# ABMs in Philosophy of Science

Conor Mayo-Wilson

Models and Simulations in Epistemology and Philosophy of Science
June 16th, 2014

## Review

First Unit: Is rational disagreement possible?

## Review


We asked this question from the perspective of an individual:

Can **I** be rational and yet disagree with a peer?

## Review

First Unit: Can **I** be rational and yet disagree with a peer?

Second Unit: Is disagreement good **for the scientific community**?

## Kitcher's Model

Kitcher:

- Individual and collective rationality may diverge.
- Diversity of paradigm (or theory adoption) may raise community's chance of success.

[Kitcher, 1990]

## Longino

Longino: Diversity of research methodology is characterized by different standards for explanation:

- What can be considered a cause
- How variables are measured
- Experimental techniques to establish causation
- Statistical techniques for analysis

[Longino, 2006]

## Longino

Longino: Diversity is valuable because different explanations possess different virtues and limitations.

[Longino, 2006]

## Wylie

Wylie: Diversity of research methodologies is characterized by

- Different research questions: What is in need of explanation?
- Different "background" theories for interpreting data.
  - Different background theories can arise from differences in gender, race, class, ethnicity, sexual orientation, and more.

[Wylie, 2012]

## Wylie

Wylie: Diversity of research methodology is valuable because

- Different questions ⇒ Obviously important
- Different "background" theories ⇒
  - Question dominant assumptions
  - Inversion Thesis: Some who are disadvantaged are in a better position to know certain facts.

## Kitcher, Longino, and Wylie

Different benefits of diversity are extolled by these authors.

- Kitcher: When at least two theories are in competition, diversity may raise community's collective chance of success.
- Longino and Wylie: Diversity of approaches might yield results of different value.

Question: Is there a way to model the virtues of diversity extolled by Longino and Wylie?

## ABMs of Scientific Communities

Today: Two ABM of scientific communities developed by Weisberg and Muldoon [2009] and Hong and Page [2004]:

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

## Outline



## 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

## Goal

> *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?

## 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

## What's New?

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.

## Basic Terms

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

- Approaches
- Significance
- Epistemic Landscape

## Basic Terms

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!

## Basic Terms

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

## Basic Terms

"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.

## Basic Terms

**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 $\mathbb{R}^{\geq 0}$.

## Interpreting the Basic Terms

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

Consider significance first.

## Significance and Utility

**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.

## Significance and Utility

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

## Significance and Utility

In economics and in philosophy, researchers often appeal to 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.

Here, your preferences might be over different benefits conferred by different types of theories.

## Significance and Utility

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?

## Approaches

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

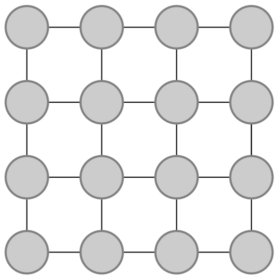
Given the interpretation, it is reasonable to assume that some approaches are more similar than others.

- 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.

## Approaches as Networks

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

## Basic Terms



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

## Approaches as Networks

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).

## Approaches as Networks

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.

More on this later ...

## Approaches as Lattice 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.

## Generalizing

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.
  - More on this later . . .

## The Landscape

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.

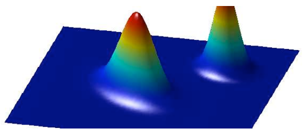
## The Landscape

Weisberg and Muldoon [2009] assume that

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

## The Landscape

- 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.



## The Landscape

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

## Strategies

**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



## Success

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

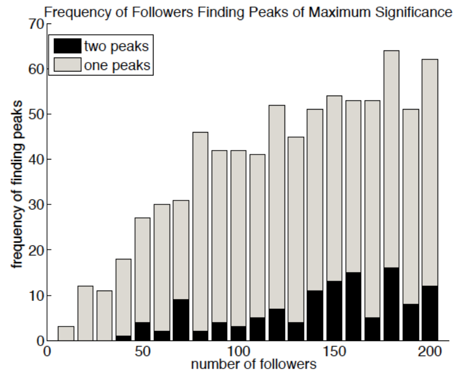
Note, there are other obvious criteria that one could also use:

- Average significance of approaches pursued by scientists
- Minimum significance

# The Perils of Following

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

Figure 5: Frequency of Convergence on Approaches of Maximum Significance




# The Perils of Following

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

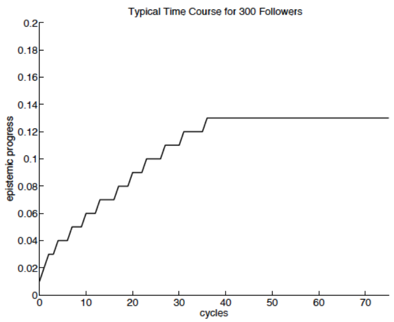



Figure 6: Typical time-course for followers on a 2 Gaussian epistemic landscape. 300 followers.

# Following and Imitation

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

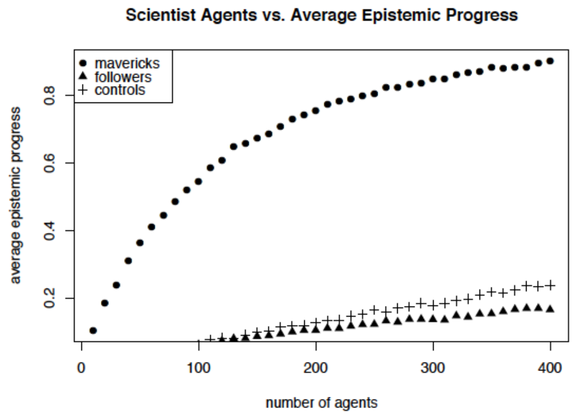
Following is a type of imitation, and imitation rules perform well in many the game-theoretic models.

## Benefits of being a Maverick

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

## Benefits of being a Maverick

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




## Mavericks and Anti-Imitation

Why might this be surprising?

Imagine what would happen if you always tried to behave in a way that deviates from your social practice.

## 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.

## 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

## Meeting the Models Goals

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

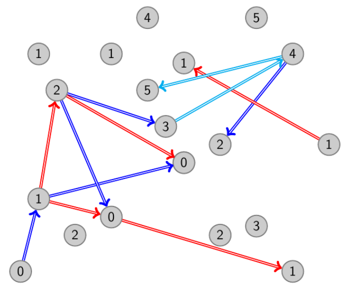
- If diversity is understood with respect to approaches, the model seems to explain the benefits of diversity:
  - Followers don't diversify enough, and so they get locked in particular parts of the landscape. Thus, they don't find the peaks.
  - In contrast, Mavericks explore a good chunk of the landscape.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

# Research Approaches

Weisberg and Muldoon [2009] assume

- Research approaches = Lattice network
- The approach network is the same for all agents.

# Dropping Assumptions

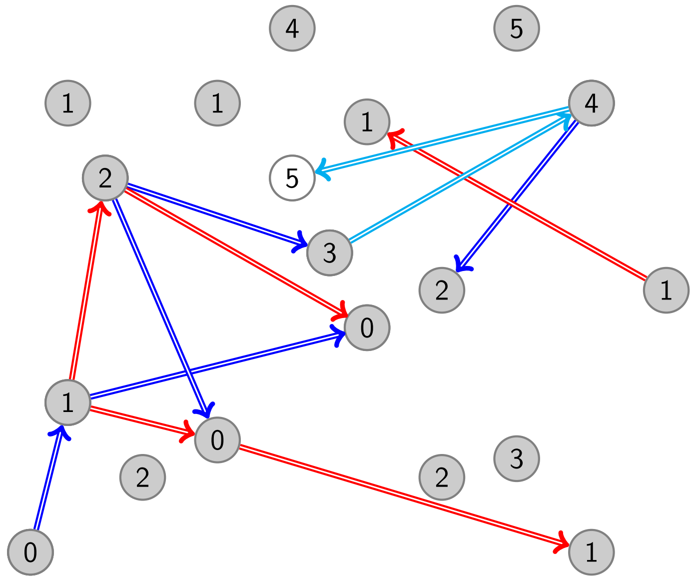


Hong and Page [2004] drop both assumptions:

- Different Perspectives: Even if they currently adopt identical research approaches, two researchers may nonetheless 'see' different possibilities for the next approach.
- No common structure to all perspectives.

Here's an example of how the model might work . . .

# The Model



## Framework

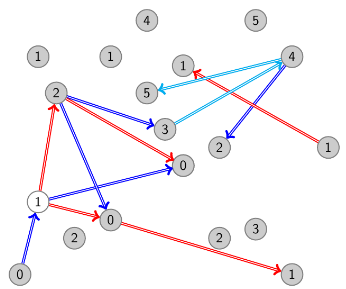
- A set of solutions $X$
  - The nodes in the previous graph
- Each solution $x$ has some value $V(x)$
  - The numbers on the nodes.
- A set $\Phi$ of problem-solving approaches
  - An element of $\Phi$ is simply all edges of a single color in the previous graph.

## Framework

Hence, given any initial starting point $x_0$ and approach $\varphi$, there is some final solution that is found by a problem-solver employing $\varphi$.

Call this state $x(\varphi)$.

## Stopping State




If the dark blue, problem-solving approach started in the bottom left node (labeled zero), then it would reach the highlighted final solution with value one.

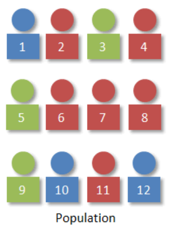
## Framework

- Each solution $x$ has some non-zero probability $\nu(x)$ of being the first one tried by the group.
  - The starting node in the graph
- So there is some expected value of the final solutions found by any given approach $\varphi$

$$\mathbb{E}_\nu[\varphi] := \sum_{x \in X} V(x(\varphi)) \cdot \nu(x)$$

- The best approach is the one maximizing this expected value.

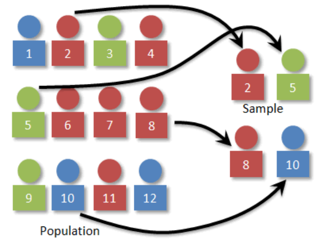
## Framework




The General Population: Hong and Page assume that there is some large population of agents of size $N$.

- For each problem-solving approach $\varphi$, there is some probability $\mu(\varphi)$ that any given agent employs it.

## Framework




Selecting a Team: Given the large general population $N$, Hong and Page consider two ways of picking a team of $N_1 < N$ many members:

- Randomly
- Select the $N_1$ agents with the approach yielding the highest expected value.

## Assumptions

- **Difficulty:** No approach always finds the best solution, regardless of where it starts.
- **Diversity:** For every non-optimal solution, there is some approach that "sees" a better solution.
- **Uniqueness:** The best approach is unique.

## Assumptions

How does a group with different approaches works together?

- Hong and Page [2004] claim: "The proof of the lemma that follows does not depend on what the specific model is."
- That's not quite right. They make one major assumption, which I'll call
- **Non-Stopping:** A group stops only if and only if no one sees a better solution.

**Theorem (Hong and Page [2004])**

*With probability one, there is a large enough population $N$ such that if one wants to form a team of size $N_1 < N$, then the a randomly formed team's expected performance is higher than that composed of the best problem solvers.*

**Proof Sketch:**

- By **Diversity**, for each non-optimal solution, there is an approach that "sees" a better one.
  - So given a non-optimal solution $x$, let $\varphi(x)$ be the solution that won't "get stuck" at $x$
- Pick a very large team size $N_1$.
- Then in a randomly formed team of size $N_1$, each of the non-sticking approaches $\{\varphi(x)\}_{x \in X}$ will be employed by someone.
- By **Non-Stopping**, a randomly formed team of size $N_1$ will always find the "best" solution.

**Proof Sketch:** What about the "best" problem solvers?

- A sufficiently big general population will contain at least one agent using **a** best approach.
- In fact, you can pick a population size $N$ large enough, so that it will contain at least $N_1$ many problem- solvers employing best approaches.
- By **Uniqueness**, a team of the $N_1$ "best" problem solvers will consist of agents who all employ the same approach.
- By **Difficulty** and **Non-Stopping**, those agents won't always find the optimal solution.
- Hence, because each initial state has non-zero probability $\nu(x)$, the expected payoff of the randomly- formed team is strictly higher than that of the best problem solvers.

## Intended Applications

Hong and Page think their work has implications for

- Affirmative action in college admissions and hiring
- Scientific practice

Question: Are their assumptions plausible in group decision-making in these cases?

First, consider Non-Stopping, which is a biconditional.

## Non-Stopping 1

Non-Stopping 1: If there is some agent whose approach "sees" a better solution, then the group will not stop.

- Entails that too many cooks **cannot** ruin the broth: no individual's approach is made worse by group work.
- Entails that one agent alone can sway an entire group to continue searching.

## Non-Stopping 2

Non-Stopping 2: If no agent individually "sees" a better solution, then the group will not either.

- Entails that two individuals, employing the "same approach" cannot find a solution unless they could each find the solution individually.
  - Fails to take into account financial, time, and computational costs.
- Entails approaches are not "synergetic": two individuals can "see" an approach only if one of them could have done so individually.

Next, consider Diversity, which says that for each non-optimal solution, there is at least one approach that will get an individual "un-stuck" from that such solution.

When are humans so elver?

## NetLogo

Programming Concepts:

- Randomization
- The Behaviorspace

## References I


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.

Kitcher, P. (1990). The division of cognitive labor. *The Journal of Philosophy*, 87(1):5—22.

Longino, H. E. (2006). Theoretical pluralism and the scientific study of behavior. In *Scientific pluralism.*, pages 102–131. University of Minnesota Press, Minneapolis.

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

Wylie, A. (2012). Standpoint matters. *Presidential Address of the American Philosophy Association*, 86th.