![JASSS logo]

Nicolas Payette (2011)

For an Integrated Approach to Agent-Based Modeling of Science

Abstract

The goal of this paper is to provide a sketch of what an agent-based model of the scientific process could be. It is argued that such a model should be constructed with normative claims in mind: i.e. that it should be useful for scientific policy making. In our tentative model, agents are researchers producing ideas that are points on an epistemic landscape. We are interested in our agents finding the best possible ideas. Our agents are interested in acquiring credit from their peers, which they can do by writing papers that are going to get cited by other scientists. They can also share their ideas with collaborators and students, which will help them eventually get cited. The model is designed to answer questions about the effect that different possible behaviors have on both the individual scientists and the scientific community as a whole.

Keywords:
Agent-Based Models, Science Dynamics, Social Networks, Scientometrics, Evolutionary Computation

Introduction

1.1   Let us start with the "science system" as a whole. How would you describe it? A first approximation might be something like "researchers interacting with nature and with each other in order to produce knowledge." You might have a few qualms about that definition. "Knowledge," for instance, is a very loaded word—especially if you are a philosopher. "In order to" might also sound a little bit teleological, and one should be careful about ascribing goal-directed behavior to a system. But, these caveats notwithstanding, I propose that we set out for the following task: unpack this definition into a formal model. And by "formal model," I really mean "computer program": an agent-based simulation that we will be able to play around with. What do we stand to gain by doing that? Well, besides the fact that you get a better understanding of something by building a model of it, what I ultimately have in mind is policy-making, or at least *normative claims*. Knowledge production (whatever that may be) is certainly something that we should try to maximize, and having a good model of how the scientific system works would allow us to test the effect of potential policies before actually applying them. Though many isolated aspects of the science system have already been studied using ABMs (see Payette 2012 for a review) what we don't have yet is an integrated model that would provide us with a common structure in which to investigate a wide range of questions. This paper is obviously not the place to fully flesh one out, but what we can do here is try to come up with a sketch of what such a model could be or, as Claudio Cioffi-Revilla (2010) would say, introduce the "cast of principal characters."

Researchers as Agents

2.1   Agents, of course, will be our first order of business. In our tentative definition, we referred to them as "researchers," but that is fairly broad. If we limit ourselves to academia (which I plan to do), that leaves us with professors, students and maybe other research personnel (a paid MRI technician, for instance). Should we consider them all? It depends, of course, of what we want our model to track. I have already said that the ultimate goal of ABMs of science should be to find ways of optimizing "knowledge production," but in order to do that, we need a way to measure it. Measuring knowledge production is not just a problem for ABMs of science: it is also a problem in the real world, and the way we do it, for better or for worse, is to measure the scientific papers output. So, leaving aside for now the question of how to assess the quality of scientific papers, we at least know that our agents are going to have to produce some[1]. This will not only allow us to measure the performance of our system under different conditions, but also to calibrate our model against actual scientometric data—a very desirable feature. So, coming back to the

question of who our agents should be, one possible answer is: those who actually write papers—which means mostly professors and graduate students (as opposed to undergrads). Of course, researchers do not operate in a vacuum: aside from the aforementioned technicians, one could argue that journal editors, administrative personnel or even politicians play an important role. And that is just the humans: if one is prepared to consider institutions as agents, one could add departments, faculties, universities, funding agencies, equipment suppliers, countries, etc. I have no quarrel with that, but we do need to limit the scope of our model somehow, so I suggest we focus on the production of scientific papers and consider adding other agents only insofar as they are directly involved in that process.

2.2   So now that we have our agents, what is it they *do*, aside from writing papers? Well, if you consider that papers are the output of the system, you probably need to ask what its input is. Going back, again, to our tentative definition, we said that researchers interact "with nature and with each other." Let us take each of these in turn, starting with "nature."

🌍 The Epistemic Landscape

3.1   Our agents need to be situated in some sort of environment about which they are going to acquire information[2] Anyone who has read *JASSS* has seen countless models where agents are located on a grid where they are trying to harvest resources and/or avoid predators, etc. This is probably not exactly what we need (though there are exceptions, scientists nowadays tend to stay in their labs) but the spatial metaphor can still be useful. Let us suppose that researchers are trying to come up with ideas[3] about how the world works. These ideas can easily be represented by points in a multidimensional space. If you are not convinced, think about Borges' (1944) Library of Babel :[4] a fictional library that contains the set of all possible books of an arbitrary maximum length written with a fixed set of characters. Using ASCII (which is obviously not what Borges had in mind), the string "E=mc^2" occupies the position (69, 61, 109, 99, 178) in the space of possible five-character long ideas. Now think of the maximum length of a paper in *Nature*, and we are probably in business.[5] I have used the Library of Babel as an illustration, but you should think of epistemic space as a conceptual space more than a syntactic space, meaning that: "energy is equal to mass multiplied by the square of the speed of light" is very close to "E=mc^2" in epistemic space, even though the characters used are very different (same goes for equivalent propositions in different languages).

3.2   Now, as each of us knows, not all ideas are created equals. In other words, the space of possible ideas is not flat: it needs one more dimension, one that represents the "objective value"[6] associated with an idea. If you think about that dimension as "height," we now have what is called an *epistemic landscape*. Our agents are trying to find the peaks in that landscape, having to cross valleys of low value and to avoid getting stuck on local maxima.

3.3   If we allow each agent to entertain multiple ideas at once[7], it is not the researchers themselves that will be moving around the landscape: it is the population of ideas that will spread out across it. But it does not mean the researchers are not active: they are the ones who create, test, modify, exchange, and ultimately read and write papers about ideas. How do they do that? Let us take each of these actions in turn. The "exchange" action will actually take us to the "interact with each other" part of our initial definition. And we will save the "paper writing and reading" actions for last.

3.4   The creation of completely new ideas, at the start of a simulation, should be the simplest process of all: just pick points at random locations in space. Agents, at first, will not know the objective value of their ideas. The best they can do is assign to each, at random, a subjective value—a guess, if you like. What they are interested in, however, is the objective value of their ideas, and the way to learn about these values is to test the ideas. But testing is a time consuming process, and not all ideas can be tested at once: the agents need a way to decide which ideas they are going to test at each time step of the simulation. Maybe it is the ideas that look the most promising, i.e., those that have the highest subjective value. In any case, a test of an idea will allow the agent to obtain an approximation of the objective value of the idea, and the agent will adjust its subjective evaluation accordingly.

3.5   This is a good place to say a few words about the evaluation of the system's performance. Remember that we want to use the simulation to evaluate the effect of scientific norms and institutional policies. Now, given what we just said on the subjective and objective values of ideas, there are two questions we can ask about our researchers:

- Have they found the good ideas? We can track the objective values of the ideas the scientists have found so far (look at the mean, the max, etc.) and how long it takes the community to get to the best ideas.
- How good is their evaluation of the ideas they have? That is simply the difference between their subjective evaluations of the ideas and their "real," objective value. The smaller this difference, the closer the scientific community is to a "true picture of the world".

3.6   But how is it that the agents can achieve these objectives? All that we have seen so far is that they start by generating random ideas and testing some of them, a process which, by itself, would not lead to any improvement. The key to improvement is that researchers are periodically allowed to generate new ideas, and that these new ideas are generated by modifying the best ones they have so far, thus exploring the regions of the space that look the most promising. They modify a previous idea by applying a certain amount of "noise" to it: take each component of the idea's coordinates and randomly nudge it in one direction or another. They can then drop the worst ideas they have and keep the best ones. If the epistemic landscape is smooth enough, that should lead them slowly to areas of better value.[8]

Science as a Social Process

4.1 What we have just described is the individual process of exploration but, as we have previously acknowledged, science is a social process: researchers interact, and that interaction plays a significant role in the search for better ideas. To describe these interactions, I am going to introduce a notion that I think plays a significant role: credit. You could also call it "peer recognition," or something else, but since it is a notion I get from David Hull (1988), I am going to use his term for it. There is a fairly big assumption here, and it is that *credit* is what individual scientists are after. More than truth? Well, yes, though coming up with theories closer to the truth is a pretty good way of getting credit. (And credit, by the way, can eventually be converted to money: good jobs, grants, etc.) So, how do you get credit? In the context of the model: by writing papers. Or, to be more precise, by having other researchers reading and citing the papers that you wrote. And the more credit they already attribute to you, the more likely they are to read new papers you write. So there is something like a positive feedback loop going on here. There is more to be said about the production (and reading) of papers, but we will come back to it in a moment. Before that, I want to talk about two slightly more direct forms of idea exchange: education and collaboration.

4.2 There is a turnover in science. Eventually, researchers retire—some after a long and fruitful career, some because they just can't get tenure and finally give up. The ones that are active at any given time get the huge privilege of being allowed to fill the minds of their students with their own ideas. (At least, in our model, that's how it works.) Your students are vectors for your ideas: the more of them you attract, the better the chances that they will spread your own theories, and thus allow you to get more credit, which will help you attract even more students, and so on.

4.3 Collaboration is a similar process, except that it goes both ways: not only do you get to share your ideas with your collaborators, they get to share theirs with you[9]. And credit also plays a role: the more credit you have, to more likely you are to attract lots of collaborators. Assuming that links with other researchers can be created and dropped, what you get is a dynamic social network of collaboration. That network can be analyzed. An obvious question to ask is whether the structure of the generated network corresponds to the social structure usually found in the real world. Another question is whether the features of the network have any influence over the performance of the system. For example, previous research using ABMs (Zollman 2007; Grim 2009) tends to show that too dense a network of communication between researchers can cause a system to settle early on a local maximum.

Bibliometric Aspects

5.1 We now, finally, turn to the reading and writing of papers. The structure of a paper in our model is very simple: the author proposes that some particular idea has some particular value, and gives a list of references to support that conclusion. The readers of the paper then attribute credit to both the author of the paper and to the authors of the papers that are cited in it. Exactly how much credit depends on the reader's evaluation of the quality of the paper: i.e., how much the reader agrees with the proposed idea and the value stated for it. We will not go into details of the process of idea evaluation, but we can say that, as a general rule, an idea that is similar to an already well-regarded idea will be well received, and vice versa. This could easily be implemented via a rule-based approach or an artificial neural network.

5.2 Now, when writing a new paper, an author has an important question to answer: who to cite? Multiple factors have to be taken into account: e.g., the cited papers should actually be related to what the current paper is about, claiming support from already accepted theories in your field is important, self-citation is an often-used strategy, etc. A somewhat lesser known factor is what Hull calls "conceptual inclusive fitness." This is an analogy from evolutionary biology, where "inclusive fitness" refers to the fact that altruistic behavior towards one's close relatives is advantageous from a gene's eye view: it promotes the replication of those genes that are shared. Similarly, from an idea-transmission standpoint, helping your students or collaborators—with whom you've shared lots of ideas—is advantageous, and our model should take that into account.

5.3 We could also complexify the model further by adding co-authorship, and since co-author networks are an important feature of real-life science, it would probably be a good idea to do so. Börner, Maru and Goldstone (2004) have explored the creation of co-authorship networks in a model they call TARL (for topics, aging, and recursive linking), but in their model, coauthors are randomly assigned to one-another when a paper is produced. In our model, we presumably would have to take credit, collaboration links, number of shared ideas, etc. into account. A decision would also have to be made about which idea to present in the paper: a previously existing idea from one of the authors, or a new idea that they generate together? Same goes for references: how does a group of coauthors decide who to cite? I don't have exact answers for these questions here, but they do illustrate a point: agent-based modeling, when done at a sufficiently low level, forces the modeler to ask questions that could otherwise be abstracted away (though one has to be careful when trading away simplicity for completeness).

5.4 Let us briefly recapitulate what we have described so far. Our agents are researchers, producing ideas that are points on an epistemic landscape. We are interested in our agents finding the best possible ideas. Our agents are interested in acquiring credit from their peers, which they can do by writing papers that are going to get cited by other scientists. They can also share their ideas with collaborators and students, which will help them eventually get cited. If we think in terms of networks, there are a lot of them present here: a collaboration network, a citation network, and potentially a co-authorship network. We also have a student-supervisor hierarchy network, and also an idea hierarchy network (given that each new idea has a parent idea). All those

networks can be analyzed and compared to the corresponding "real world" networks (though in the case of the idea network, this is more difficult and requires a good deal of interpretation.)

Conclusion

6.1 It will come as no surprise to the reader that I am currently working on a model that fits very closely with what I have been describing in this paper. Still, despite the fact that I had a specific implementation in mind while describing this model, I do believe that the very high level overview I gave here captures some of the features that an agent-based model of science should have: researchers, organized in various networks, trying to find out about nature while exchanging information about it either through direct contact or through publication. Of course, the details may vary. You could choose to work with some other form of representation of the "ideas" of the scientists (symbolic formulas); the cognitive architecture of the scientists could be a lot more elaborate (artificial neural networks are an obvious option), etc.

6.2 Given the complexity of the overall system, one could argue that would it be better approached by a succession of local models. Is it feasible to capture all those properties in one encompassing simulation? What would such a complex model actually tell us? I expect that what it will come down to in the end is predictive power. If we have a model that allows for a good approximation of the effect of complex science policies that are hard to test with simpler models, then we will have something very valuable. But we won't know until we build it.

Notes

[1] Nigel Gilbert (1997), in what is usually considered to be the first agent-based model of science, went as far as having the papers themselves be the main agents in his simulation and have them "choose" their authors amongst a pool of available researchers.

[2] This is not always the case, however. In Gilbert (1997), for example, papers produced by the agents are different from one another, but are not *about* anything. Another example is Edmonds (2007), where agents are trying to produce logical formulas with no empirical content.

[3] "Idea" is an intentionally vague term. It could be taken to stand for theories, beliefs, models, propositions, research strategies, experimental parameters, etc.

[4] Dennett (1995, 107-108) also uses this example to illustrate the space of all possible genetic codes—something that, for reasons I will not go into here, is not unrelated to our current discussion.

[5] Computational resources being what they are, an actual simulation would be limited to a much smaller number of dimensions, however.

[6] Or "quality," "significance," "empirical adequacy," "degree of truth," etc.

[7] Some existing models (e.g., Weisberg and Muldoon 2009) have their agents holding only one position at a time: the position associated with the agent's current approach.

[8] Some readers might have noticed the similarity between this process and that of "gradient ascent" used in evolutionary computation. This is not a coincidence: this later field is a major inspiration for the present model. See Luke (2010) for details on evolutionary computation.

[9] Of course, students also share ideas with you. To account for that fact, we will assume that, after their education is completed, they have a very good chance of becoming your collaborators.

References

BÖRNER, K., Maru, J. T. & Goldstone, R. L. (2004). "The simultaneous evolution of author and paper networks." *Proceedings of the National Academy of Sciences of the United States of America* 101 Suppl 1 (April 6), 5266-5273. [doi:10.1073/pnas.0307625100]

BORGES, J. L. (1944). *Ficciones*. Argentina: Editorial Sur.

CIOFFI-REVILLA, C. (2010). A methodology for complex social simulations. *Journal of Artificial Societies and Social Simulation*, 13(1), 7 http://jasss.soc.surrey.ac.uk/13/1/7.html [doi:10.2139/ssrn.2291156]

DENNETT, D. C. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. Simon & Schuster.

EDMONDS, B. (2007). Artificial Science: A Simulation to Study the Social Processes of Science. In B. Edmonds, K. G. Troitzsch & C. H. Iglesias (Eds.), *Social Simulation: Technologies, Advances and New Discoveries* (pp. 61-67). IGI Global. [doi:10.4018/978-1-59904-522-1.ch005]

GILBERT, N. (1997). A simulation of the structure of academic science. *Sociological Research Online 2*(2). http://www.socresonline.org.uk/2/2/3.html. [doi:10.5153/sro.85]

GRIM, P. (2009). Threshold Phenomena in Epistemic Networks. In *Complex Adaptive Systems and the Threshold Effect: Views from the Natural and Social Sciences*. http://www.aaai.org/ocs/index.php/FSS/FSS09/paper/viewPaper/916.

HULL, D. L. (1988). *Science as a process: an evolutionary account of the social and conceptual development of science* University of Chicago Press. [doi:10.7208/chicago/9780226360492.001.0001]

LUKE, S. (2011). *Essentials of Metaheuristics*. Lulu. http://cs.gmu.edu/~sean/book/metaheuristics/. Archived at: http://www.webcitation.org/5zmtyFdML.

PAYETTE, N. (2012). Agent-Based Models of Science. In A. Scharnhorst, K. Börner & P. van den Besselaar (Eds.), *Models of Science Dynamics*, Complexity Series. Springer. [doi:10.1007/978-3-642-23068-4_4]

WEISBERG, M., & Muldoon, R. (2009). Epistemic landscapes and the division of cognitive labor. *Philosophy of Science, 76*(2). 225-252. [doi:10.1086/644786]

ZOLLMAN, K. J. S. (2007). The communication structure of epistemic communities. *Philosophy of Science 74*(5). 574-587. [doi:10.1086/525605]