ABMs in Philosophy of Science

Conor Mayo-Wilson

Models and Simulations in Philosophy
June 11th, 2013

Kitcher's and Strevens' Models

1. Strevens' models are very general and need not be interpreted as describing scientific practice. Are there other social practices that institute a reward system similar to the priority rule? Are the benefits in such cases "winner-confers-all" as well?
2. Who are the decision-makers in Kitcher's and Strevens' model?
3. Are Kitcher and Strevens' models "agent-based"?
   - In what ways are the decision-makers in their models similar to and different from the ABMs we have studied concerning the evolution of norms?
   - In general, in what other ways are their models similar to and different from the ABMs we have studied concerning the evolution of norms?

Today: An ABM of scientific communities developed by Weisberg and Muldoon [2009]:

- Boundedly-rational agents:
  - Finite memories.
  - Limited action set on each play.
  - Updating behavior is simple and algorithmic.
- Interactions among agents determine payoffs:
  - Primarily: By causing particular sections of the landscape to be explored.

Outline

1 Review and Discussion Questions

Outline

1. Review and Discussion Questions
2. The Model
   - Goals
   - The Landscape
   - The Scientists
   - Evaluation of the Model

Outline

1. Review and Discussion Questions
2. The Model
   - Goals
   - The Landscape
   - The Scientists
   - Evaluation of the Model
3. Results
   - Criteria for Success
   - Evaluation of Exploration Strategies

Outline

1. Review and Discussion Questions
2. The Model
   - Goals
   - The Landscape
   - The Scientists
   - Evaluation of the Model
3. Results
   - Criteria for Success
   - Evaluation of Exploration Strategies
4. References

What do Weisberg and Muldoon hope to explain with their model?

Goal

What do Weisberg and Muldoon hope to explain with their model?

- Why diversity of research methodology is beneficial
- Why risk-taking can be good for a scientific community as a whole

We will argue that to be maximally effective, scientists need to really divide their cognitive labor, coordinating in such a way to take account of what other scientists are doing. We also show, albeit in a preliminary way, that a mixed strategy where some scientists are very conservative and others quite risk taking, leads to the maximum amount of epistemic progress in the scientific community.

Weisberg and Muldoon [2009], pp. 3.

What's New?

Didn't Kitcher and Strevens already reach the same conclusion?
What's new?

First, by using an ABM, Weisberg and Muldoon are checking the robustness of Kitcher and Strevens' results by relaxing assumptions.

In particular, they relax assumptions about

- Agents' rationality
- Agents' knowledge

Second, and more importantly, Weisberg and Muldoon's model is intended to represent a different kind of scientific practice than is modeled by Kitcher and Strevens . . .

What's New?

In one kind of scenario, scientists choose between different approaches, all of which aim at the same narrow goal . . .

Another type of scenario in which scientists divide their cognitive labor involves research on the same topic broadly construed, but with small differences in the activities and goals of particular scientists. For example, within the research program of synthetic biology (Benner, 2003), a group of chemists successfully synthesized novel DNA nucleotides that function analogously to naturally occurring DNA bases. This initial synthesis by one group of scientists (Liu et al., 2003) led another to incorporate these bases in to a strand of DNA, creating what they called xDNA (Gao, Liu, & Kool, 2005). . . . These individual episodes of research were independent, but they built off of one another . . . [U]nlike in Watson and Cricks elucidation of DNAs structure, a significant discovery made by one did not signal the end of the specific research topic.

Different Target System

So, unlike Kitcher and Strevens, Weisberg and Muldoon aim to model scientific inquiry when:

- One discovery does not end a research program.
- Researchers build on previous results.
- Scientists employ different techniques within the same field, broadly construed.
- Different techniques have different value.
- Scientists may have different goals.

You tell me! Explain the **informal** interpretation of the following terms:

- Approaches
- Significance
- Epistemic Landscape

In Weisberg and Muldoon [2009], an approach is intended to represent:

- The research questions being investigated
- The instruments and techniques used to gather data
- The methods used to analyze the data
- The background theories used to interpret the data

An approach is pretty complicated!

The remaining terms in the model can be interpreted in any number of ways . . .

"Significance", as the name suggests, represents the value of some research approach to a problem.

But that value could be either **scientific** or **personal**:

- Scientific value - Concerning truth, fruitfulness of an approach, etc.
- Personal - Publications, grants, graduate student workers, etc.

Basic Terms

An epistemic landscape specifies the significance/value of each approach.

You tell me! Explain the **formal** interpretation of the following terms:

- Approaches
- Significance
- Epistemic Landscape

Basic Terms

- Approaches - A point in the plane $\mathbb{R}^2$
- Significance - A non-negative real number.
- Epistemic Landscape - A function from $\mathbb{R}^2$ to $\mathbb{R}^{\geq 0}$.

Let's reflect on these formal definitions for just a bit ...

Consider significance first.

Question: Why should significance be a real number?

Significance and Utility

Research might be significant in any number of ways ...

- It might provide evidence for some hypothesis
- It might rule out a hypothesis
- It might simplify or unify existing theories/hypotheses
- It might suggest a new experimental technique
- It might create valuable technology (e.g. for medicine)
- Etc.

Question: Why should these values be capable of being compared, let alone numerically quantified?

At the outset of the class, you read Alexander [2007]'s discussion of representation theorems that show that,

- If your preferences satisfy certain plausible axioms of rationality,
- Then we can treat you as if your assigned numerical utilities to various outcomes.

At the outset of the class, you read Alexander [2007]'s discussion of representation theorems that show that,

- If your preferences satisfy certain plausible axioms of rationality,
- Then we can treat you as if your assigned numerical utilities to various outcomes.

This is why numerical utilities have appeared in all the games we've discussed thus far in the class.

Weisberg and Muldoon might appeal to similar theorems here.

Alternatively, they might restrict the interpretation of significance to some collection of values that are comparable and quantifiable.

Numerical Approaches

Similarly, why are approaches represented by pairs of numbers?

Recall, an approach is intended to represent:

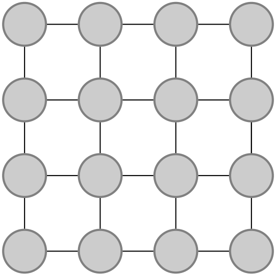
- The research questions being investigated
- The instruments and techniques used to gather data
- The methods used to analyze the data
- The background theories used to interpret the data

Features of Approaches

Given the interpretation, it is reasonable to assume that some approaches are more similar than others.

- Research questions can be more and less similar.
- Measurement instruments can be more and less similar.
- Statistical methods to analyze data can be more and less similar.
- Background theories to interpret data can be more and less similar.

Define an approach network in which two approaches are connected by an edge precisely if they are similar to a specified degree.



What Weisberg and Muldoon, in effect, assume is that the approach network is a lattice network.

Weisberg and Muldoon never assume that approaches have any other features of pairs of real numbers:

- Approaches are never added, multiplied, etc.
- Approaches are never compared (though their significances are).

Of course, one might wonder why Weisberg and Muldoon [2009] assume that approaches are arranged in a two- dimensional lattice network, rather than some other type of networks . . .

They offer a few different reasons a footnote:

- Computational simplicity
- "Conceptual clarity"
- Non-arbitrariness - What other types of assumptions would make the model more realistic?

Approaches as Lattice Networks

> *Our primary motivation for adopting the three-dimensional landscape was conceptual clarity and computational simplicity. More complex landscapes can be generated easily and they often yield additional local maxima. However, without making more specific real-world commitments about what the topography of a particular landscape represents, we believe that the prudent course is to keep the landscapes simple. Future investigations could profitably explore landscapes of higher dimensionality and greater ruggedness.*

Weisberg and Muldoon [2009], pp. 7.

Two quick comments:

- Weisberg and Muldoon don't disagree that the model could be checked for robustness by considering other types of approach networks.
    - Lattice networks are most applicable in representing spatial proximity.

Two quick comments:

- Weisberg and Muldoon don't disagree that the model could be checked for robustness by considering other types of approach networks.
  - Lattice networks are most applicable in representing spatial proximity.
- The approach network is the same for all agents.
  - In contrast, see Hong and Page [2004] and Hong and Page [2001], who model agents that might see different approaches as being available at any given time.

What assumptions should we make about the epistemic landscape?

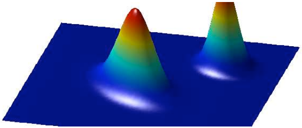
For example, should we assume that

- Neighboring approaches have similar significances?
- There is a uniquely best approach?
- Few approaches have high significance?
- Etc.

Weisberg and Muldoon [2009] assume that

- Neighboring approaches have similar significances.
- There are multiple approaches that are best.
- Few approaches have high significance.

- For simplicity, Weisberg and Muldoon [2009] choose (multi-variate) normal distributions to represent the distribution of significance over various approaches.
- They assume each landscape has two peaks.



Why does this meet the three criteria?

- Neighboring approaches have similar significances
  - as normal distributions are continuous.
- There are multiple approaches that are best
  - as there are two peaks.
- Few approaches have high significance
  - as the peaks are narrow (i.e., small variance), and so most of the landscape has zero significance

You tell me! Describe three strategies for exploring the landscape.

You tell me! Describe three strategies for exploring the landscape.

- Simple hill climbing
- Followers
- Mavericks

Strategies

What are some common features of these methods? In particular, what do they are share in common with respect to

- Memory?
- How far ahead they plan?
- How they consider others' results?

Evaluation of the Model

As specified, the model seems to meet most of Weisberg and Muldoon's criteria:

- One discovery does not end a research program. ✓
- Researchers build on previous results. ✓
- Scientists employ different techniques within the same field, broadly construed. ✓
- Different techniques have different value. ✓
- Scientists may have different goals. ✗

Outline

1. Review and Discussion Questions
2. The Model
   - Goals
   - The Landscape
   - The Scientists
   - Evaluation of the Model
3. Results
   - Criteria for Success
   - Evaluation of Exploration Strategies
4. References

Success

How do Weisberg and Muldoon [2009] evaluate the success of the the strategies?

How do Weisberg and Muldoon [2009] evaluate the success of the the strategies? Three criteria for success:

- Does the community have at least on scientist who finds each peak?
- If so, how long do they take to find the peak?
- **Epistemic Progress:** What percentage of the approaches with non-zero significance are explored?

Other Criteria of Success

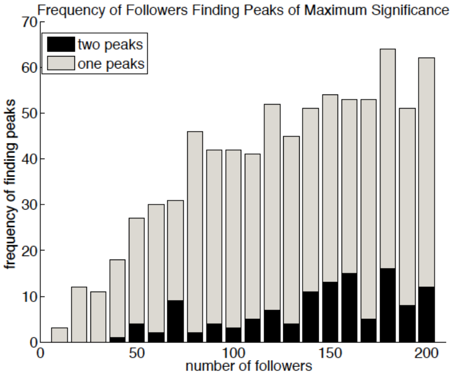
Note, there are other obvious criteria that one could also use:

- Average significance of approaches pursued by scientists
- Minimum significance

Success Criterion 1: Followers often fail to find one or more peaks.

Success Criterion 1: Followers often fail to find one or more peaks.

Figure 5: Frequency of Convergence on Approaches of Maximum Significance



Success Criterion 2: They often also over a small portion of the landscape: they explore a small section, and then stop.

Success Criterion 2: They often also over a small portion of the landscape: they explore a small section, and then stop.

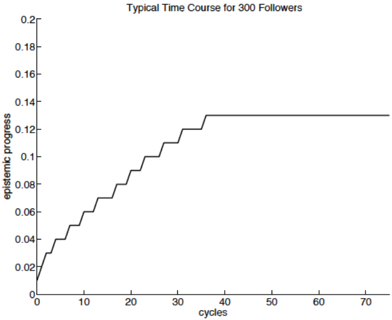


Figure 6: Typical time-course for followers on a 2 Gaussian epistemic landscape. 300 followers.

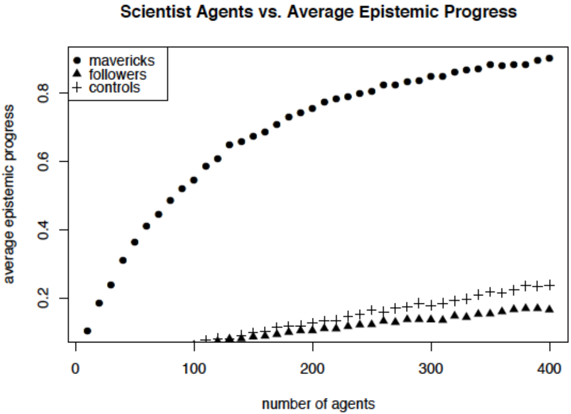
This may not surprise you because Weisberg and Muldoon prepare you up for this conclusion, but . . .

Following is a type of imitation. And think about how well imitation rules have performed in the game-theoretic models that we've studied thus far.

Benefits of being a Maverick

Success Criterion 1: In contrast, Mavericks find the peaks nearly all the time and, (though perhaps unsurprisingly),

Success Criterion 2: They explore considerably more of the landscape.



Why might this be surprising?

Again, consider what would happen in the game-theoretic models we studied if agents tried to behave in a way that deviates from that of their neighbors in a network.

The Perks of Being a Mav Follower

- In "mixed" populations of both Mavericks and Followers, the addition of one maverick can substantially improve the performance of the remaining followers.

The Perks of Being a Mav Follower

- In "mixed" populations of both Mavericks and Followers, the addition of one maverick can substantially improve the performance of the remaining followers.
- Why? Mavericks explore areas of the epistemic landscape that followers can subsequently explore.

Recall, Weisberg and Muldoon hoped to explain the following:

Meeting the Models Goals

Recall, Weisberg and Muldoon hoped to explain the following:

- Why diversity of research methodology is beneficial
- Why risk-taking can be good for a scientific community as a whole

The superior performance of the mavericks, who might face substantial personal costs in real life, show the benefits of risk-taking.

Meeting the Models Goals

What about diversity?

Here, the results seem a bit more mixed.

Meeting the Models Goals

- If diversity is understood with respect to approaches, the model seems to explain the benefits of diversity:
    - Followers don't diversify enough, and so they get locked in particular parts of the landscape. Thus, they don't find the peaks.
    - In contrast, Mavericks explore a good chunk of the landscape. So they do find peaks.

Meeting the Models Goals

- If diversity is understood with respect to approaches, the model seems to explain the benefits of diversity:
  - Followers don't diversify enough, and so they get locked in particular parts of the landscape. Thus, they don't find the peaks.
  - In contrast, Mavericks explore a good chunk of the landscape. So they do find peaks.
- If diversity is understood with respect to exploration methods (e.g. maverick, follower, hill-climber), then the model actually predicts homogeneity is better.
  - Switching an agent from a follower to a maverick always raises the probability of finding the peaks and increased epistemic progress.

References I

Alexander, J. M. (2007). *The structural evolution of morality*. Cambridge University Press Cambridge.

Hong, L. and Page, S. E. (2001). Problem solving by heterogeneous agents. *Journal of Economic Theory*, 97(1):123—163.

Hong, L. and Page, S. E. (2004). 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.

Weisberg, M. and Muldoon, R. (2009). Epistemic landscapes and the division of cognitive labor. *Philosophy of Science*, 76(2):225—252.