Taking Agent-Based Social Simulation to the Next Level Using Exascale Computing: Potential Use Cases, Capabilities & Threats

Extended Abstract

Matt Hare
The James Hutton Institute
Aberdeen, United Kingdom
matt.hare@hutton.ac.uk

Doug Salt
The James Hutton Institute
Aberdeen, United Kingdom
doug.salt@hutton.ac.uk

Ric Colasanti
University of Glasgow
Glasgow, United Kingdom
ricardo.colasanti@glasgow.ac.uk

Richard Milton
University College London
London, United Kingdom
richard.milton@ucl.ac.uk

Mike Batty
University College London
London, United Kingdom
m.batty@ucl.ac.uk

Alison Heppenstall
University of Glasgow
Glasgow, United Kingdom
alison.heppenstall@glasgow.ac.uk

Gary Polhill
The James Hutton Institute
Aberdeen, United Kingdom
gary.polhill@hutton.ac.uk

ABSTRACT

Exascale computing (10^{18} FLOPS) officially arrived in 2022 when Oak Ridge's Frontier achieved that performance benchmark, and other countries seek their own exascale capabilities. High-end computing is typically used by the natural sciences, but empirical Agent-Based Social Simulation (ABSS) is a social science application. Empirical ABSS has a long history, but was prominent during the Covid crisis. In future crises, policy options could be evaluated within rapid policy design windows using exascale computing. We report on a group model-building exercise, co-constructing a causal loop model, to explore visions of the potential of exascale computing in ABSS, identifying potential use cases, capabilities, capacity requirements and threats.

KEYWORDS

Exascale Computing; Agent-Based Social Simulation; Use Cases; Threats

ACM Reference Format:
Matt Hare, Doug Salt, Ric Colasanti, Richard Milton, Mike Batty, Alison Heppenstall, and Gary Polhill. 2024. Taking Agent-Based Social Simulation to the Next Level Using Exascale Computing: Potential Use Cases, Capabilities & Threats: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024,* IFAAMAS, 3 pages.

1 INTRODUCTION

We report here on our engagement with the empirical Agent-Based Social Simulation (ABSS) community on how exascale computing could support their needs, and how it might affect future directions for ABSS research. Use of High Performance Computing (HPC) is rare in the ABSS community. Social scientists face a steep learning curve simply to build a working agent-based model in the first place [1, 2]. Though it need not necessarily be so, HPC environments require even more computing skills to access. Even then, agent-based models have features (such as varying population sizes) that make their memory and CPU-time unpredictable, while access to and continuing use of HPC facilities is usually predicated on being able to make such predictions reliably [9]. These two points combine to exacerbate a third, cultural issue around HPC regarding the suitability of the code and the scientific problem it is investigating for the advanced computing machinery on which it will be run. These are points that we can expect to be more pronounced in exascale environments that cost hundreds of millions of euros to build.

Fourteen years after a review of possibilities for exascale computing in ABSS [5], exascale computing officially arrived. Exascale computers are about a billion times faster than a personal computer, and a thousand to a million times faster than a university's computing cluster. ABSS became particularly prominent during the Covid crisis [e.g. 3, 10], but work was still hampered by constraints on computing resources. Exascale computing power is such that the computation could be reduced to seconds in future crises.

We report here on envisioning how exascale could transform ABSS use, for which we have held the first of a series of group model-building exercises. As well as use cases, the exercise elicited exascale ABSS capabilities, capacity requirements and threats.

2 METHOD

The group model building took place within a two-hour workshop hosted at an international conference in 2023. The method centred on co-constructing a raw causal loop model (CLM), - a directed cyclic graph [12]. This group model building methodology was

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

Table 1: The top two nodes of each type in terms of different degree metrics. Degree indicates the total number of edges a node has (a proxy for overall importance), divided into out degree (a proxy for its influence on other nodes) and in degree (a proxy for other nodes' influence on it).

Node	*Node Type*	*Out Degree*	*In Degree*	*Degree*
Running simulations faster	Exascale capability	9	5	14
(Faster) identification of the best ways of doing something	Exascale capability	3	2	5
Realtime error correction and visualization of testing	Use case	1	4	5
General emulator of simulations	Use case	3	1	4
Writing higher level code for ABSS	Capacity required	3	2	5
AI data mining	Capacity required	4	1	5
Risk that gatekeepers over-protect access to exascale computing	Threat	2	1	3
Over-complication of models	Threat	2	1	3

chosen in favour of alternatives for various reasons. First, it represents the diverse perspectives of the participants while allowing them to direct the development of knowledge represented by the model. Second, as a contrived method [7], it permits the elicitation of tacit as well as explicit knowledge. Third, as a systems approach, it provides an integrated view of interconnections among elements of participants' visions. Fourth, it provides a formal representation amenable to network analysis and qualitative reasoning.

The raw CLM was co-constructed using VUE software [11] projected on two large screens always visible to the participants. The procedure used was similar to the approaches described in [4, 12], but reversed the order of the cause and consequence steps:

1. Individually, participants thought about and then wrote down what frustrated them about ABSS development and use, and how they thought exascale computing might provide something better. In plenum, the Knowledge Engineer (KE) drew their answers as 'start' nodes in the CLM, encoded as 'exascale capabilities'.

2. For each start node, the KE asked participants, "What does *this node* bring us in terms of ABSS development and use?" The KE then created new nodes for each answer, linking them to the start node via edges, and encoding the new nodes as 'exascale capabilities' or 'use cases' as appropriate.

3. For each new node, the KE asked participants, "What capacities do you require to realize the improvement represented by *this node*?" Further nodes were created, linked by new edges to the appropriate node and encoded as 'capacities'.

4. The KE asked participants to identify potential threats from exascale ABSS, encoding answers as 'threat' nodes linked to other nodes in the CLM as appropriate.

After the workshop, the raw CLM was then cleaned up and refined by the KE (e.g. to ensure nodes had been correctly categorised) to create the final CLM, in a process supported by 'blind judge' verification [e.g. 6].

3 RESULTS AND DISCUSSION

The final CLM, integrating the different participants' visions on the potential of exascale ABSS, was composed of 58 nodes and 63 edges. The results of applying network centrality metrics of degree, out-degree, and in-degree to the final CLM can be seen in Table 1. This metric is useful in reflecting important nodes identified by the participants that greatly influenced, and/or were greatly influenced by, causal chains in the CLM. By far the most important node was the capability to run simulations much faster. Next important in terms of influencing other causal chains i.e. out-degree, was the capacity required for AI data mining in order to improve the explainability of complex exascale ABSS and to support development of higher level coding blocks, e.g., for controlling parallel processing in exascale ABSS.

Being able to run simulations faster provides the capability to carry out more model refinement iterations. The CLM showed how this could lead to faster identification of the best ways of doing something and thus to new use cases such as using exascale ABSS to identify the the best formalization of a given social theory - a current challenge for ABSS [8].

The exercise was successful in eliciting other new use cases e.g., the development of generative simulation emulators/meta-models trained on billions of simulation runs. Interestingly, faster debugging and visualization of results of testing was also raised as a use case - something exascale computing 'gatekeepers' would expect to have been done *before* running a model on such an expensive machine.

Finally, the CLM highlighted threats resulting from exascale ABSS development. In addition to increased energy consumption and the potential end of 'normal', giga-scale ABSS, it identified threats such as the encouragement of spuriously over-complicated models and inequitable access to exascale resources due to overprotective gatekeeping.

Further workshops will follow up this work to identify practical pathways for taking ABSS to the next level in terms of achieving the desirable visions identified, and mitigating key threats. Such identification will be important since some participants warned against over-optimism, questioning whether exascale computing might lose out to another technological paradigm shift (e.g. quantum computing) before it can mature. Also, the following challenge was posed: is sheer computing power really the main bottleneck to taking ABSS up to the next level?

ACKNOWLEDGMENTS

We acknowledge support from the Engineering and Physical Sciences Research Council (grant number EP/Y008839/1), and the Scottish Government Rural and Environment Science and Analytical Services Division (project reference JHI-C5-1).

ETHICS STATEMENT

The final causal loop model (CLM) captured two areas leading to ethical concerns.

i) Developing the capacity to produce workable institutional regulations to ensure equitable access to exascale resources will be very important. Without this, as mentioned in [9], the result could be reduced access for social scientists who, for example, would not necessarily, by training, have the technical capacity to provide technical use-justifications. As mentioned in the workshop, inequitable access might also particularly affect those *not* from the 'Global North'.

ii) The CLM also revealed concerns as to who will end up using exascale ABSS and for what purposes. Using exascale to develop fine-grained ABSS working at large population or geographical scales will require both large amounts of input data and increased capacities in terms of AI-based machine learning tools to understand the resulting models. The concern was that both the scale of complexity of exascale ABSS and the use of AI to explain such models may concentrate ABSS ownership in larger, more powerful actors who may use it for purposes that might not be transparent.

Both areas of ethical concern would benefit from increased capacities for the monitoring of such possible trends as exascale ABSS is operationalized over time.

REFERENCES

[1] Lilian Na Alessa, Melinda Laituri, and Michael Barton. 2006. An "all hands" call to the social science community: establishing a community framework for complexity modeling using agent based models and cyberinfrastructure. *Journal of Artificial Societies and Social Simulation* 9, 4 (2006), 6. https://www.jasss.org/9/4/6.html

[2] Li An, Volker Grimm, and Billie L. Turner II. 2020. Editorial: meeting grand challenges in agent-based models. *Journal of Artificial Societies and Social Simulation* 23, 1 (2020), 13. https://doi.org/10.18564/jasss.4012

[3] Jennifer Badham, Pete Barbrook-Johnson, Camila Caiado, and Brian Castellani. 2021. Justified stories with agent-based modelling for local COVID-19 planning. *Journal of Artificial Societies and Social Simulation* 24, 1 (2021), 8. https://doi.org/10.18564/jasss.4532

[4] Johannes Halbe and Claudia Pahl-Wostl. 2019. A methodological framework to initiate and design transition governance processes. *Sustainability* 11, 3 (2019), 844. https://doi.org/10.3390/su11030844

[5] Charles M. Macal and Michael J. North. 2008. Agent-based modeling and simulation for exascale computing. *SCIDAC Review* Summer 2008 (2008), 34–41.

[6] N. A. M. Maiden and M. Hare. 1998. Problem domain categories in requirements engineering. *International Journal of Human-Computer Studies* 49 (1998), 281–304. Issue 3. https://doi.org/10.1006/ijhc.1998.0206

[7] P. McGeorge and G. Rugg. 1992. The uses of 'contrived' knowledge elicitation techniques. *Expert Systems* 9, 3 (1992), 149–154. https://doi.org/10.1111/j.1468-0394.1992.tb00395.x

[8] Hannah Muelder and Tatiana Filatova. 2018. One theory – many formalizations: testing different code implementations of the Theory of Planned Behaviour in energy agent-based models. *Journal of Artificial Societies and Social Simulation* 21 (2018), 5. Issue 4. https://doi.org/10.18564/jasss.3855

[9] Gary Polhill. 2022. Antisocial simulation: Using shared high-performance computing clusters to run agent-based models. *Review of Artificial Societies and Social Simulation* December 2022 (2022). https://rofasss.org/2022/12/14/antisoc-sim/

[10] Jason Thompson, Rod McClure, Tony Blakely, Nick Wilson, Michael G. Baker, Jasper S. Wijnands, Thiago Herick De Sa, Kerry Nice, Camilo Cruz, and Mark Stevenson. 2022. Modelling SARS-CoV-2 disease progression in Australia and New Zealand: an account of an agent-based approach to support public health decision-making. *Australian and New Zealand Journal of Public Health* 46, 3 (2022), 292–303. https://doi.org/10.1111/1753-6405.13221

[11] Tufts University. 2023. *VUE User Guide.* https://sites.tufts.edu/vue/

[12] Jac A. M. Vennix. 1996. *Group Model Building: Facilitating Team Learning Using System Dynamics.* Wiley, Chichester, New York, USA.