SUPPORTING AGENT-BASED MODELING AND SIMULATION IN DEMOGRAPHY

Oliver Reinhardt

Institute of Computer Science
University of Rostock
Albert-Einstein-Straße 22
Rostock, 18059, GERMANY

ABSTRACT

Continuous-time microsimulation is an important method for predicting demographic changes in the next decades. However, these modeling and simulation studies face various challenges. One is to integrate adequately dynamics at different organizational levels, i.e. at micro level (e.g., individuals), and at macro level (e.g., population), and linking those dynamics in continuous time and discrete space. With more expressive models, demography has started moving from data- to hypothesis-driven development of individual-based models, which impacts modeling and simulation processes and methods used. This work aims at addressing various challenges of continuous-time agent-based simulation in demography.

1 A MODELING LANGUAGE FOR LINKED LIVES

Today, agent-based models in demography are either implemented from scratch using a general purpose programming language. The implementation of complex models, especially with scheduling in continuous time, is a tremendous effort. Also, model mechanisms are often closely intertwined with simulation details, making it difficult to understand the model, and to reuse model and simulation code. Alternatively, agent-based simulation tools such as NetLogo are used. However, these tools usually do not support continuous-time, and are not designed for the specific challenges of demographic models, e.g., age-dependent transition rates for fertility and mortality.

The Modeling Language for Linked Lives (ML3) was developed (Warnke et al. 2015) to allow for a succinct modeling of life courses of individuals, and link those life courses in continuous time and discrete space. ML3 follows an agent-based metaphor, where agents are characterized by their age, various attributes, and dynamic links to other agents, forming a dynamically evolving social network. Behavior is described by stochastic rules, guard-rate-effect triplets (see Figure 1 for an example), with Continuous-time Markov-Chain (CTMC) semantics. It should be noticed that the rate of certain demographic events, e.g., mortality or fertility, depends strongly on a persons age. Therefore, ML3 models describe an inhomogeneous CTMC, instead of the standard homogeneous CTMC used in other application fields, e.g., cell biology. This adds complexity and requires the adaptation of existing algorithms. Unlike in many other simulation systems that work population-based, individuals in ML3 are very heterogeneous and thus need to be treated individually, making many efficient simulation techniques used in stochastic simulation algorithms inapplicable. However, their individual linkages to other individuals allows to exploit that changes are only locally induced for a more efficient simulation. The algorithm (Reinhardt and Uhrmacher 2017) keeps track of changes and dependencies, and thereby reduces the rescheduling of events. While the speedup is significant, the exploration of further possibilities, e.g., by applying approximate methods or parallel simulation, is still warranted to support a thorough analysis of models by executing simulation experiments.

```
1  Person
2    |  ego.sex = "f", ego.hasPartner()
3    @  fertilityRate[ego.age]
4    -> new Person(
5         sex := ["m", "f"].random(),
6         parents := [ego, ego.partner] );
```

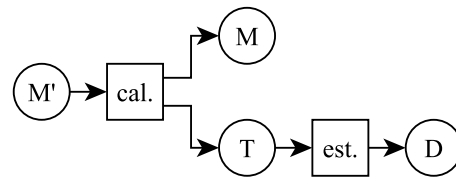Figure 1: An ML3 rule describing fertility as a guard-rate-effect triple (who-when-what).

Figure 2: Experpt of a provencance model, showing the calibration (*cal.*) of a simulation model *M*)and to a target *T*, resulting in the calibrated model *M'*. The target is estimated (*est.*) from data *D*.

2   TOWARDS STREAMLINED AND REPRODUCIBLE SIMULATION STUDIES

In addition to efficiency, another challenge simulation studies in this domain face is the integration of data driven and hypotheses driven development of models, which requires complex simulation experiments being conducted flexibly and replicably. In these experiments, the model must be compared to micro- and macro-level data, and analyzed using various statistical methods, e.g., for sensitivity analysis or uncertainty quantification. To streamline this intricate experimental process, a flexible and efficient computational environment is necessary. To this end, we created a binding between ML3 and SESSL – a domain specific language for specifying and executing simulation experiments. We demonstrated the potential of the approach by fitting two kinds of meta-model to an agent-based model of return migration, and by conducting simulation-based optimization on a model of social care (Reinhardt et al. 2018).

To fully understand and trust an agent-based model, individual simulation experiments are not enough. A thorough and accessible documentation of the model, its foundation in theory and data, and its validation is necessary. The ODD protocol (Grimm et al. 2010) addresses this need by prescribing certain content and structure for textual documentation of agent-based models. However, while ODD describes the model itself, the process of its creation and the experiments conducted remain undocumented. Therefore, we suggest complementing the ODD protocol with a provenance model of the simulation model, in an approach we call ODD+P (Reinhardt et al. 2018). A provenance model captures the history of a product as a directed graph, where nodes represent the processes and artifacts involved, and edges describe their relations (see Figure 2). We demonstrated, using a model of migration from Senegal to France, how a provenance model can provide a clear and concise overview about a model's foundations, creation, and validation.

REFERENCES

Grimm, V., U. Berger, D. L. DeAngelis, J. G. Polhill, J. Giske, and S. F. Railsback. 2010. "The ODD Protocol: A Review and First Update". *Ecological Modelling* 221(23):2760–2768.

Reinhardt, O., J. Hilton, T. Warnke, J. Bijak, and A. M. Uhrmacher. 2018. "Streamlining Simulation Experiments with Agent-Based Models in Demography". *Journal of Artificial Societies and Social Simulation* 21(3):9.

Reinhardt, O., A. Ruscheinski, and A. M. Uhrmacher. Appearing 2018. "ODD+P: Complementing the ODD Protocol with Provenance Information". In *Proceedings of the 2018 Winter Simulation Conference*, edited by M. Rabe et al. Piscataway, New Jersey: IEEE.

Reinhardt, O., and A. M. Uhrmacher. 2017. "An Efficient Simulation Algorithm for Continuous-Time Agent-Based Linked Lives Models". In *Proceedings of the 50th Annual Simulation Symposium*, edited by S. Jafer et al., 9:1–9:12. San Diego, California: SCS.

Warnke, T., A. Steiniger, A. M. Uhrmacher, A. Klabunde, and F. Willekens. 2015. "ML3: A Language for Compact Modeling of Linked Lives in Computational Demography". In *Proceedings of the 2015 Winter Simulation Conference*, edited by L. Yilmaz et al., 2764–2775. Piscataway, New Jersey: IEEE.