dr. Osgood has additionally helped contribute novel new innovations to ease the Agent-Based modeling process, and has helped introduce novel techniques that hybridize Agent-based models with System Dynamics and Social Networks Analysis approaches, which combine simulation models with decision analytic approaches, and which leverage such models using data gathered from wireless real-time epidemiological monitoring systems. Dr. Osgood has taught tutorials on agent-based modeling and System Dynamics modeling internationally, and has served as a course instructor, guest lecturer, and plenary speaker on simulation modeling for health for the NIH-sponsored Institute for Systems Science and Health. The material presented in the bootcamp was also the focus of his recent MIT graduate-level course in Agent-Based modeling for health. Prior to joining the U of S faculty, Dr. Osgood worked for many years in a number of academic and industry positions, including on industry & academic projects applying modeling to tobacco and environmental epidemiology, health informatics, and multi-framework modeling for natural resource policy-making. He also served as instructor for several undergraduate- and graduate-level courses at MIT. Screencasts of some of Dr. Osgood's courses and talks are available at http://tinyurl.com/OsgoodModelingVideos

Teaching Assistants

The course will be staffed with 4 graduate-level teaching assistants, who will provide assistance both during the tutorial sessions and during the open times and post-tutorial brainstorming sessions. To better address questions of participants from a wide variety of backgrounds, the teaching assistants will be drawn from both health science and technical backgrounds.

Objectives and Audience

The tutorial is designed to give a detailed, hands-on introduction to the application of agent-based modeling to health decision making. The tutorial is specifically designed for health researchers who have no background in modeling, but who are interested in applications of modeling in the health area. While the tutorial will assume some familiarity with general terms from the health sciences, *no background in computer programming or mathematics is required*. The tutorial is also suitable for those researchers previously familiar with some form of modeling but who are seeking to develop skill in agent-based modeling conducted with the AnyLogic software package. Course material presented here has been refined from presentations over the past 18 months to students with a health science background. The large majority of the tutorial will be focused on Agent-Based Modeling in AnyLogic. There will also be a module that examines agent-level discrete event modeling (capturing, for example, patient flow through a hospital or an outpatient clinic). A final module will discuss the tradeoffs between multiple systems science approaches, and the ways in which such methods can be synergistically combined to yield greater insight. The tutorial will emphasize the practical benefits extending from the use of system science methods, and how such methods can be used in conjunction with traditional techniques.

The primary objectives of the tutorial are the following:

- To identify for participants the basic steps of the agent-based modeling process;
- To understand how agent-based models can assist with health science needs, including support for intervention planning, program evaluation, contact tracing, surveillance instrument design, vaccination schedule design;
- To provide participants a chance to discuss project ideas with experienced modelers;
- To provide hands-on knowledge on how to build basic models in AnyLogic;
- To provide participants with references of many existing models and identify resources where they can retrieve further information;
- To provide a roadmap for exploring and extending existing model-based research to gain better insights into chronic, infectious, and zoonotic disease dynamics;
- To provide an opportunity to understand how distinct modeling approaches typically pursued in isolation can be effectively combined within a single project or within the same model.
- To highlight limitations and challenges associated with agent-based models compared with those of other systems science methods;
- To convey an understanding of how agent-based models can be used together with traditional tools of the health sciences, including epidemiological and biostatistical methods, administrative data, systematic reviews and meta-analyses;

Teaching Style

The tutorial will employ a hands-on teaching style where participants will learn concepts while interactively exploring existing modeling software, model structure and model results. The instructor will use examples to highlight unique features of agent-based models, and the types of questions that on which they offer particular advantage, and limitations of such models. Registered participants will be given access to examples prior to the tutorial, and will be provided usable working copies of AnyLogic for the duration of the tutorial.

Additional times will be set aside for experimenting with the modeling package, pursuing exercises, raising points for clarification, and brainstorming with modelers and TAs concerning possible modeling projects.

Tutorial Contents

The tutorial will consist of a set of plenary sessions, accompanied by additional optional sessions that cover additional background material in greater depth. There will also be open sessions to allow students to ask additional questions on tutorial material. Each of these types of sessions is discussed below.

Hands-on Plenary Sessions

Topics covered in plenary sessions will include the following:

Focus	**Notes**
Orientation	What to expect & what not to expect from this course. Using the class time gaps well
Motivations for health applications of Systems Science methods and agent-based modeling techniques in particular	State & individual-based organization: Comparison & contrasts of agent-based lens with aggregate stock & flow modeling approaches; Understanding Emergence in Agent Based Models & stock and flow models; the effects of scale & scale invariance; Network structure & dynamics, Heterogeneity; Stochastics; Longitudinal data vs. cross-sectional time series; Adaptation; Learning; capacity to use of ABMs ABM's "synthetic ground truth" to evaluate statistical estimators and sampling methodology. Modification of an existing model.
Agent based modeling process	Brief overview of the agent based modeling (ABM) process, including the overview, Design concepts & Details (ODD) protocol, Phenomenology & Pattern-oriented modeling.
Scoping a model	Thinking in a structured way about the state of an agent (distinguishing static from dynamic), causal interactions (both structured and random) between the parties, higher level environmental influences, multi-level interactions. High level summary measures/statistics. Representing Inter- and intra-individual Feedbacks in Agent-Based models. Key question: What is relevant to answer the research questions being asked? Endogenous/Exogenous/Ignored classification. Interaction Scenarios. Breadth vs. depth. Opportunity costs. Incremental development. Key problem features pointing to agent-based or aggregate approaches.

AnyLogic ABM core concepts	Core concepts for creating runnable agent-based models; separation of models & scenarios ("experiments") Design vs. execution environment. The notion of "builds", and components with representation at both design & execution time.
Introduction to the AnyLogic user interface	AnyLogic declarative languages: Focusing on the "what", hiding (much of) the "how" AnyLogic user interface. Association of presentations with classes. Experiment objects. Data sets & output (charts & tables). Lack of trajectory files. Simulation time horizon. Access to Java code. Libraries & Library documentation. Modification-based & incremental development. Declaring agent populations. Real-time display: Dynamic properties.
Computational Architecture of Agent-Based Models	Key classes for all types of models. Primary classes: Main, Agent classes, Presentations, Experiments. Relationship of classes, run-time objects & design-time objects. Expressing agent heterogeneity with subclassing vs. parameterization.
Computational Components	Parameters & (State) variables. Functions. "Action" based code fragments (entering & leaving state, at transition, startup & destruction of Main & Person objects). Accessing the "Main" object via get_Main().
The importance of the language (Java) underlying declarative specification	The limits of AnyLogic's declarative specification ("How much Java code will I actually have to write?") Peeks under the hood: Imposed & elective (Java source files). Snippets of Java code: Where & why. (Startup, Destroy, etc.)
Specifying discrete intra-agent dynamics	Statecharts (Visual UML State Diagrams) & transitions. Branching, sources & sinks. Transitions: Rates, Timeouts, Conditions, Message-Based.
Discrete inter-agent dynamics	Messages: Sending & handling. Contagion of Pathogens & Ideas. Network neighbour selection options.
Continuous intra- and inter-agent dynamics	Stocks & flows in ABM: Why & how. CT & Viral Dynamics.
Agent characteristics & heterogeneity	Parameters, Variables (dynamic & static). Creating heterogeneous populations.
Events	Static events. Dynamic events & their uses. Event scheduling.
Environments	Networks & their relationships to populations. Topological (network) & geometric (regular & irregular spatial) context
Network environments	Networks: Structure and Dynamics • Network embedded agents, AnyLogic's built-in network types, and creating custom networks.

	- Network classes (Poisson Random, Ring lattice, Small World networks, Scale Free networks) - Mathematics of scale free networks - Accounting for network structure in aggregate models Impact on network dynamics Impact of network fitting and aggregation on results.
Structural dynamics	Implementing dynamically varying populations with birth, death, migration & dynamic networks.
Meta-population modeling	Networked or geographic positioning of (possibly sublocated) populations
The Embedded Environment: Focus on Regular Spatial Embedding	**Regular spatial embedding (Cellular Automata)**: Neighbourhoods & inter-agent communication. Color. Boundary Conditions. Alternative tessellation schemes. Discrete & continuous geometries.
Representing Spatial Heterogeneity & Irregular geometries	Elephants, CWD. **Irregular spatial embedding**: "Network-Based Modeling" in AnyLogic. Capturing spatial locations for processing. Dealing with resources & resource competition/mutual exclusion. Spatial movement between discrete locations. Defining paths & resources. Static & dynamic resources.
Agent mobility & movement	Regular vs. irregular, discrete vs. continuous
Customizing model appearance before & during execution	Creation of custom user-interface elements to - Present data via graphs & tables - Dictate assumptions via user input "widgets" (sliders, checkboxes, text fields, etc.) - Initial assumptions (e.g. regarding parameter values) Live changes to assumptions
Representing Interventions	Representing diverse static & dynamic interventions for simulation experiments
Analyzing & reporting data	Reporting data for analysis - Built-in Statistics - Custom statistics & stratification - Event use
Inputting and Outputting data	Moving data to/from files, databases, Excel, R, etc.) (ad-hoc and pre-planned)
Sensitivity Analyses	Parameter, structural, automated & manual, one way & multi-way Parameter uncertainty Model uncertainty Stochastics: Uncertainty in evolution over time 2D Histograms: Understanding the dynamic response of the model under uncertainty - Empirical fractiles - Plots of per-realization behaviour

Model parameter estimation & Calibration	Diverse sources of parameters. Challenges of network data. Opportunities for employing novel data collection mechanisms

Calibration	Adjusting assumptions to best match historical data & behavioural reference modes
	• Basic concepts of calibration
	  ○ Parameter space
	  ○ Objective function & the implied error model & MLE
	  ○ Weighing multiple matches
	  ○ Optimization algorithm
	• Calibration in presence of stochastics
	• Assessing convergence & handling pathologies
	• Alternatives to calibration: A glimpse of Bayesian approaches

Discrete Event Modeling & Simulation	Simulating the flow and interaction of agents and resources in facilities
Model computational resource footprint	Common performance vulnerabilities Measuring performance using profilers Effective strategy for improving model performance
Debugging in AnyLogic	Via Eclipse & AnyLogic Resesarch/Professional Debugger. Distinguishing the failure from the fault. Tracing. Isolation. Binary search. Common defects. Remote debugging via Eclipse. Basic principles & practices of debugging Tips and techniques for effective debugging Debugging build problems Mechanisms for logging & tracing Multiple debugging methodologies Use of the AnyLogic debugger and an external debugger
Understanding & combining multiple systems science modeling types: Tradeoffs & synergies	Complementary nature of ABM, classic stock & flow aggregate SD models, and social network analysis. Terminology (Individual & Agent-Based models), Emergence. Levels of description. The myth of "bottom up" description. Independent representation & interactions of comorbidity. Intervention granularity, expressiveness, performance scaling, expressing heterogeneity (both static & dynamic), capturing history information, transfer effects, capturing detailed dynamics (e.g. spatial & network spread, memoryful processes), calibration & parameterization, ease of specification.
Multiple systems science methods	Complementary strengths and weaknesses Building AnyLogic models incorporating multiple modeling approaches. Using several sorts of models together as synergistic lenses. Immuno-epidemiological modeling for influenza & Chlamydia as an example.

	Presentation of effective hybrid modeling strategies
Confronting Dynamic Decision Problems by Combining Simulation & Decision Analysis	Linkages of simulation models with other modeling techniques such as multiattribute decision theory. Theoretical and practical basis for tying together decision trees & simulation models.
Best Practices: Processes	Peer reviews, Version control (Versioning, Check-in (commit)/check out (locking)), Incremental delivery, YAGNI Principle, Continuous Integration & smoke testing, Refactoring. Unit checking. Brief exposure to principles of testing (test drivers, harnesses, mocks/fakes, unit tests), Risk Management. Continuous process improvement.
Best Practices: Technical	Eliminating manifest constants. Assertions. ABM as realization of mathematical process; documenting mathematical description & Checking adherence of code to mathematical specification. Reducing software engineering complexity via modularity & model transparency. Trajectory Recording, abstraction using methods & classes, separating interface specification & implementation. Commenting. Using intention-revealing naming conventions. Individual versioning. Mocks & Fakes. Incremental delivery. Change & configuration management. Exceptions vs. return codes. Clean coding suggestions. Test-first coding.

Optional Additional Sessions

Additional background sections will be offered for practitioners interested in seeking broader background or deeper understanding of some modeling practices and supporting technologies. The particular set of sessions to be offered will depend on the balance student interest expressed in the candidate sessions. While we cannot guarantee that all topics listed below will be covered during the bootcamp, dedicated sessions will cover material of interest to participants in a prioritized fashion, and videos will be made available for those topics which could not be included within the bootcamp.

- Motivation for systems science methods in general, and agent-based modeling in particular
- Best practices for model building
- Implementing dynamically varying populations with birth, death, migration.
- Model performance
  - Common performance vulnerabilities
  - Measuring performance using profilers
  - Effective strategy for improving model performance
- Helpful bits of Java for AnyLogic users who lack programming exposure but are interested in securing added flexibility when creating models, and deeper understanding of model plumbing.
  - Basic introduction to Java
  - Methods & functions
  - Classes & objects

    - Events
    - Types
    - Expressions
    - Statements
    - Using code in external Java libraries
    - Capturing hierarchies of related agents or resources: Subtyping and subclassing
- Simulating the flow and interaction of agents and resources in facilities: Discrete event modeling and visualization in AnyLogic

Deliverables

By the conclusion of the tutorial, each participant will be provided with the following:

- Slides
- Exercises
- Example models
- Models built by the participant
- Purpose-built tools to assist in AnyLogic model debugging & database output
- References to more extensive learning opportunities (full courses, articles, books and model libraries)
- Information on how to access videos of course contents and related online material
- Optional free, custom-built software to help do "bookkeeping" on model versions, model inputs, and model outputs, associated files, and related commentary.

Optional post-tutorial modeling brainstorming session

The instructor and teaching assistants will be available to participants for a day of modeling brainstorming. This session will provide participants a chance to integrate and assimilate teachings from the tutorial session, and discuss application to their area of interest.

Student volunteers are being sought for helping out with the bootcamp; such volunteers receive free admission to the bootcamp. Please contact the address above for more information.