Agent-based models in the philosophy of science

2022 Summer School of the Vienna Doctoral School of Philosophy

Samuli Reijula, University of Helsinki
samuli.reijula@helsinki.fi
https://www.samulireijula.net

Events this Summer (in Helsinki)

- Nordic Network for the Science of Science
  www.nordicscisci.net

- Institutional epistemology workshop 2022
  www.institutionalepistemology.net

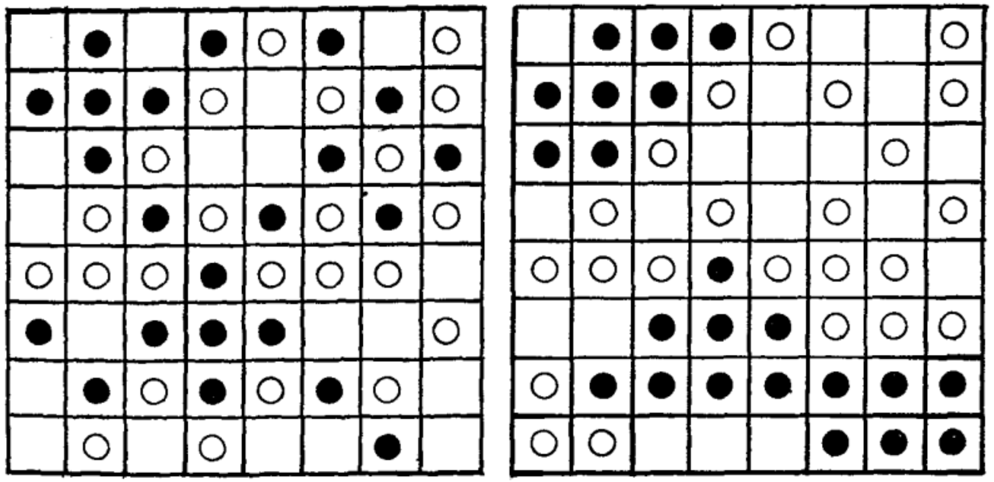
Background materials

a) An introductory course to Python (e.g. codeacademy)

a) Readings:
  - Epstein, J. M. (2008). Why model? *Journal of artificial societies and social simulation*, 11(4), 12. https://www.jasss.org/11/4/12.html
  - Aydinonat, N. E., Reijula, S., & Ylikoski, P. (2021). Argumentative landscapes: the function of models in social epistemology. *Synthese*, 199(1), 369-395.
  - Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, 41(2), 203-226.

(1)
Modeling as philosopher’s tool

The Schelling model



(a) (b)

Models everywhere?

- Mental models (Johnson-Laird 1983, 2009) vs. scientific models
- Material models vs. formal models
- Theoretical vs. data-driven models
  - Modeling mechanisms/laws vs. modelling patterns in data
- Equation-based vs. agent-based models (ABMs)
- …

- Different ways of exploring a model: analytical derivation vs. simulation

- Philosophy as modeling? (Godfrey-Smith 2006, 2012; L.A.Paul 2012; Williamson 2017; Wimsatt 2007)

Agent-based models

Traditional Tools	*Agent-Based Objects*
Precise	Flexible
Little process	Process oriented
Timeless	Timely
Optimizing	Adaptive
Static	Dynamic
1, 2, or ∞ agents	1, 2, . . . , N agents
Vacuous	Spacey/networked
Homogeneous	Heterogeneous

Miller & Page 2007, p.79

- Agent-based models = ABMs
    - In ecology and physics, IBMs, "individual-based models"
- Components: Population of agents interacting in an environment
    - Agent behavior typically rule-based
    - Interactions both between agents and an agent and environment
- Characteristics
    - Easy to introduce complexity, e.g. diversity among agents, resources, or environment sites
        - (vs. equation-based models)
    - Useful for exploring micro→macro relations

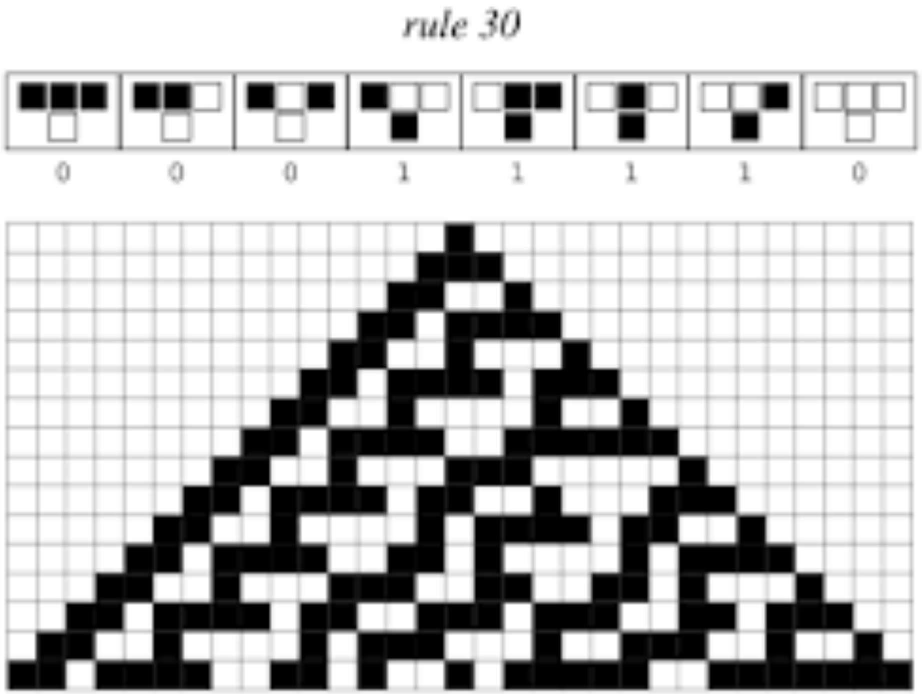
Agent-based models (ii)

Traditional Tools	*Agent-Based Objects*
Precise	Flexible
Little process	Process oriented
Timeless	Timely
Optimizing	Adaptive
Static	Dynamic
1, 2, or ∞ agents	1, 2, . . . , N agents
Vacuous	Spacey/networked
Homogeneous	Heterogeneous

Miller & Page 2007, p.79

- …
- Typically ”solved” by simulation
  - Simulation = “use of the partially autonomous behaviour of a surrogate artefact to mimic some process in the target system” (Beisbart 2012)
  - Example: Axelrod 1997: the model is a Markov chain, but a complicated one → not feasible to analytically derive results
- … but not always a computer simulation (original Schelling model on a checkerboard)!

Simple to complex



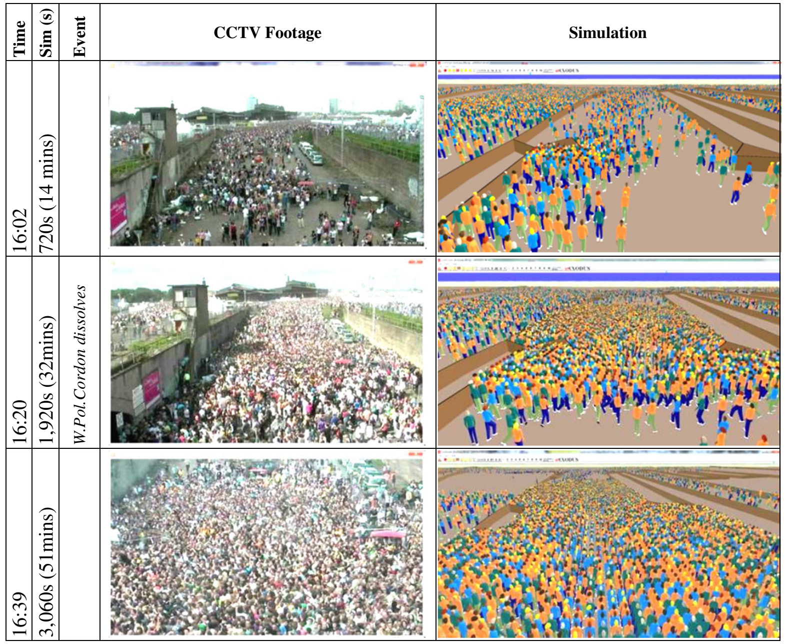
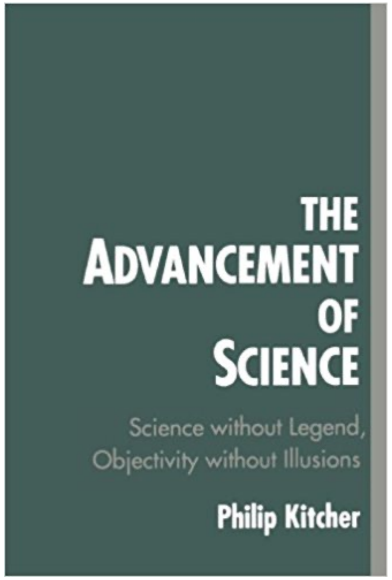
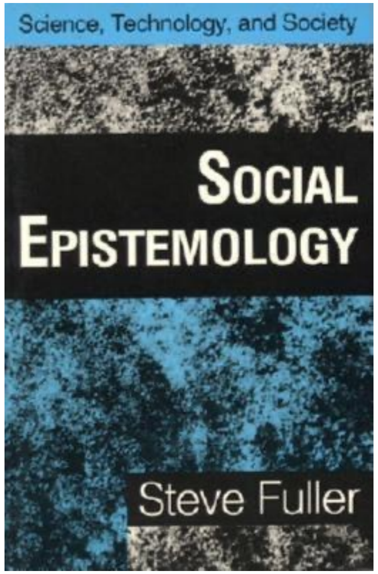


Figure 5. Timeline: comparison of the conditions on the main access area (video stills source [8]).

Pretorius & al. 2015





”

The general problem of **social epistemology**, as I conceive it, is to identify the properties of epistemically well-designed social systems.

”

“

How should the pursuit of knowledge be organized, given that […] knowledge is pursued by many human beings, each working on a more or less well-defined body of knowledge and each equipped with roughly the same imperfect cognitive capacities, albeit with varying degrees of access to one another's activities?

“

Models of social organization of science

1st generation

- (Early opinion dynamics models: French 1956 → Lehrer and Wagner 1981)
- Resource allocation in theory choice (Kitcher 1990)
- In science studies: Gilbert 1977

2nd generation (mostly ABMs)

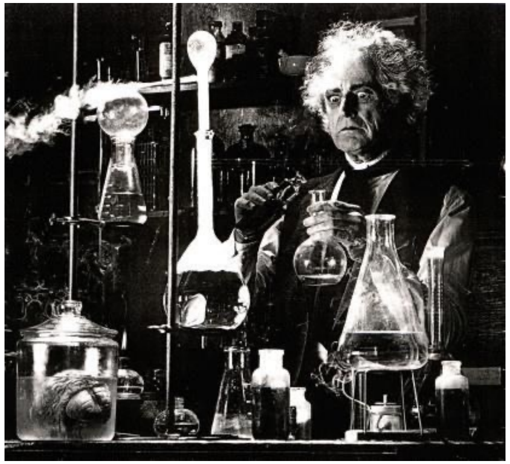
- Opinion dynamics with truth signal (Hegselmann & Krause 2006→)
- Epistemic network models (Zollman 2007→)
- Epistemic landscape models (Weisberg & Muldoon 2009 →)

And more

- Group problem solving models (Grim et al. 2019; Reijula & Kuorikoski 2021)
- …

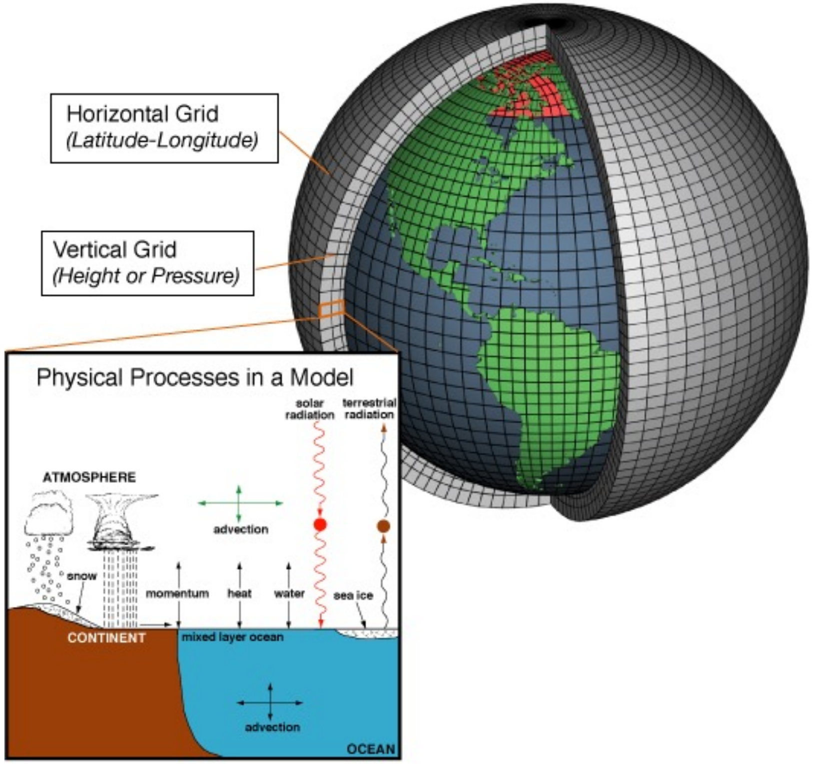
Reviews

- Reijula & Kuorikoski 2019: Modeling epistemic communities
- Seselja 2020: Exploring Scientific Inquiry via Agent-Based Modelling

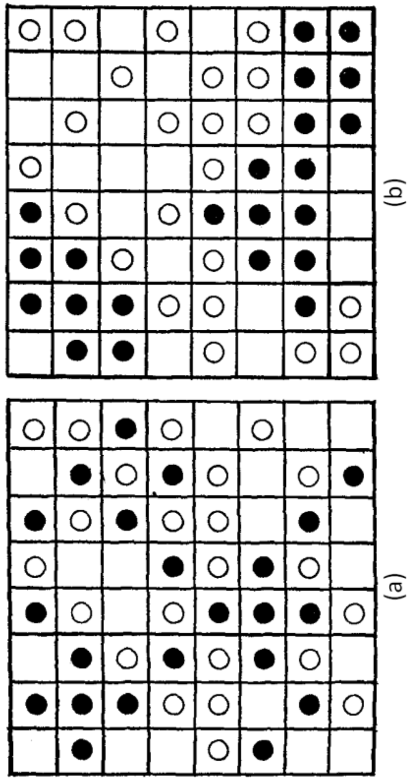


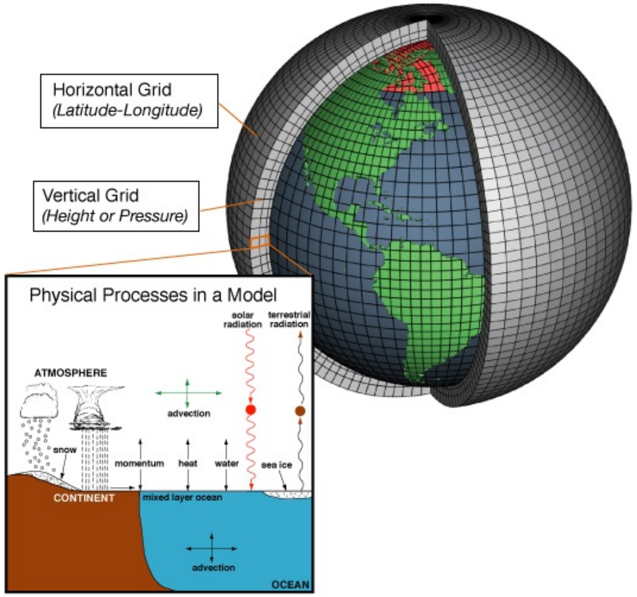
A contrast: climate modeling

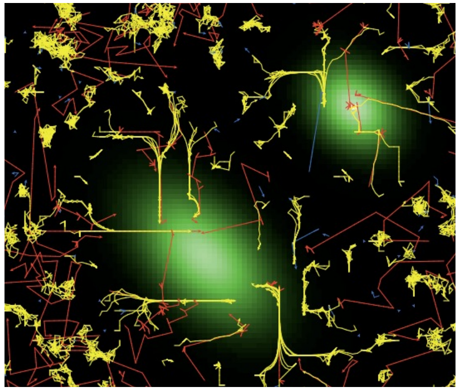
- Different weather and climate models rely on similar principles:
  - 1. Sample the state of the fluid at a given time
  - 2. Use the equations of fluid dynamics and thermodynamics to estimate the state of the fluid at some time in the future
- Masses of measurements of temperature, moisture, wind fields, …
- .. From various sources: weather balloons, satellites, aircrafts, ships … used as initial conditions for predictions
- The Earth discretized: size of grid, temporal steps
- Run on supercomputers: solving differential equations --> weather predictions for future time states



Do different kinds of models have anything in common?







When is a model successful, a good model?

→ We first need to answer what a model is for

Modeling is fun

Modeling is a great form of theorizing

Working with models can make you a better thinker

Epistemology of computational modeling

Why do we model?

Used both in science
and in philosophy!

Theoretical models can be used for surrogative reasoning

Analytical;
Computer
simulation;
Analog model

Vs.
data models,
phenomenological
models

”Use one sort of thing as a
surrogate in our thinking
about another”
(Swoyer 1991)

Why is surrogative reasoning possible?

- Surrogative reasoning: Learning about the **target** by investigating the **source** (model)

- Option 1: Models **carry information** about their targets
  1. Establish the representation relation between source and target
  2. → Manipulating the model allows us to learn about the target …
  3. …just like in material experiments

Simulations as experiments

Norton & Suppe 2001	“Computer modeling constitutes a form of world building, where controlled experiments impossible to perform physically can be run in virtual environments”
Mäki 2005	”Models are experiments, experiments are models”
Morrison 2009	"The paper presents an argument for treating certain types of computer simulation as having the same epistemic status as experimental measurement"
Barberousse et al. 2009	”Simulations can be used as experiments because they represent phenomena”

Miller and Page 2007. p.88

6.10 Economic *E. coli* (*E. coni*?)

The interaction between the theorist and the computational model provides an ideal medium from which theoretical insights can be gleaned (Tesfatsion, 1997, 2006). Agent-based object models give the theorist some rather intimate experiences with the phenomena of interest. As we have outlined, these artificial worlds are fully observable, recoverable, and repeatable, and thus they are a fertile playground from which theories can be created, refined, and tested. Like many theoretical tools, computational models have the potential to produce insights well beyond those needed to implement the original model.

A fanciful, but perhaps ultimately enlightening, use of agent-based object models is as an "animal" model for the social sciences. The ability to experiment with animal models like *E. coli* in biology and *Drosophila* in genetics, has led to great advances in our understanding of human systems. Unfortunately, there is not an obvious choice of an animal model for social systems research. Indeed, even human-based experiments are relatively new in fields like economics, where their results are just beginning to facilitate the process of scientific creative destruction. While the possession of a simple animal model is not necessary for scientific progress—both economics and astronomy were developed around passive-observation-based methodologies—having an animal model may lead to new scientific opportunities.

A problem

How can we learn
something genuinely new about
the world
without directly empirically investigating
the target system itself?

My view: models are not experiments

- The **praxis of simulation modeling** resembles experimentation
  - In material experiments you have "causal control", in models "theoretical control"

- But the analogy breaks down at the level of epistemology

- Principle (first order formulation): "NO CAUSAL INTERACTION, NO INFORMATION FLOW"
  - 'Getting new information concerning target system T':
    Learning about a property P_j of T which is not a logical consequence of the properties of T $\{P_1 \ldots P_i\}$ that we already know about

Heuristic uses for models

Perhaps modeling serves only heuristic purposes

a) Conceptual precision & innovation:
   - Making concepts precise: E.g. Different things we might mean by ’cognitive diversity’
   - Sharpening questions
b) Introducing new ways of thinking about a problem
   - E.g. scientific problem solving as an epistemic landscape
c) Highlight new factors of interest & new aspects of a problem
d) Remind us that things are complex (especially for policy purposes)
   - In many situations, several competing processes in play

But I think more can be said…

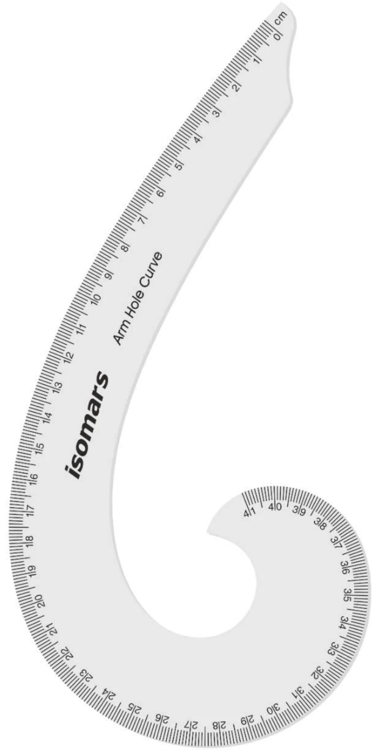
Feynman on the French curve

At MIT, in mechanical drawing class, someone wondered if the curves on French curve had a formula

Jokingly, Feynman replied that "sure, the French curve is made so that at the lowest point on each curve, no matter how you turn it, the tangent is horizontal"

"All the guys in the class were holding their French curve up at different angles, holding their pencil up to it at the lowest point and laying it along, and discovering that, sure enough, the tangent is horizontal. They were all excited by this 'discovery' — even though they had already gone through a certain amount of calculus and had already 'learned' that the derivative (tangent) of the minimum (lowest point) of *any* curve is zero (horizontal). The didn't put two and two together. They didn't even know what they 'knew' "

(Feynman: *Surely you're joking mr. Feynman*, p.24)



What the scientists say

Kokko 2007, p.7 ”Models do not investigate nature. Instead, they investigate the validity of our own thinking, i.e. whether the logic behind an argument is correct.”

Krugman 1998, p.1834 The main function of modelling is “keeping things straight” and “helping to focus and form intuitions” in contexts involving the adding-up of constraints, indirect chains of causation and feedback effect.

Williamson on the epistemic role of models

Williamson, Timothy 2017 "Model-Building in Philosophy", in Blackford and Broderick (eds.), *Philosophy's Future: The Problem of Philosophical Progress.* Oxford: Wiley

"[I]n philosophy, too, one form of progress is the development of better and better models […] Not only can philosophy make progress through model-building, it has been doing so for quite some time"

- For example, in epistemology, a standard model of epistemic uncertainty is a lottery

"When we explore a model by valid deductive reasoning from the model description, we learn necessary truths of the general conditional form "If a given case satisfies the model description, then it satisfies this other description too."

"That broadly logico-mathematical knowledge has the virtue of precision, but by itself is less than we want, since it says nothing unconditional about how close the original phenomenon (such as predator-prey interaction) comes to satisfying the model description."

"If we started in total ignorance of the target, we could hardly expect to learn much about it by modelling alone."

God does not simulate

- What is the epistemic role of models, then?

- Alternative interpretation of surrogate **reasoning**
  - Modeling assumptions are motivated by the target
  - The model helps you to investigate what logically follows from such assumptions
- Models are external scorekeeping devices that improve the scope and reliability of the inferences we make from modeling assumptions
- Although models would be of no use for logically omniscient agents, they often are necessary for boundedly rational agents like us
- They **increase understanding** (see Ylikoski & Kuorikoski 2010): They extend our inferential reach, and allow us to explore counterfactual scenarios

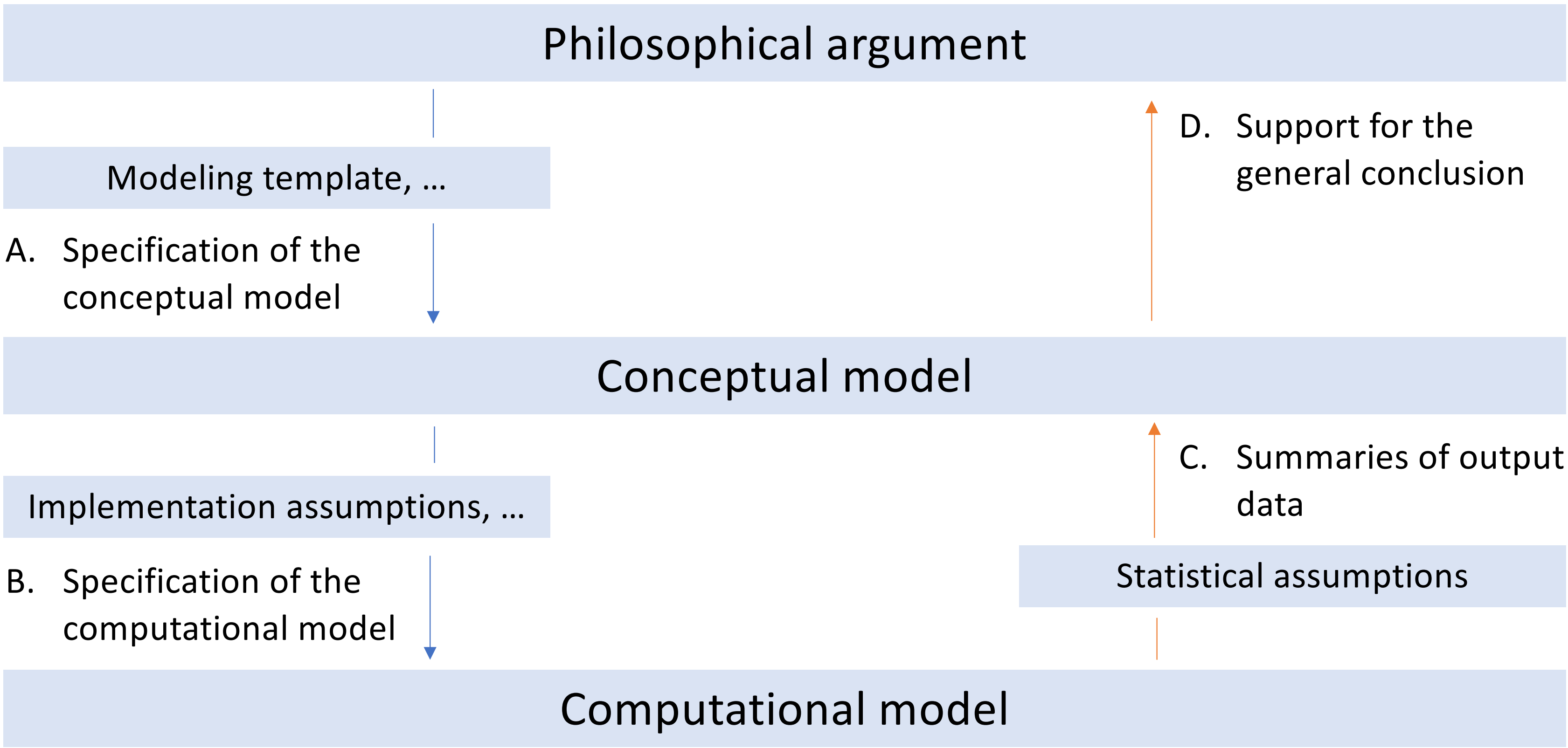
- Furthermore, formal models enable a social way of working, others can check and improve your work (Edmonds 2019)

An unsettled debate

- The debate on the epistemic powers of models has not been settled
- My reason for holding the scorekeeping view: Considerations of the burden of proof
    - Scorekeeping account avoids magical epistemology (learning about target without causal interpretation)
    - Scorekeeping account is in line with scientists' judgments on the role of models
    - Still waiting for counterexamples to the "no causal interaction, no information flow" principle
- (see Kuorikoski & Reijula, unpublished):

Using models as argumentative devices

Argument and the layers of modeling



Conceptual model: premises and conclusions

Premises		Conclusions
• [P1] Agents: motivation, perception, decision-making, prediction • [P2] Their form of organization (such as network structure) and interactions, • [P3] Initial states and distribution of agent properties, and • [P4] The structure of the environment … • (see Railsback & Grimm 2019, p.99)	→ ? →	• [C1] System dynamics (equilibria, cycles) • [C2] Spatial configurations (e.g., segregation patterns), • [C3] Thresholds (e.g., tipping points) • …

Establishing validity

- ABMs are **inferentially opaque** (Humpreys 2004): we don’t directly see how (and whether) the conclusions follow from the set of premises
  - Example1: Weisberg & Muldoon 2009: ”Mavericks **stimulate** the followers to make considerable epistemic and total progress”
    - Alexander et al. 2015: Not a valid conclusion to draw. Insufficient data analysis.
  - Example2: Discrepancy between follower agents in the conceptual model and in the computational model (Alexander et al; Pöyhönen 2017)

- Strategies for establishing validity
  - Using well-known and tested submodels
  - Unit testing
  - Rigorous data analysis at multiple levels
  - Process tracing, reality checks
  - …

What kind of conclusions do you aim for?

- How does the practice of simulation modeling resemble experimentation?
  1. Introduce controlled variation into the modeling assumptions
  2. Apply Mill's method of difference
  3. → Draw conclusions about **difference-making**
- **Minimal difference making:** A pair of sets of premises and conclusions, where there's a difference in one premise only, and a difference in the conclusion of interest
  - notice: not causal control, but theoretical control
- E.g., Schelling model:
  "20% similar wanted / [30% similar wanted] → no segregation / [segregation"

Note:
The similarity with material experiments to be found on the level of experimental praxis, not epistemology

When does an ABM support a difference-making claim

1. Validity
   - Conclusions from the computational model must be logical consequences of the modeling assumptions
   - The computational model must correctly implement conceptual model
   - Findings from the conceptual model must support the general conclusion

2. Independent variables of interest must make a different to outcomes of interest

3. Insignificant/tractability assumptions must *not* make a difference to *stability* of such a dependency
   - (~ invariance)

Generalizing difference-making results

- **Minimal difference-making**: dependency between IV and DV observed when all other parameters held at specific values
- Establishing minimal difference-making is rarely enough for the argumentative purpose
  - Concerns about the relevance of the findings (Rosenstock et al. 2017)
- → Wiggle your assumptions
  - What happens across changes to substantial assumptions → **Generality**
  - What happens across changes to non-substantial assumptions → **Derivational robustness**

- Strategies
  - Sensitivity analysis / robustness analysis (Frey & Šešelja 2018)
  - Building model families (Aydinonat & Ylikoski 2014)

From model to the world?

Models as argumentative devices: ”Models examine the validity of our thinking”
→ how do we get to results about real-world targets?

- Indirect strategy for representing
  - ”Models are not pictures” → Do not map model components onto components of real world target-systems
  - Instead, your simulation findings apply to scenarios (possible worlds) where the modeling assumptions are true
- Advantages
  - A natural way to deal with abstractions & omissions
  - No mystery in applying a model to widely different application domains
  - Connection to empirical research: ABMs as platforms for integrating (empirical) information from various sources

Literature

- Aydinonat, Emrah, Samuli Reijula and Petri Ylikoski (2021). Argumentative landscapes: the functions of models in social epistemology. *Synthese*, 199, 369-395.
- Frigg, R. & Nguyen: *Modelling nature : an opinionated introduction to scientific representation*
- Beisbart, C. 2018. Are Computer Simulations Experiments? And If Not, How Are They Related to Each Other? *European Journal for Philosophy of Science* 8 (2): 171-204.
- Boge, F.2019. Why computer simulations are not inferences, and in what sense they are experiments. *European Journal for Philosophy of Science*.
- Bueno, O. 2014. Computer simulations: an inferential conception, *The Monist* 97, 378-398
- Edmonds (2019). The evolution of empirical ABMs. https://www.slideshare.net/BruceEdmonds/the-evolution-of-empirical-abms-137463946
- Giere, R. 2009. Is computer simulation changing the face of experimentation? *Philosophical Studies* 143: 59–62.
- Hughes, R. I. 1997. Models and representation., *Philosophy of science* 64, S325-S336.
- Kuorikoski, Jaakko and Samuli Reijula. Making it count. An inferentialist account of computer simulation. https://osf.io/preprints/socarxiv/v9bmr
- Mayo-Wilson, C., & Zollman, K. J. (2021). The computational philosophy: simulation as a core philosophical method. *Synthese*, 199(1), 3647-3673.
- Reijula and Kuorikoski 2019. Modeling epistemic communities, in M. Fricker, P.J. Graham, D. Henderson, N. Pedersen, J. Wyatt (eds.) *The Routledge Handbook of Social Epistemology*, 240-249.
- Schelling, T. 1971. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1: 143–186
- Šešelja, D. (2022). Agent-based models of scientific interaction. Philosophy Compass, e12855.
- Weisberg, M. 2013. *Simulation and Similarity*. Oxford University Press
- Winsberg 2010: *Science in the Age of Computer Simulation,* University of Chicago Press.

The anatomy of an ABM

Why Python

- Python is an interpreted scripting language compatible with various programming styles: object oriented programming, functional programming …
  - Basic data type: list (in this sense, belongs to the LISP lineage)
- … But wouldn't it be easier to use something like Netlogo?
- Advantages of using Python (or R):
  - Huge, reliable ecosystem
    - Ready-made tools available for almost any purpose (numpy, scipy, scikit-learn …)
    - Support: courses; tutorials; googling answers almost any question you have
  - General-purpose programming language
    - No unnecessary constraints on what you can do
    - You get general programming skills
  - Possible to interface with compiled languages, if you need speed
  - …
- Why Jupyter notebooks
  - Literate programming (Knuth)
  - Good for scientific computing / data-science applications
    - Visualizations in the notebook itself
    - Easy to move the project to a virtual machine
  - …

Walk-through of a model: Does diversity beat ability

PNAS

Groups of diverse problem solvers can outperform groups of high-ability problem solvers

Lu Hong[†‡§] and Scott E. Page[¶]

[†]Michigan Business School and [¶]Complex Systems, University of Michigan, Ann Arbor, MI 48109-1234; and [‡]Department of Finance, Loyola University, Chicago, IL 60611

Edited by William J. Baumol, New York University, New York, NY, and approved September 17, 2004 (received for review May 25, 2004)

We introduce a general framework for modeling functionally diverse problem-solving agents. In this framework, problem-solving agents possess representations of problems and algorithms that they use to locate solutions. We use this framework to establish a result relevant to group composition. We find that when selecting a problem-solving team from a diverse population of intelligent agents, a team of randomly selected agents outperforms a team comprised of the best-performing agents. This result relies on the intuition that, as the initial pool of problem solvers becomes large, the best-performing agents necessarily become similar in the space of problem solvers. Their relatively greater ability is more than offset by their lack of problem-solving diversity.

equal ability, functionally diverse groups outperform homogeneous groups. It has also been shown that functionally diverse groups tend to outperform the best individual agents, provided that agents in the group are nearly as good (1). These results still leave open an important question: Can a functionally diverse group whose members have less ability outperform a group of people with high ability who may themselves be diverse? The main result of our paper addresses exactly this question.

Consider the following scenario: An organization wants to hire people to solve a hard problem. To make a more informed decision, the organization administers a test to 1,000 applicants that is designed to reflect their individual abilities in solving such a problem. Suppose the applicants receive scores ranging from

Modeling Cognitive Diversity in Group Problem Solving

Samuli Reijula, Jaakko Kuorikoski

University of Helsinki

Diversity-beats-ability theorem

- **The question**

  Page 2014: "Understanding what produces and maintains cognitive diversity takes on practical importance as the challenges that we face—climate change, poverty, and managing the world's financial system—become more complex"
- **The claim**

  Diversity-beats-ability theorem (DAB) (Hong and Page 2004, Page 2008): Groups of diverse problem solvers can outperform groups of high-ability problem solvers
- **The mechanism**

  Groups of high-ability problem solvers tend to be less diverse. Thus they tend to rely on similar methods of problem solving
- **The evidence**

  Analytical and computational models

Our research questions

- Do the models support the truth of DAB in real life? Under what conditions does diversity beat ability?
- Replications, critique, and further developments:
  1. One-dimensional case: From *"Ringworld model"* (Hong and Page 2004) to *Stairway model* (Reijula and Kuorikoski 2021)
  2. Multidimensional case: Original binary string model
  3. The analytical theorem in Hong and Page 2004
- **The bottom line**

  Our findings suggest that the model-based arguments for DAB by Hong and Page are unsound

Ringworld model

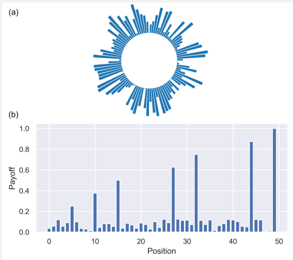


Figure: Problem spaces: (a) Random landscape used by Hong and Page 2004; (b) Stairway landscape

- Replication of Hong and Page 2004 shows that it is not possible to amplify the weak DAB effect
- Diagnosis: expert heuristics cannot function on random landscapes (Kauffman and Levin 1987). H&P merely compare exhaustive search to non-exhaustive
- Conclusion: the model does not capture *the trade-off* between diversity and ability

Stairway model

- A modified ringworld model with a structured problem space
- Demonstrates the trade-off between diversity and ability: depending on properties of the problem space, favors either high-ability of diverse groups
- Distinction between two characteristics of a problem: easy-difficult vs. simple-complex

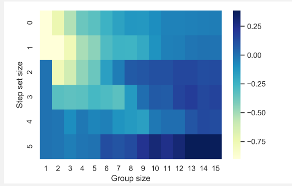


Figure: Stairway landscapes: High-ability-vs-random group performance differential. Lighter shades stand for high-ability group advantage

Binary string model

- Used by Page 2008 to explain DAB
  - Better implementation of the intuitions underlying DAB
- **Replication**: no evidence of DAB
- **Diagnosis**
  - Again, no structure for heuristic search to exploit
  - No interaction between different flipset heuristics
    → Stunted search paths

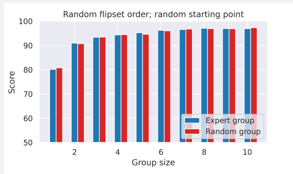


Figure: Diversity and ability in binary string model

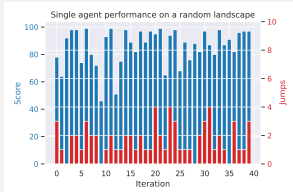


Figure: Single agent performance in the binary string model. Score mean: 85.9

The analytical theorem

We argue: No responsible way to infer from a weak conceptual possibility, proven at the limit, to any finite real-life case in which expertise and diversity may trade-off

References

[1] Lu Hong and Scott E. Page.
Groups of diverse problem solvers can outperform groups of high-ability problem solvers.
Proceedings of the National Academy of Sciences of the United States of America, 101(46):16385–16389, 2004.

[2] Stuart Kauffman and Simon Levin.
Towards a general theory of adaptive walks on rugged landscapes.
Journal of Theoretical Biology, 128(1):11–45, 1987.

[3] Scott E. Page.
The difference: how the power of diversity creates better groups, firms, schools, and societies.
Princeton University Press,, Princeton, N.J. ;Woodstock, paperback edition, 2008.

[4] Scott E. Page.
Where diversity comes from and why it matters?
European Journal of Social Psychology, 44(4):267–279, 2014.
Publisher: Wiley Online Library.

[5] Samuli Reijula and Jaakko Kuorikoski.
The Diversity-Ability Trade-Off in Scientific Problem Solving.
Philosophy of Science, 88(5):894–905, 2021.

Contact Information

- samuli.reijula@helsinki.fi; @samulipo
- jaakko.kuorikoski@helsinki.fi

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Beamer template: Jacobs Landscape Poster

Walk-through of a model: Does diversity beat ability
(Reijula & Kuorikoski 2021)

- A model of group problem solving
  - Introduced by Hong & Page 2001, 2004; Page 2008
  - Famous result: **Diversity-beats ability theorem**: Groups of diverse problem solvers can outperform groups of high-ability problem solvers
  - The Reijula & Kuorikoski 2021 paper re-examines the findings by Hong & Page and develops a version of the model more appropriate for studying the ability-diversity tradeoff

- Building blocks of the computational model
  1. Python boilerplate: Module imports, helper functions etc.
  2. Main classes: Environment, Agents, Groups, Population
  3. Parameter set
  4. Simulation “experiment”
     1. Sweeping across parameter values
     2. Sufficient repeats to distinguish pattern from noise
  5. Visualization and analysis of outputs
- Each notebook cell ends with a tests section
  - Making sure the components function as expected

Homework for tomorrow

1. Read carefully the article: Axelrod, R. (1997). The dissemination of culture
2. Write down the assumptions constituting the conceptual model outlined in the paper
   - Environment
   - Agents and their properties
   - Rules of interaction
   - Dynamics (= what happens when)
   - …
3. What are the central findings
4. and possible ways of extending the model, according to Axelrod?