Mathematical Philosophy, Science and Public Policy

Stephan Hartmann

Munich Center for Mathematical Philosophy
LMU Munich

Climbing Ladders, Building Bridges: Mathematical Philosophy at Work

German Center for Research and Innovation
New York, September 2014

Mathematical Philosophy

- Mathematical Philosophy is the study of **philosophical problems** with the help of **mathematical methods**.

Mathematical Philosophy

- Mathematical Philosophy is the study of **philosophical problems** with the help of **mathematical methods**.
- Which methods? Logic, probability theory, set theory, topology,. . . , modeling and simulation,. . . Which method is useful will depend on the problem at hand. The mathematical philosopher is (just as the physicist) a consumer of methods developed in other fields.
- Which problems? Mathematical Philosophy addresses problems from many different philosophical fields (such as epistemology, philosophy of science, metaphysics, ethics, social and political philosophy, etc.). See the Coursera course "Introduction to Mathematical Philosophy" by H. Leitgeb & S. Hartmann.
- Many of the great philosophers were indeed also mathematicians. . .

Gottfried Wilhelm Leibniz (1646–1716)

Philosophy, Science and Public Policy

- Many problems that mathematical philosophers address are directly related to problems in the sciences, e.g. in the decision sciences and in cognitive and social psychology.
- Some problems are also related to topics that concern public policy.
- Hence, Mathematical Philosophy is, at least to some extent, an interdisciplinary endeavor.
- There are many ways how science and philosophy are related, and there is a lot of debate about it at the moment (e.g. in physics). Here I want to stress that mathematical philosophers import methods from the sciences to solve philosophical problems as well as scientific and policy-related problems (with a philosophical flavor), partly in interaction with the scientists and perhaps in a different way than scientists would address these problems.
- The proof of the pudding is in the eating. . .

Overview

II. Decision-Making in the European Union

The Problem

- The European Union currently comprises 28 member states. Some are large (Germany has a population of 82 Mio.), others are small (Malta has a population of .42 Mio.).
- The Council of Ministers (or the Council of the EU) is an institution of the EU that makes many important decisions.
- Question: Which weights should the different states get in the decision-making process?
- The Treatises of Nice (2003) and Lisbon (2007) addressed this question and came up with specific proposals which were or will be implemented.
- We ask: What would be a fair assignment of weights?
- We will explore which assignments result if we adopt fairness conceptions from the philosophical literature.

Some Facts about the European Union

- Founded in 1952 by six states (Belgium, France, Italy, Luxembourg, the Netherlands, and West Germany)
- Successive enlargements in the following years. In 2004, ten new states joined the EU. Meanwhile the EU has 28 member states. (Croatia joined in 2013.)
- The current total population is $P \approx 505$ Mio.
- The government structure of the EU is quite complex: parliament, Council of Ministers, and, of course, the national governments.
- Several treaties regulate the decision making: Nice (2003) and Lisbon (2007).

The Treaty of Nice (2007)

A proposal is accepted by the Council of Ministers if the following three conditions hold:

1. Majority of countries: 50% + 1, if the proposal is made by the Commission; or else at least two-thirds (66.67%)
2. Majority of population: 62%
3. Majority of voting weights: 74%

The weights are to be negotiated and changed with every enlargement of the EU.

Some Considerations

What kind of union is it?

1. A union of states: Equal representation
2. A union of people: Proportional representation
3. Something in-between: degressive proportionality (qualified majority)

The Swedish Proposal

- Let $x_i = P_i/P$ be the ratio of the population in state i (P_i) and the total population of the EU (P).
- Let r_i be the ratio of the number of representatives of state i and the total number of representatives.

The Swedish Proposal

$r_i \sim \sqrt{x_i}$

The Swedish Proposal

- Let $x_i = P_i/P$ be the ratio of the population in state i (P_i) and the total population of the EU (P).
- Let r_i be the ratio of the number of representatives of state i and the total number of representatives.

The Swedish Proposal

$r_i \sim \sqrt{x_i}$

- Chirac: "What is the political significance of the square root?"
- So why should we accept this (or any other) proposal? How can it be justified?

Models of Representation

We parametrize different models of representation as follows: $r_i \sim x_i^\alpha$ where α ranges from 0 to 1:

- $\alpha = 0$: equal representation
- $\alpha = 0$: proportional representation
- $\alpha = 1/2$: the Swedish Proposal

Our goal is to find the optimal value of α.

To do so, we have to specify a criterion.

The Welfarist Framework

- Every decision is associated with a certain utility U for each state.
- If this utility is positive, then the state votes for the motion. If this utility is negative, then the state votes against of the motion.
- If the proposal is accepted (for a given α), then each state gets the original utility.
- Illustration:

$$U(\text{Poland}) = .2 \quad , \quad U(\text{Portugal}) = .3 \quad , \quad U(\text{Germany}) = -.7$$

- To work the idea out and to determine the optimal value of α, some additional model assumptions have to be made.

Simulation Procedure

The model is too hard to study analytically. We therefore run a series of computer simulations. These simulations proceed as follows.

(i) Fix the initial value of $\alpha = 0$.
(ii) Specify a normal distribution $\mathcal{N}(0, \sigma)$ for the utilities.
(iii) Draw a utility for each state from it.
(iv) The states decide accordingly and the decision rule will be applied.
(v) If the motion is accepted, each state gets the original utility, if not nothing happens.
(vi) Repeat steps (iii) to (v) many times and finally average all obtained utilities for each state for that value of α.
(vii) Increase the value of α successively and repeat the whole procedure until one reaches $\alpha = 1$.

Evaluation

- We arrive at a family of utilities $U(\alpha) := (u_1(\alpha), \dots, u_{28}(\alpha))$, which needs to be evaluated.
- To do so, we have to ask what it is that we want to optimize.
- Candidates: The total utility ("utilitarian measure"), an egalitarian measure, a Rawlsian measure,...
- One would expect that different measure choices lead to different results.

Results

1. The utilitarian measure maximizes for $\alpha \approx 1/2$.
2. Interestingly, the egalitarian measure also maximizes for $\alpha \approx 1/2$.
3. Even more interestingly, $\alpha \approx 1/2$ is also optimal if alternative modeling frameworks are used (such as the voting power framework). $\rightarrow$ Model independence

Our approach therefore justifies the Swedish proposal.

III. Anchoring in Deliberations

Motivation

- A group has to decide a factual issue.
- Example: A committee has to decide what the maximally tolerable amount of CO_2 emissions is.
- It turns out that there is disagreement amongst the group members and so the group starts to deliberate.
- In the course of the deliberation, the group members try to convince each other, but everyone is also open to change her mind in the light of the arguments of others.
- In many cases (but of course not always), a consensus emerges eventually.

The Anchoring Effect

- Question: Does the final outcome of the deliberation depend on the order in which the group members speak?
- Empirically one sees that often one speaker dominates the group, jumps up first and sets the bar for the rest of the discussion. She or he anchors the deliberation.
- More specific question: Does the anchoring effect also occur in groups of individually rational members, i.e. in groups where all members change their mind according to rational rules.
- Put differently: Is it possible that a group of rational agents is irrational? (This is a question of philosophical appeal.)
- N.B.: The anchoring effect shows up in many different contexts and has been studied, for example, by Tversky and Kahneman.

The Procedure

- A group of n members has to fix the value of a real-valued parameter x.
- We assume that the group members are ordered from 1 to n.
- Each group member i submits an initial value $x_i^{(0)}$ and additionally assigns a value to the reliability of her judgment.
- The deliberation proceeds in K rounds, and each round proceeds in n steps.
- In round 1, step 1, the first group member presents her arguments for her initial submission (i.e. for $x_1^{(0)}$).

The Procedure Cont'd

- Based on this, the other group members (i) assign a reliability to her and then (ii) update their original submissions, i.e. the second group member updates her initial assignment (i.e. $x_2^{(0)}$) taking $x_1^{(0)}$ as well as her own reliability assignment and the reliability she assigns to the first group member into account.
- The first group member does not update her assignment.
- In step 2, the second group member presents her arguments for her (now already once updated) assignment, and so on.
- In round 2, the same procedure applies. We assume, however, that the reliability assignments improve, i.e. that the group members become increasingly better in judging the reliability of their fellow group members.

The Reliability Assignments

- Every group member i (for $i = 1, \ldots, n$) has a reliability $r_i \in (0, 1)$.
- The reliabilities that are actually used by the group members have discrete values: high (H), medium (M), and low (L).
- Every group member i has access to her own reliability and assigns herself a reliability of H if $r_i \geq 2/3$, L if $r_i \leq 1/3$ and M otherwise.
- Every group member has a second order competence $c_i \in (0, 1)$ to assess the reliability of the others. This quantity is "updated" in the course of deliberation.
- Based on c_i and the value of r_j, group member i assigns a reliability to group member j.

The Updating Procedure

1. I am H (M, or L) and the presenter is my *peer*, i.e. she is also H (M, or L). Then:
$$x_i^{(1)} = \frac{1}{2}\left(x_i^{(0)} + x_1^{(0)}\right)$$
My reliability value remains H (M, or L).
2. I am H, the presenter is L. Then:
$$x_i^{(1)} = x_i^{(0)}$$
My reliability remains H.
3. I am L, the presenter is H. Then:
$$x_i^{(1)} = x_1^{(0)}$$
My reliability changes to H.

The Updating Procedure Cont'd

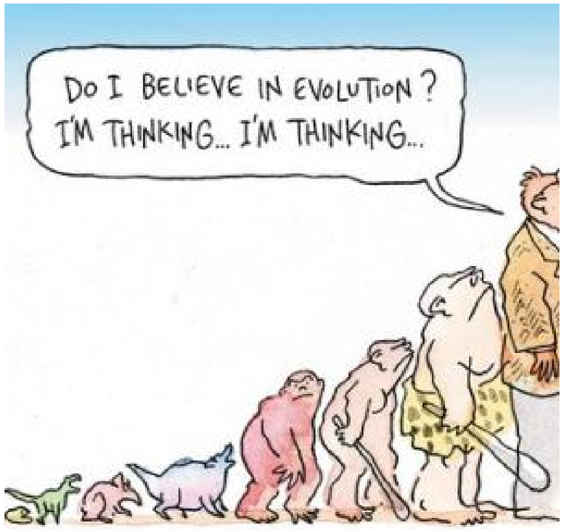
4. I am H (M), the presenter is M (L). Then:
$$x_i^{(1)} = \frac{1}{4}\left(3x_i^{(0)} + x_1^{(0)}\right)$$
My reliability value remains H (M).
5. I am L (M), the presenter is M (H). Then:
$$x_i^{(1)} = \frac{1}{4}\left(x_i^{(0)} + 3x_1^{(0)}\right)$$
My reliability value changes to M (H).

Note that after one round, we have effectively generated a Lehrer-Wagner matrix.

Results

- We run computer simulations to explore the consequences of our model. More specifically, we calculate the probability P_F that the consensus value is closest to the initial value submitted by the first group member (i.e. the group member who speaks first) for various distributions of the reliabilities.
- We find that the probability that the initial assignment of the first speaker is higher than for all other group members, quite independently of the reliabilities. And indeed, the anchoring effect shows up in deliberations under a wide range of conditions.
- The anchoring effect in deliberations is also hard to correct. We are currently working on simple deliberation procedures and study how they fare in computer simulations. If they are promising, we would also like to study them in experiments.

IV. Scrutinizing Scientific Reasoning



Types of Reasoning

Here are three well-known types of reasoning:

1. Deduction
2. Induction
3. Abduction ("Inference to the best explanation")

We use these types of reasoning in ordinary reasoning and in science.

Question: How are these types of reasoning justified?

Here I will focus on an assessment of one new type of reasoning: the no-alternatives argument.

Assessing Scientific Theories

- Question: How are scientific theories assessed? What is a good scientific theory?
- Traditional answer: Theories are assessed in the light of empirical data.
- However: This does not work too well in fundamental physics. In some cases one has to wait very long for an empirical confirmation (e.g. in the case of the Higgs), and in other cases it is not clear whether there will ever be empirical data (e.g. in the case of superstring theory).
- What can be done? Are there "non-empirical" ways of assessing scientific theories?
- Some people believe so, and I will assess the corresponding argument structure.
- Before, however, I want to show that traditional accounts do not apply.

Theory Assessment I: The HD-Model

According to the hypothetico-deductive model a theory or hypothesis H is confirmed by a piece of evidence E if and only if E is predicted by H (i.e. if E is a deductive consequence of H) and if E is observed.

Some remarks:

1. Note that a theory can only be confirmed empirically.
2. The hypothetico-deductive model has a number of problems, e.g.
   - The Tacking Problem: If E confirms H, then it also confirms H∧X.
   - Confirmation is a yes-no matter: there are no degrees of confirmation.
   - Etc.

Theory Assessment II: Bayesian Confirmation Theory

According to the Bayesian Confirmation Theory a theory or hypothesis H is confirmed by a piece of evidence E iff the observation of E raises the probability of H. Here is an illustration.

- A Bayesian agent entertains the following hypothesis:
- H: Tomorrow the sun will shine in New York.
- She assigns a prior probability of, say, $P(\mathrm{H}) = .8$ to it.
- Next, she listens to the radio and hears the weatherman saying that tomorrow the sun will shine in Philadelphia. Hence, she takes as evidence for H the following proposition:
- E: The weatherman says that tomorrow the sun will shine in Philadelphia.
- She then updates her beliefs according to Bayes Theorem: $P'(\mathrm{H}) = P(\mathrm{H}|\mathrm{E})$. Here $P'(\mathrm{H})$ is the posterior probability of H.
- We might then obtain, say, $P'(\mathrm{H}) = .9$: E confirms H.

Discussion

- Bayesian Confirmation Theory can be applied in a straightforward way to empirical testing.
- However, there are other ways of confirming or supporting a theory. Examples: (i) a theory explains a novel phenomenon for which it was not constructed. (ii) a theory coheres well with accepted theories. (iii) a theory belongs to a class of theories that have been successful (e.g. it is a gauge theory).
- We ask: Can these other ways be accounted for in Bayesian Confirmation Theory?
- Another example: The No Alternatives Argument.

The No Alternatives Argument (NAA)

Scientists often argue like this:

1. Hypothesis H satisfies several desirable conditions (incorporates various principles, coheres with other theories,. . .)
2. Despite a lot of effort, the scientific community has not yet found an alternative to H.
3. Hence, we have one reason in support of H.

We ask:

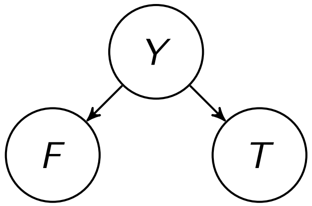
- How good are NAAs?
- Under what conditions do they work?

Examples from Fundamental Physics

There are many examples of NAAs in fundamental physics, mainly because discriminating empirical evidence is hard to come by. Here are three:

1. The Higgs Mechanism
   This mechanism was invented in 1964. Its empirical confirmation had to wait till 2012. However, there was little doubt in the scientific community that the model was in principle correct. There was no alternative. . .
2. String Theory
   This theory cannot (yet) be tested empirically. What speaks in its favor are (mostly unproven) coherence arguments and the NAA.
3. Cosmic Inflation
   This theory enjoys a very limited degree of empirical confirmation (at least this was true until a couple of months ago). Trust in the theory crucially relies on the NAA. This is a nice example which can be used to study how empirical and non-empirical (NAA) confirmation can work together.

Formalizing NAAs: Bayesian Networks at Work



Some Results

- Whether or not a NAA succeeds depends on our beliefs about the number of alternative theories.
- If we are sure that there are infinitely many alternatives theories, then the NAA has no force.
- The same holds for inference to the best explanation which might explain why this type of reasoning is less controversial in ordinary reasoning than in fundamental physics.

Outlook

- I have given three examples of Mathematical Philosophy at work.
- There are many more, and we hope to inspire more people to work on problems at the intersection between philosophy, science and public policy.
- Come and visit us in Munich!

Thanks for your attention!

The talk is based on joint work with Claus Beisbart (Bern), Luc Bovens (LSE), Richard Dawid (MCMP), Jan Sprenger (Tilburg), and Soroush Rafiee Rad (Amsterdam/MCMP).