NEW MATHEMATICS AND SOFTWARE FOR AGENT-BASED MODELS

JOHN C. BAEZ AND NATHANIEL D. OSGOOD

1. INTRODUCTION

An agent-based model (ABM) is a computational construct characterizing the dynamics of one or more open populations of individual agents, each associated with parameters, state, rules and actions, and interacting with each other and with an environment in which they are embedded [7]. Such agents are often spatially situated and mobile, linked by one or more networks, and hierarchically nested. While ABMs increasingly inform health policy decision-making, our work within this arena, and that of many other researchers, clearly demonstrates that extant tools and software frameworks impose many barriers to effective ABM construction and use. In a domain where interdisciplinary teamwork is of central importance, current ABMs lack precise and publishable specifications, are commonly opaque to non-technical stakeholders, and are often difficult to navigate even for technical team members. They also provide poor support for modifying and reusing models, and often impaired performance. All this and more raises numerous challenges to the collaboration within and between diverse teams critical for impactful modeling.

A new generation of modeling methods grounded in applied category theory offers the potential to overcome many of these barriers. Our recent work has demonstrated the promise of these methods as part of the AlgebraicJulia open-source software project [3, 4, 6]. In this application, we seek to extend our work to ABMs. We plan to develop models that offer visual transparency for stakeholders, modularity and reuse, easy optimization, alignment between types of heterogeneity, formal migration in support of model evolution, and a syntax/semantics separation supporting flexibility in model operation across interdisciplinary teams.

These new methods embody the following principles:

(1) **Modularity:** Models of specific subsystems may be constructed individually by different domain experts, then coupled together using appropriate principles, supporting ongoing collaboration by the parties.
(2) **Hierarchy:** Multiscale models, such as of spatially distributed human/animal populations, can be constructed in a multi-scale way that mirrors the structure of the system.
(3) **Robustness:** Models are well-defined mathematical structures, which may be specified, optimized, evolved, and visualized using high-level operations. Simulation code is generated from the specification, reducing programming effort and errors, and accelerating the feedback cycle with stakeholders.
(4) **Clarity, accessibility, and transparency:** Rigorous yet intuitive diagrammatic languages, such as wiring diagrams, support clear communication of model structure. This allows input and critique from all members of a modelling team, regardless of mathematical experience.

2. Objectives and Specific Aims

The overall objective of our proposed work is to reduce barriers to team-based and richly data-informed health dynamic modeling by using categorical methods to realize visually and mathematically transparent, modular, flexible, readily evolvable, efficient and reproducible agent-based modelling. We propose five lines of work. (1)–(3) will be undertaken at the ICMS facility during the event, while (4)–(5) will be conducted remotely and continue after the event:

(1) Developing a mathematical framework for ABMs.
(2) Implementating this framework in AlgebraicJulia.
(3) Creating five ABMs using our framework for teaching purposes, and beginning to develop a full-scale production-quality ABM within this framework. Here we shall draw on Osgood's training library of over 100 diverse health science models.
(4) Building a real-time collaborative web-based interface to let teams of users simultaneously and collaboratively construct, view, modify and interact with ABMs.
(5) Evaluating, refining and disseminating the above via training bootcamps, incubators, hackathons, Osgood's online video series [9] and sponsored projects.

While development will concentrate on these five lines of work in successive periods of the schedule, we anticipate contributors from all spheres of work to be continually present through all 6 weeks of the proposed event.
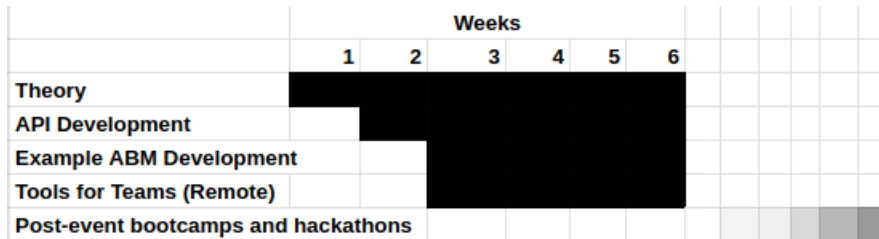

FIGURE 1. Event timeline, showing successive staging of work

3. The Team

The team is:

(1) Nathaniel D. Osgood: computer scientist and public health modeler with dozens of peer-reviewed publications on ABMs; provincial director of COVID-19 modeling in Canada; supervised creation of ABMs used for COVID-19 decision-making across Canada and Australia. Contributed to development of StockFlow.jl using AlgebraicJulia as a framework for "stock and flow models" in epidemiology [3, 4, 8, 10].
(2) John C. Baez: mathematical physicist and applied category theorist; helped develop the "decorated cospan" framework widely used in AlgebraicJulia [2], and helped apply this to stock and flow models in epidemiology [3, 4].
(3) William Waites: computer scientist and member of the Digital Health and Wellness Research Group at Strathclyde; developed compositional and rule-based models of disease transmission [11, 12].
(4) Xiaoyan Li: Doctoral student in Computer Science with extensive peer-reviewed contributions in health applications of dynamic modeling; lead

developer of the StockFlow.jl package using AlgebraicJulia for stock and flow health models [3, 4, 8, 10].

(5) Sean Wu: Epidemiology (PhD) and public health (MPH) modeler; senior scientist at Merck; creator of an ABM package in R and a prototype ABM implementation in AlgebraicJulia [13].

(6) Evan Patterson: applied category theorist specializing in scientific computing, software systems and data science; helped design the AlgebraicJulia framework [1] and software using this framework for both Petri net [6] and stock and flow models [3] of epidemiology.

(7) Kristopher Brown: research software engineer; led software development for ABMs and rewrite rules in AlgebraicJulia [5, 13].

(At most 6 team members will be on site at any time. Patterson and Brown will share a spot, and the Topos Institute will cover any extra travel expenses for them.)

4. Benefits

Based on our above-proposed work in building, evaluating and refining model libraries and updating our course exercises for use with the systems being developed here, we will work to secure widespread dissemination of our ACT-based modeling tools and supporting material via our bootcamps, hackathons and online materials. The applicants and attendees have already led a successful and well-subscribed bootcamp on infectious disease modeling using ACT. Applicant Osgood has a demonstrated history of leading of dozens of previous ABM bootcamps across Australia, Canada, the US and beyond, and has > 3500 subscribers and $> 700,000$ views on his teaching channel [9]. Thus, we are confident that successful delivery of this work will support effective opportunities for feedback and evaluation, uptake and dissemination of the categorical methods described here. The work will further offer significant insights for researchers in computer science (via conference publications, AlgebraicJulia contributions, language design/engineering), computational epidemiology, and ripple-through benefits for public health (PH) decision-makers and population health.

Methodological contributions will come in the form of an open-source project to support health agent-based modeling. For computational epidemiologists, this work is likely to illuminate little-recognized avenues to advance their field, and offer support of more transparent, shareable, reusable, versatile and less error-prone models. For health scientists, we anticipate our work will expand data availability and reduce cognitive and computational burdens and collaborative hurdles associated with modeling, leading to more rapid insight regarding intervention tradeoffs and indirectly improving PH decision-making via our extensive modeling application collaborations. While the effects are more indirect, we expect that decision-makers in PH (with whom we maintain extensive collaborations) will benefit from this research via earlier warnings of emergent PH concerns, more reliable and transparent models, improved ability to explore scenarios, greater confidence regarding intervention tradeoffs, and better understanding of drivers underlying observed PH trends, and enhanced decision making.

5. Methodological Details

Adapting the approach employed to create software for stock and flow [3, 4] and Petri net [6] models, we will encode the syntax of ABMs using copresheaf categories, which are well supported in AlgebraicJulia [1]. In standard fashion, a functor into Set will serve to encode a particular ABM. Various functors will then allow various rewriting-based semantics to execute or analyze this ABM. Semantics that

we seek to support (in decreasing order of priority) include single-realization simulation, ensemble simulation, sensitivity analyses, calibration, parameter sampling via Approximate Bayesian Computation, and execution using Sequential Monte Carlo methods. Drawing on insights from our previous work, we plan to use Julia macros to create a declarative domain-specific language supporting the specification of ABMs. Our work will draw on the success of a recent effort by our collaborators [13] to implement a simple version of individual-based models using the AlgebraicJulia package for rewrite rules [5], but will move far beyond what was accomplished in that work by supporting both higher-level ABM constructs critical for model transparency and a repertoire of mechanisms for agent-based modeling. Constructs that we aim to support include—but are not limited to—statecharts, mode-dependent wiring diagrams, and ODEs for characterizing and evolving state, populations, queries over them, and scenario definitions. We will additionally seek to support (in order of decreasing priority) hierarchical nesting of agents, network and spatial embedding, and between-agent interaction mechanisms, such as message exchange. If there is insufficient time to adequately support all mechanisms, we shall pursue them according to priority and ease of characterization within the developed mathematics and API. During the event, we plan to implement these methods in five ABMs for pertussis, COVID-19, gestational and type 2 diabetes, opioid-related harms, and antimicrobial resistance.

References

[1] AlgebraicJulia: Bringing compositionality to technical computing. Available at https://www.algebraicjulia.org.

[2] J. C. Baez, K. Courser and C. Vasilakopoulou, Structured versus decorated cospans, *Compositionality* **4**, 3 (2022)

[3] J. C. Baez, X. Li, S. Libkind, N. D. Osgood and E. Patterson, Compositional modeling with stock and flow diagrams, to appear in *Proceedings of Applied Category Theory 2022.*

[4] J. C. Baez, X. Li, S. Libkind, N. D. Osgood and E. Redekopp, A categorical framework for modeling with stock and flow diagrams, to appear in *Mathematics for Public Health, Springer*, Berlin.

[5] K. Brown, E. Patterson, T. Hanks, and J. Fairbanks, Computational category-theoretic rewriting, in *Graph Transformation: 15th International Conference, ICGT 2022*, 2022, pp. 155–172.

[6] S. Libkind, A. Baas, M. Halter, E. Patterson and J. P. Fairbanks, An algebraic framework for structured epidemic modelling, *Philosophical Transactions of the Royal Society A* **380** (2022), 20210309.

[7] G. W. McDonald and N. D. Osgood, Agent-based modeling and its tradeoffs: an introduction and examples, arXiv:2304.08497.

[8] N. Meadows, X. Li and N. D. Osgood, Hierarchical and upstream-downstream composition of stock and flow models, arXiv:2305.02136.

[9] N. Osgood, Teaching videos.

[10] Stockflow.jl. Available at https://github.com/AlgebraicJulia/StockFlow.jl.

[11] W. Waites, M. Cavaliere, D. Manheim, J. Panovska-Griffiths, and V. Danos, Rule-based epidemic models, *Journal of Theoretical Biology* **530** (2021), 110851.

[12] W. Waites, M. Cavaliere, V. Danos et al, Compositional modelling of immune response and virus transmission dynamics, *Philosophical Transactions of the Royal Society A* **380** (2022), 20210307.

[13] S. L. Wu, S. Libkind, K. Brown, E. Patterson and J. Fairbanks, Individual.jl: Rewriting individual-based models for epidemiology using graph rewriting, to appear in *Proceedings of Applied Category Theory 2022.*