

## SPE316 Final Project Guidelines

### Fall 2022

#### Pick an Approach

You will either substantively extend an existing ABM model from any political, economic, social or other literature or create a new ABM de novo based upon theoretical grounds or policy contexts. Either approach will necessitate you to work with a CAS framework.

If you extend an existing ABM model, you must add at least one significant module, ie add a network component, a behavioral game component, increasing agent attributes and resultant behavior, or a macro dynamic component. Sub models and components must be architected in a CAS framework. If you build an ABM de novo, please use existing theory, policy challenge or opportunity that has not been formalized before and architect an ABM. Either way, the NetLogo library and online community is a great place to start!

Final project contents and specifications per below:

- 1. Introduction**
- 2. Research Question**
  - a. Start with a clear question, problem, or hypothesis
  - b. Why might an ABM/CAS approach be better?
- 3. Diagram Model Architecture**
  - a. Translate your natural language theory and narrative to logical diagrams, often called a concept of operations (conops) diagram
- 4. ODD** (fill out the usual framework per below)
- 5. Coding**
  - a. Define high level pseudo code first, then start coding details and submodules according to your architecture diagram and conops; provide code in the appendix.
- 6. The Model User Interface**
  - a. Describe what will your model look like, such as the world, monitors, plots, input and parameter sliders, simulation controls, etc. and why?
- 7. Baseline, Simulation and Scenarios**
  - a. Write out case narrative, ICs, parameters, etc and resulting outputs, compare and contrast cases discursively, quantitatively and qualitatively
  - b. What CAS behavior(s) result?
- 8. Data Analysis on Results**
  - a. Behavior Space for results robustness
  - b. Regression of quasi global sensitivity analysis
  - c. Equilibrium behavior via phase portraits
  - d. Heat Maps of policy lever sensitivity
- 9. Strategy**
  - a. Behavior search for strategic insights how to optimize/minimize/maximize state/global variables or fitness function of your model to answer your policy question(s) and hypotheses.
- 10. Conclusions**
  - a. What worked, what didn't work, what can you
  - b. Summary of ABM insights from your model
  - c. Why they are important for decision/policy makers
- 11. References, Appendix, etc.**

## ODD

### 1. Purpose

*Question:* What is the purpose of the model? **Answer:** ...

*Explanation:* Every model has to start from a clear question, problem, or hypothesis. Therefore, ODD starts with a concise summary of the overall objective(s) for which the model was developed. Do not describe anything about how the model works here, only what it is to be used for. We encourage authors to use this paragraph independently of any presentation of the purpose in the introduction of their article, since the ODD protocol should be complete and understandable by itself and not only in connection with the whole publication (as it is also the case for figures, tables and their legends). If one of the purposes of a model is to expand from basic principles to richer representation of real-world scenarios, this should be stated explicitly.

### 2. Entities, state variables, and scales

*Questions:* What kinds of entities are in the model? By what state variables, or attributes, are these entities characterized? What are the temporal and spatial resolutions and extents of the model?

**Answer:** ...

*Explanation:* An entity is a distinct or separate object or actor that behaves as a unit and may interact with other entities or be affected by external environmental factors. Its current state is characterized by its state variables or attributes. A state variable or attribute is a variable that distinguishes an entity from other entities of the same type or category, or traces how the entity changes over time. Examples are weight, sex, age, hormone level, social rank, spatial coordinates or which grid cell the entity is in, model parameters characterizing different types of agents (e.g., species), and behavioral strategies. The entities of an ABM are thus characterized by a set, or vector (Chambers, 1993; Huse et al., 2002), of attributes, which can contain both numerical variables and references to behavioral strategies.

One way to define entities and state variables is the following: if you want (as modelers often do) to stop the model and save it in its current state, so it can be re-started later in exactly the same state, what kinds of information must you save? If state variables have units, they should be provided. State variables can change in the course of time (e.g. weight) or remain constant (e.g. sex, species-specific parameters, location of a non-mobile entity). State variables should be low level or elementary in the sense that they cannot be calculated from other state variables. For example, if farmers are represented by grid cells which have certain spatial coordinates, the distance of a farmer to a certain service centre would not be a state variable because it can be calculated from the farmer's and service centre's positions.

Most ABMs include the following types of entities:

– *Agents/individuals.* A model can have different types of agents; for example, wolves and sheep, and even different sub-types within the same type, for example different functional types of plants or different life stages of animals. Examples of types of agents include the following: organisms, humans, or institutions. Example state variables include: identity number (i.e., even if all other state variables would be the same, the agent would still maintain a unique identity), age, sex, location (which may just be the grid cell it occupies instead of coordinates), size, weight, energy reserves, signals of fitness, type of land use, political opinion, cell type, species-specific parameters describing, for example, growth rate and maximum age, memory (e.g., list of friends or quality of sites visited the previous 20 time steps), behavioral strategy, etc.

– *Spatial units (e.g., grid cells).* Example state variables include the following: location, a list of agents in the cell, and descriptors of environmental conditions (elevation, vegetation cover, soil type, etc.) represented by the cell. In some ABMs, grid cells are used to represent agents: the state and behavior

of trees, businesses, etc., that can be modeled as characteristics of a cell. Some overlap of roles can occur. For example, a grid cell may be an entity with its own variables (e.g., soil moisture content, soil nutrient concentration, etc., for a terrestrial cell), but may also function as a location, and hence an attribute, of an organism.

– *Environment.* While spatial units often represent environmental conditions that vary over space, this entity refers to the overall environment, or forces that drive the behavior and dynamics of all agents or grid cells. Examples of environmental variables are temperature, rainfall, market price and demand, fishing pressure, and tax regulations.

– *Collectives.* Groups of agents can have their own behaviors, so that it can make sense to distinguish them as entities; for example, social groups of animals, households of human agents, or organs consisting of cells. A collective is usually characterized by the list of its agents, and by specific actions that are only performed by the collective, not by their con-stitutive entities.

In describing spatial and temporal scales and extents (the amount of space and time represented in a simulation), it is important to specify what the model's units represent in reality. For example: "One time step represents one year and simulations were run for 100 years. One grid cell represents 1 ha and the model landscape comprised 1,000 x 1,000 ha; i.e., 10,000 square kilometers".

### 3. Process overview and scheduling

Questions: Who (i.e., what entity) does what, and in what order? When are state variables updated?

How is time modeled, as discrete steps or as a continuum over which both continuous processes and discrete events can occur? Except for very simple schedules, one should use pseudo-code to describe the schedule in every detail, so that the model can be re-implemented from this code. Ideally, the pseudo-code corresponds fully to the actual code used in the program implementing the ABM.

**Answer: ...**

*Explanation:* The "does what?" in the first question refers to the model's processes. In this ODD element only the self-explanatory names of the model's processes should be listed: 'update habitat', 'move', 'grow', 'buy', 'update plots', etc. These names are then the titles of the submodels that are described in the last ODD element, 'Submodels'. Processes are performed either by one of the model's entities (for example: 'move'), or by a higher-level controller that does things such as updating plots or writing output to files. To handle such higher-level processes, ABM software platforms like Swarm (Minar et al., 1996) and NetLogo (Wilensky, 1999) include the concept of the 'Model', or 'Observer', itself; that is, a controller object that performs such processes.

By "in what order?" we refer to both the order in which the different processes are executed and the order in which a process is performed by a set of agents. For example, feeding may be a process executed by all the animal agents in a model, but we must also specify the order in which the individual animals feed; that is, whether they feed in random order, or fixed order, or size-sorted order. Differences in such ordering can have a very large effect on model outputs (Bigbee et al., 2006; Caron-Lormier et al., 2008).

The question of when variables are updated includes the question of whether a state variable is immediately assigned a new value as soon as that value is calculated by a process (asynchronous updating), or whether the new value is stored until all agents have executed the process, and then all are updated at once (synchronous updating). Most ABMs represent time simply by using time steps: assuming that time moves forward in chunks. But time can be represented in other ways (Grimm and Railsback, 2005, Chapter 5). Defining a model's schedule includes stating how time is modeled, if it is not clear from the 'Entities, State Variables, and Scales' element.

#### 4. Design concepts

*Questions: There are eleven design concepts. Most of these were discussed extensively by Railsback (2001) and Grimm and Railsback (2005; Chapter. 5), and are summarized here via the following questions:*

*Basic principles. Which general concepts, theories, hypotheses, or modeling approaches are underlying the model's design? Explain the relationship between these basic principles, the complexity expanded in this model, and the purpose of the study. How were they taken into account? Are they used at the level of submodels (e.g., decisions on land use, or foraging theory), or is their scope the system level (e.g., intermediate disturbance hypotheses)? Will the model provide insights about the basic principles themselves, i.e. their scope, their usefulness in real-world scenarios, validation, or modification (Grimm, 1999)? Does the model use new, or previously developed, theory for agent traits from which system dynamics emerge (e.g., 'individual-based theory' as described by Grimm and Railsback [2005; Grimm et al., 2005])? **Answer: ...***

*Emergence. What key results or outputs of the model are modeled as emerging from the adaptive traits, or behaviors, of individuals? In other words, what model results are expected to vary in complex and perhaps unpredictable ways when particular characteristics of individuals or their environment change? Are there other results that are more tightly imposed by model rules and hence less dependent on what individuals do, and hence 'built in' rather than emergent results? **Answer: ...***

*Adaptation. What adaptive traits do the individuals have? What rules do they have for making decisions or changing behavior in response to changes in themselves or their environment? Do these traits explicitly seek to increase some measure of individual success regarding its objectives (e.g., "move to the cell providing fastest growth rate", where growth is assumed to be an indicator of success; see the next concept)? Or do they instead simply cause individuals to reproduce observed behaviors (e.g., "go uphill 70% of the time") that are implicitly assumed to indirectly convey success or fitness? **Answer: ...***

*Objectives. If adaptive traits explicitly act to increase some measure of the individual's success at meeting some objective, what exactly is that objective and how is it measured? When individuals make decisions by ranking alternatives, what criteria do they use? Some synonyms for 'objectives' are 'fitness' for organisms assumed to have adaptive traits evolved to provide reproductive success, 'utility' for economic reward in social models or simply 'success criteria'. (Note that the objective of such agents as members of a team, social insects, organs—e.g., leaves—of an organism, or cells in a tissue, may not refer to themselves but to the team, colony or organism of which they are a part.) **Answer: ...***

*Learning. Many individuals or agents (but also organizations and institutions) change their adaptive traits over time as a consequence of their experience? If so, how? **Answer: ...***

*Prediction. Prediction is fundamental to successful decision-making; if an agent's adaptive traits or learning procedures are based on estimating future consequences of decisions, how do agents predict the future conditions (either environmental or internal) they will experience? If appropriate, what internal models are agents assumed to use to estimate future conditions or consequences of their decisions? What tacit or hidden predictions are implied in these internal model assumptions? **Answer: ...***

*Sensing. What internal and environmental state variables are individuals assumed to sense and consider in their decisions? What state variables of which other individuals and entities can an individual perceive; for example, signals that another individual may intentionally or unintentionally send? Sensing is often assumed to be local, but can happen through networks or can even be assumed to be global (e.g., a forager on one site sensing the resource levels of all other sites it could move to). If agents sense each other through social networks, is the structure of the network imposed or*

*emergent? Are the mechanisms by which agents obtain information modeled explicitly, or are individuals simply assumed to know these variables? **Answer: ...***

*Interaction. What kinds of interactions among agents are assumed? Are there direct interactions in which individuals encounter and affect others, or are interactions indirect, e.g., via competition for a mediating resource? If the interactions involve communication, how are such communications represented? **Answer: ...***

*Stochasticity. What processes are modeled by assuming they are random or partly random? Is stochasticity used, for example, to reproduce variability in processes for which it is unimportant to model the actual causes of the variability? Is it used to cause model events or behaviors to occur with a specified frequency? **Answer: ...***

*Collectives. Do the individuals form or belong to aggregations that affect, and are affected by, the individuals? Such collectives can be an important intermediate level of organization in an ABM; examples include social groups, fish schools and bird flocks, and human networks and organizations. How are collectives represented? Is a particular collective an emergent property of the individuals, such as a flock of birds that assembles as a result of individual behaviors, or is the collective simply a definition by the modeler, such as the set of individuals with certain properties, defined as a separate kind of entity with its own state variables and traits? **Answer: ...***

*Observation. What data are collected from the ABM for testing, understanding, and analyzing it, and how and when are they collected? Are all output data freely used, or are only certain data sampled and used, to imitate what can be observed in an empirical study ("Virtual Ecologist" approach; Zurell et al., 2010)? **Answer: ...***

*Explanation: The 'Design concepts' element of the ODD protocol does not describe the model per se; i.e., it is not needed to replicate a model. However, these design concepts tend to be characteristic of ABMs, though certainly not exclusively. They may also be crucial to interpreting the output of a model, and they are not described well via traditional model description techniques such as equations and flow charts. Therefore, they are included in ODD as a kind of checklist to make sure that important model design decisions are made consciously and that readers are aware of these decisions (Railsback, 2001; Grimm and Railsback, 2005). For example, almost all ABMs include some kinds of adaptive traits, but if these traits do not use an explicit objective measure the 'Objectives' and perhaps 'Prediction' concepts are not relevant (though many ABMs include hidden or implicit predictions). Also, many ABMs do not include learning or collectives. Unused concepts can be omitted in the ODD description.*

*There might be important concepts underlying the design of an ABM that are not included in the ODD protocol. If authors feel that it is important to understand a certain new concept to understand the design of their model, they should give it a short name, clearly announce it as a design concept not included in the ODD protocol, and present it at the end of the Design concepts element.*

## **5. Initialization**

### **Detailed Dashboard**

*Questions: What information do you need to convey in your dashboard/user interface? Why do you need that information to answer your research question(s)? What will your world be? What monitors and plots do you need? What input parameters/sliders and other simulation controls will you need? What is the initial state of the model world, i.e., at time  $t = 0$  of a simulation run? In detail, how many entities of what type are there initially, and what are the exact values of their state variables (or how were they set stochastically)? Is initialization always the same, or is it allowed to vary among simulations? Are the initial values chosen arbitrarily or based on data? References to those data should be provided. **Answer ...***

*Explanation:* The dashboard is the user interface to your model. Just like any software, it must be designed to achieve its users goals, answering questions and providing insights. Model results cannot be accurately replicated unless the initial conditions are known. Different models, and different analyses using the same model, can of course depend quite differently on initial conditions. Sometimes the purpose of a model is to analyze consequences of its initial state, and other times modelers try hard to minimize the effect of initial conditions on results.

## **6. Input data**

*Question:* Does the model use input from external sources such as data files or other models to represent processes that change over time? **Answer:** ...

*Explanation:* In models of real systems, dynamics are often driven in part by a time series of environmental variables, sometimes called external forcings; for example annual rainfall in semi-arid savannas (Jeltsch et al., 1996). “Driven” means that one or more state variables or processes are affected by how these environmental variables change over time, but these environmental variables are not themselves affected by the internal variables of the model. For example, rainfall may affect the soil moisture variable of grid cells and, therefore, how the recruitment and growth of trees change. Often it makes sense to use observed time series of environmental variables so that their statistical qualities (mean, variability, temporal autocorrelation, etc.) are realistic. Alternatively, external models can be used to generate input, e.g. a rainfall time series (Eisinger and Wiegand, 2008). Obviously, to replicate an ABM, any such input has to be specified and the data or models provided, if possible. (Publication of input data for some social simulations can be constrained by confidentiality considerations.) If a model does not use external data, this element should nevertheless be included, using the statement: “The model does not use input data to represent time-varying processes.” Note that ‘Input data’ does *not* refer to parameter values or initial values of state variables.

## **7. Submodels**

*Questions:* What, in detail, are the submodels that represent the processes listed in ‘Process overview and scheduling’? What are the model parameters, their dimensions, and reference values? How were submodels designed or chosen, and how were they parameterized and then tested? **Answer:** ...

*Explanation:* The submodels are presented in detail and completely. The factual description of the submodel, i.e., equation(s) and algorithms, should come first and be clearly separated from additional information. From what previous model this submodel was taken or whether a new submodel was formulated, and why, can be explained. If parameterization is not discussed outside the ODD description, it can be included here. The parameter definitions, units, and values used (if relevant) should be presented in tables.

Any description of an ABM and its submodels will seem *ad hoc* and lack credibility if there is no justification for why and how formulations were chosen or how new formulations were designed and tested. Because agent-based modeling is new and lacks a firm foundation of theory and established methods, we expect ODD descriptions to include appropriate levels of explanation and justification for the design decisions they illustrate, though this should not interfere with the primary aim of giving a concise and readable account of the model. Justification can be very brief in the Overview and Design concepts sections, but the complete description of submodels is likely to include references to relevant literature, as well as independent implementation, testing, calibration, and analysis of submodels. ODD-based model descriptions consist of the seven elements described above; however, in most cases it will be necessary to have a simulation experiments or model analysis section following the model description.