Unpublished Reviews

Philosophy Reviewed

Wednesday, July 11, 2018

Minds and Machines on Causality and the Brain June 2018, Volume 28, Issue 2,

This volume of Minds and Machines is the product of a conference, which seems largely to have determined the contributions. Although purportedly about science, the essays are often principally directed at those philosophers of science who do not understand the banalities of the sciences they write about or are interested in. (Scientists tend to like this kind of stuff, because it is people saying what the scientists know or think. Everyone likes cheerleading.) Only one of the essays, Romeijin and Williamson's, makes any contribution a brain scientist could conceivably use.

Romeijin and Williamson, Intervention and Identification in Latent Variable Modeling

The authors actually do something. They show that if X, Y, L are binary, and L is the common cause of X, Y, and X, Y are measured and L is unmeasured and there and there are no other causal relations between X and Y, then an exogenous perturbation of the distribution of L allows identification of p(X | L) and p(Y | L) (and of course, p(X,Y | L) for all values of L, without knowledge of the distributions of L before and after perturbation except that the distributions are different.

Of course, it isn't true if the relation between X, Y and L is linear, or if besides the common cause, X influences Y, or if L has more than two values, etc.

The authors give no empirical example that realizes their result. Still, they did something.

Colombo and Naftali, Discovering Brain Mechanisms Using Network Analysis and Causal Modeling

This is a book--or rather multi-paper--report. While there is nothing new in it, the essay is sensible and measured. Unfortunately, it is not up-to-date on search methods for causal signaling relations between brain regions estimated by fMRI or EEG, not even close. Things are happening, fast.

Perhaps philosophers writing what are essentially judgemental review essays ought to talk first with some of the people actually doing the work?

Winning and Bechtel, Rethinking Causality in Biological and Neural Mechanisms: Constraints and Control

Aside from a foray into the "causality power" i (read "oomph") bit of metaphysics, this essay, like many of Bechtel's, is a paradigm of saying the scientifically banal without furthering anything. Banalities, of course, are generally true.

- Gottfried Vosgerau and Patrice Soom, Reduction Without Elimination: Mental Disorders as Causally Efficacious Properties

Here is the upshot:

"our proposal is to analyze mental disorders as higher-level dispositional properties that cause specific symptoms under specific conditions, and that are token-identical to complex physical states. This proposal secures the causal efficacy of mental disorders and their crucial role in explanations, while specifying the systematic relation to lower levels of descriptions as found in neurology and neurochemistry".

That's nice. Thank you Donald Davidson.

Matthew Baxendale and Garrett Mindt, Intervening on the Causal Exclusion Problem for Integrated Information Theory

In gyrations through discussions of the mental and the physical, I look for the takeaway. Here is theirs:

"According to IIT there is an identity between phenomenological properties of experience and informational properties of physical systems...The maximally irreducible conceptual structure (MICS) generated by a complex of elements is identical to its experience... An experience is thus an intrinsic property of a complex of mechanisms in a state."

Speaking of thoughts of complex mechanisms, I wonder what Pluto is thinking now that it's not a planet but still a complex of mechanisms.

I did not read the paper.

Sebastian Wallot and Damian G. Kelty-Stephen, Interaction-Dominant Causation in Mind and Brain, and Its Implication for Questions of Generalization and Replication

I am tired, but just in case you want to read it, you will learn again that that lots of variables affect what people do, so generalization in psychology is hard.

Posted by Unknown at 4:16 PM No comments:
Email ThisBlogThis!Share to TwitterShare to FacebookShare to Pinterest

Friday, June 29, 2018

Dinner at Princeton Comeuppance

N.B. This is a true personal memory about philosophers, not a review,

Dinner in Princeton, or How Bas van Fraassen Shut Up a Bigot, Albeit in a Politically Incorrect Way--He Had It Coming

There were things—most things—about Princeton my wife did not like in 1970, and one thing she did, working at a women's health clinic in inner city (read:, black) Trenton with Rennie Hampshire. In her fifties, Rennie was slim and energetic with the bones and eyes and skin that make for that rare type, a beautiful English woman. She had once been the wife of a famous English philosopher, A.J. Ayer, but was at the time the wife of a still more famous English philosopher, Stuart Hampshire, who was the Chairman of my department. She was a busy person in occupation and manner, never still, and she had a ferocity of spirit that set her apart from anyone else I can remember in my life. Stuart often remarked with pride—and I assume with truth—that she had been the first woman to ride a motorcycle across the United States, but now her energies were directed to a poor black community in a depressed state capital. (In those days, as you crossed the Delaware River into Trenton, you were welcomed by the most self-piteous sign: *Trenton Makes, the World Takes*. And doesn't give much back.)

The clinic, in the basement of an undistinguished building, provided gynecological services, birth control, pregnancy testing and check-ups, abortion referrals (to New York—abortion was illegal in New Jersey then), advice and help of many kinds. It was funded, so far as I could tell, almost entirely by the efforts of Rennie Hampshire. Her husband's position put her often at dinner parties and other social gatherings with the well-heeled and the immensely rich, and she used these chances shamelessly to extort contributions for the clinic. I recall watching her with admiration at a dinner party in her own home, first telling some well-to-do guest about the clinic and it's needs, and then, after dinner, pursuing him until they retired to the kitchen, one to give a check and one to receive. She was not just a fund raiser. She worked at the clinic as did any non-medical volunteer, doing whatever she could that needed doing, from counseling patients to filing records to scrubbing. My wife, Anita, helped her and loved doing it, and she loved Rennie. Rennie alone in Princeton showed Anita that Anita's world—the world of poor, simple people, the world of personal charity—counted, and counted more than the hot house of Princeton refinements, where Anita never felt welcome or at home or free to breathe.

Anita was never so happy as when leaving for Trenton, and never so exhilarated as when she returned.

Rennie had her eccentricities. She would leave small gifts—tea, or chocolates—on our porch, but on the several occasions Anita or I caught her at it, she refused to come into the house. When she and Stuart once came to dinner, she insisted on washing the dishes. I think she was determined not to be fussed over, but really did not quite understand the peculiar texture of lower middle class American formality and informality.

Of all things Princetonian, Anita most despised faculty dinner parties. Her prejudice was confirmed one evening when, after I had begged her on behalf of my career, we attended a party at the home of an eminent (about as eminent as an academic gets) professor of history of science, Thomas Kuhn. The professor's wife, whom Anita found particularly cold, greeted us at the door and threw her arms around my wife in a grand hug, exclaiming how delighted she was that we—and my wife in particular—had come. Anita reddened, not in embarrassment, but in pleasure. I followed them into the house, our hostess walking with her arm wrapped around my wife's waist. Half way to the other guests, she leaned her mouth to Anita's ear and I heard her whisper: *Do remind me, dear, what is your name?* An understandable lapse, surely, but Anita never attended another faculty party, except once.

Bas van Fraassen is a man to envy: handsome, elegant, charming, original, brilliant, sensitive and European; even nowadays, past seventy, he can pull off dressing like a rock star and can fly the trapeze. We were friends, and his warmth and charm made Anita forgive him his occupation (he professed philosophy at Yale then). He liked her too. He was visiting us in Princeton one weekend when, on Saturday morning, Anita answered a phone call from a colleague of mine, Margaret Wilson, inviting us to dinner that evening, with apologies for the short notice. Anita had the perfect excuse—we had a guest for the weekend—but she erred in revealing the guest's name. Margaret, who was a straightforward person, immediately said to bring him, and Anita was stuck. So, with Bas, we went to dinner at the Wilson's.

The party was small: the three of us, Margaret, her husband, and another colleague, George Pitcher. George was tall, slender and athletic—the only faculty member who hit home runs at the annual faculty versus student softball game—with a full head of brushy hair that fell over his forehead in a boyish cowlick but seemed never to be out of place, altogether an extraordinarily handsome man who had at least three loves, his housemate, Ed, and a pair of stray dogs that had taken up residence with them and about which he later wrote a very sentimental book. George was a polite and to appearance a gentle man, but, as I discovered when serving with him as the junior member of a committee of two that made minor personnel decisions, he was privately the kind of bigot that once flourished in Princeton, the kind for which there is only a neologism, a classist.

We sat in the Wilson's dining room, next to a cool bower, around a circular table centered with a bowl of fruit, and had drinks and nibbles and get reacquainted talk before dinner. The talk, inevitably, turned to academic gossip, who is doing what, going where, with whom, the sort of thing that fascinates academic neighbors and bores everyone else, not least Anita, who sat quietly with the bland look that I knew hid an interior woe: *not this again.* Professors, philosophy professors anyway, at parties scarcely ever talk about ideas (that's business), or politics (they all have the same, or none), or religion (what's to talk about?), or sex (not done), or money (not done), or sports (intensely not done). Its gossip, travel and high culture. Somehow, Rennie Hampshire's name came up, and I think Anita mentioned her work in Trenton. George, in his quiet but forceful and authoritative voice, began a rant: *What was she doing there, among the refuse of the city? Why not leave those people to themselves? The trouble with Rennie is that she does not respect her own class.*

The denunciation did not stop, and as it continued Anita changed in color, her smooth face lined in anger, and, I knew, frustration. Her good manners conspired with her sense of social unease; she could not speak without demonstrating her anger. I objected ineffectually, but George talked over me. The Wilson's looked uncomfortable and said nothing while George went on into the evils of crossing the borders of social strata that he seemed to think should be guarded with machine guns and barbed wire, Berlin walls of Class.

Noting Anita's distress and my clumsiness, Bas quietly reached an elegant arm to the centerpiece and removed a banana. Smiling at George, he carefully peeled back half the skin of the fruit, and then, delicately, began to lick the tip. Eyes still full on George, still silently, he pushed the end of the banana into his mouth and, moving it back and forth, began to suck. George reddened, then blushed, and fell silent for the rest of the evening, shamed, if not for the right reason, then, at least, for the wrong one.

Hayley Clatterbuck, The Logic Problem and the Theoretician's Dilemma

While surfing around I ran across Hayley Clatterbuck's, essay, The Logical Problem and the Theoretician's Dilemma, Philosophy and Phenomenological Research doi: 10.1111/phpr.12331 in a journal I usually don't read. It is almost good, just tiptoes up to goodness and gets no farther, leaving some oddities and bad arguments along the way and huge opportunities untouched. It reminded me of two occasions. On one, Judea Pearl asked a dinner party of UCLA philosophers "Why don't you guys *do* anything?" On another, after hearing two hours of lectures by Alvin Goldman on the difference between "hard wired" and "soft-wired" capacities, Allen Newell asked: "So what has your laboratory discovered about hard wired capacities?"

Clatterbuck's problem is the warrant for attribution of understanding to creatures that are not human. She rightly sees that claims of behavioral evidence for such attributions come face to face with the behaviorist version of Hempel's Theoretician's Dilemma. She first proposes that understanding can be established by having independent observable stimuli with correlated responses, inviting explanation by a mediating variable. She represents this by a graphical causal model, roughly

E1 : S1 R1
U
E2: S2 R2

with arrows S1 -> U; S2 -> U; U -> R1; U -> R2. Citing the Causal Markov Condition, which she modestly says she does not fully understand, she claims the graphical model above implies that R1 and R2 are correlated. (That is correct but since S1 and S2 are mutually exclusive, they are associated, so perhaps they should be collapsed into a single variable with 2 values; but Clatterbuck does not want the experimental treatment to be a variable common cause of the results.) She rightly goes on to object that nothing says the mediating variable U has to be some state of understanding; it could be a lot of different things. So she goes on to suggest that designs are needed in which U -> R1 is a positive association and U -> R2 is negative, or vice-versa. I guess the idea is that understanding would produce positive associations in some circumstances and negative ones in others that were perceptually similar. So here is a good reason for positing a mediating variable, essentially using Reichenbach's common cause principle, and an argument I don't fully understand for it's interpretation.

However that works out, her discussion is lexically odd. She says that the graphical model shown, which produces (with Faithfulness) an association between R1 and R2 is "syntactical" but the revised model that produces a *negative* association is "semantic." Associations are syntactic but negative associations are semantic?Since "syntactic" is a term of abuse in contemporary philosophy of science, I wonder at the rationale for her terminology. But on to something more serious.

She argues, I think, that the schema illustrates a way round the Theoretician's Dilemma. Following John Earman, Clatterbuck argues that prior evidence, call it E, provides "inductive support" for some theory t, and t entails (and hence predicts) some new phenomenon N which thus would not have been predicted without recourse to t. Earman puts the argument in Bayesian terms, as does Clatterbuck. But if t entails N then $Pr(N, E) \geq Pr(N, t, E)$ hence $Pr(N \mid E) \geq Pr(N \mid t, E)\ Pr(t \mid E) = Pr(t \mid E)$: the novel phenomenon is at least as probable on the prior evidence as is the theory on the same evidence. The Bayesian could skip the theory and go directly to the predicted phenomenon. So that doesn't work.

Towards the end of her essay, she reformulates the idea in a way that I think she takes to be just an elaboration of Earman's (bad) argument, but is not: "In a case of *within-domain extrapolation*, an empirical regularity within a domain is redescribed in terms of a theoretical relation which is then extrapolated to unobserved cases of that same domain. In a case of *cross-domain extrapolation*, empirical regularities in domain A and B are redescribed in terms of the same theoretical relation, and it is induced that what is true of A is also true of B."

I read this suggestion this way: given data from cases in one domain, find "theoretical" features of those cases that hold in another domain, and use those invariant features to predict about data in this second domain. That would be a reason for theories.

Ok, how can that be done...what is the script, the recipe? How can such invariant theoretical features be found? Nada. Full stop in Clatterbuck's essay. My graduate student, Biwei Huang, has an illustration for neuroscience. Taking fMRI images from subjects in one laboratory, she identifies strengths of some neural causal relations ("effective connections" in contemporary neuropsychology jargon) that separate autistic subjects from normals. She then uses the presence (or absence) of these connections and their strengths inferred from fMRI scans from another laboratory (another "domain") to predict autistic versus normal in the second laboratory (actually, in each of several other laboratories).

This is a pretty neat illustration of the invariance strategy whose suggestion I attribute to Clatterbuck. And it does defeat the Theoreticians Dilemma: you couldn't make comparably accurate predictions by, say, comparing the correlations among fmri signals in different brain regions in subjects in one lab with those in another lab. People have tried.

Huang's example is just that, not a general procedure for learning theoretical invariants for cross-domain classification. My colleague, Kun Zhang, has developed one, explicitly in those terms. Of course, that leaves a lot of room for work on the description of general procedures for other kinds of cross-domain invariants, of which physics provides many examples, e.g., classical thermodyanmics.

Afterword: Goldman replied to Newell that there is a division of labor: philosophers help science by posing the problems and distinctions; psychologists investigate them in the laboratory. Newell said thanks, but psychologists have no trouble doing both jobs.

References

B. Huang, Diagnosis of Autism Spectrum Disorder by Causal Influence Strength Learning from Resting-Stae fMRI Data, M.S. Thesis, Department of Philosophy, Carnegie Mellon University, 2018.

Gong, M., Zhang, K., Huang, B., Glymour, C., Tao, D., & Batmanghelich, K. (2018). Causal Generative Domain Adaptation Networks. *arXiv preprint arXiv:1804.04333*.

Philosophy Reviewed

Friday, June 22, 2018

Remarks on Constructive Empiricism and on Nora Boyd, "Evidence Enriched," Philosophy of Science, 85, 201

Counting the Deer in Princeton

Remarks on Constructive Empiricism and on Nora Boyd, "Evidence Enriched," *Philosophy of Science*, 85, 2018

Once upon a time, philosophers thought that scientific theories are collections of statements about the world. The statements have logical connections that could be studied mathematically by the idealization of formal languages, and the statements have semantic relations that could be studied mathematically by the idealization of model theory, supplemented by various accounts of how terms in the language or mathematical objects in the models relate to things one can see, hear or touch. Then along came constructive empiricism, which kept the idealized models but did away entirely with the formalized language and the logical relations it characterized and said little about how mathematical objects in the models relate to things one can see, hear or touch.

Rather belatedly, two difficulties with constructive empiricism were noticed. The first was, indeed, how the models relate to things we can see, hear or touch, a matter that is, after all, at the heart of empiricism. The answer given is so odd that one might have thought the author was just kidding. The idea is that the theorist has a mathematical data model, and either that model can be embedded in a model of the theory or it cannot be. Van Fraassen considers a theory T of the growth of the deer population in Princeton, and the theorist's data model, a graph of the variation of the deer population over time. He writes: "Since this is *my* representation of the deer population growth, there is *for me*no difference between the question whether T fits the graph and the question whether T fits the deer population growth" (256). The question of whether the mathematical model describes the actual deer population (not *for me*, but in fact) does not arise; it is not even sensible.

Suppose we ask a scientist how the curve of deer population growth in Princeton was obtained, and we are told "For each of several years, I counted the number of hoof marks in Princeton and divided by 4.'" We advise the scientist that his curve may be a severe overcount, since the same deer makes many more than 4 hoof marks. The scientist replies that there is no point to the challenges. If the critics have a different theory, construct their own data model. Constructive empiricism, after all.

Suppose a group of physicists launch a mass spectrometer aboard a satellite to record ion concentrations above the atmosphere. They fail to calibrate the instrument before launch, with the result that it returns values in wild disagreement with previous measurements. (This really happened with the Swedish Freya satellite.) Would the scientists use the data anyway to try to publish a new estimate of ion concentrations? Would referees and a journal editor not care? Of course they would care, and what the scientists actually published was a procedure for calibrating the spectrometer in-flight.

No one who takes science seriously can take seriously this constructive empiricist account of how data and theory meet. Nora Boyd does. Her essay focuses on facts familiar to anyone who has read almost any scientific paper: scientific data typically are accompanied by ancillary information that records the provenance of the measurements: what instruments were used, how they were calibrated and shielded, what resolutions of space or time or other variables were obtained, how were the data censored, or clustered or transformed, what statistical procedures were used, how were the units selected for measurement or treatment, where and when the measurements were made, whether the study was blinded or double-blinded, etc.

This sort of information is typically given in the body of scientific reports or in supplementary material or in documents attached to databanks.

Framing her story as an extension of Van Fraassen's, she claims the value of such ancillary information is twofold: it helps multiple data sources to be used for related problems or investigations or arguments and it "breaks underdetermination." I agree it does the first, but not in a way that is accommodated by constructive empiricism. I doubt it does the second in any sense except that of allowing further tests of a theory or theories; if some other theory can account for all of the same possible evidence—Quine's sense of underdetermination—combining data sets won't distinguish them. But the main thing such information does is something she ignores, something to which van Fraassen seems to think there is no point: it gives assurances that the measurements have not been made by a process that disqualifies them as premises in the assessment of a theory or theories because the measurements are not faithful to the quantities claimed to be measured; and it provides information to investigate whether such assurances are unwarranted. On constructive empiricist grounds, there is no point to such assurances and no point to arguments that quantities have been mismeasured, or to arguments that data treatments destroyed information, or to objections that in view the provenance of the data the wrong statistical procedures were used, or that the experimental design leaves open alternative explanations of the data whose possibility better designs would have eliminated etc. Boyd misses all of that, perhaps because once science is cast in a constructive empiricist framework, faithfulness to the phenomena, truth, is not the point.

Boyd's suggestion that ancillary information helps in the proper use of multiple data sets for a question, or the same data set for multiple problems is of course correct, but it is unintelligible in the constructive empiricist framework. And that is the second belatedly noticed problem with constructive empiricism. On the old-fashioned view, language provides linkages between models. Language makes the connections that a relation in one model is the same relation as in another model. As Hans Halvorson points out, there is no such connection in constructive empiricism, only so many disconnected models, so many monads. A theory that constrains quantities conditionally, Newtonian dynamics for example, has many models under different conditions. One would like to say that the force holding the planets in their orbits is the same as the force acting on pendula, and indeed Newton says just that. On the constructive empiricist reconstruction, these are just different models of the theory, and nothing identifies the property *acceleration*, in one model with the property, *acceleration*, in another. On the old-fashioned philosophy of science that is one of the services of language. Boyd tell me (private communication) that she does not endorse this part of "constructive empiricism," and she does refer to "minimal empiricism."

Minimal empiricism turns out to be bad wine in new bottles. Citing van Fraassen, she says data are acquired to a theoretical purpose, to support, or not, a particular theory, and data are empirical only with respect to such a purpose. Being empirical for a purpose is just what has been called, since longtime, being relevant to a theory or hypothesis. So what determines that relevance? No answer. If I collect data on the spread of California poppies is that relevant to a hypothesis about the acceleration of the universe? Is it if I *say* that is its purpose? Of course, there is no theory of relevance in "constructive empiricism" either. If a theory combines dynamics for the universe with dynamics for the spread of poppies, and someone's "data model" for poppies fits into it, is that evidence for the dynamics I postulate for the universe?

Boyd is a new Ph.D from Pitt HPS, and it is not fair to take her to task. Who then? Pitt HPS. They take smart young people and make them, well, without a sense of what it is personally to discover something worth discovering, even the development of an actually new idea. As Pitt HPS goes, so goes philosophy of science in America, pretty much.

Posted by Unknown at 2:36 PM No comments:
Email ThisBlogThis!Share to TwitterShare to FacebookShare to Pinterest
Tuesday, August 16, 2016

Recent Books on Causation III: Carolina Sartorio, Causation and Free Will, Oxford, 2016

Carolina Sartorio, *Causation and Free Will*, Oxford, 2016

The styles of philosophy change. Spinoza gave us axioms, from which it was patent his "theorems" did not follow. Hobbes, and Locke and Hume gave as long essays. Berkeley and Hume, dialogues. Nowadays, philosophical style is more often like a video game with unspoken rules: the reader is told the author has a goal, followed by example, counterexample, perplex after perplex, which the author dispatches one after another, like so many arcade mopes, with occasional reverses to revive the dead and kill them again. Double tap. And, then, finally, the reader reaches The Theory. Or not. Ellery Ells' endlessly annoying *Probabilistic Causality* is like that, and so, less endlessly—hers is a short, dense book--is Carolina Sartorio's *Causation and Free Will*. You can't say Ells didn't think hard about his topic, he did, and so evidently has Sartorio, but you can say that both of them, and a lot of other philosophers, could have made reading and understanding a lot easier by laying cards on the table to begin with. At least her syntax is not contrived to hide banality beneath bafflement.

Shelled and peeled, the story is this: an action is done freely by a person if (and I suppose only if) the person caused the action via a sequence of events that included, as actual causes, rational (given the person's desires and beliefs) reasons for the act and absences of reasons not to do it, absences, again, as actual causes in "a normal, non-deviant way." (p. 135).

How can absences of reasons be causes, you ask. Easy, you ate ice cream because you did not have a reason not to of the kind "I am allergic to ice cream" because you are not allergic to ice cream and you know it. So the absence of that reason was a cause of your eating ice cream. In the vernacular, we allow absences as causes all the time: my tomato plants died because I didn't water them. Of course, if metaphysicians take the vernacular literally and allow absences as causes then they will have an infinity of them in every case: my plants died because Barack Obama did not water them, and so on. Sartorio is content with that, and presumably content with an infinity of such ghost causes accompanying every cause that actually happens. Essentially, every ceteris paribus clause becomes an infinity of actual but non-actual (because absent) causes.

Absences as causes might seem gratuitous in her story. They are there because she wants to distinguish, on the one hand, between courses of action in which the agent would be sensitive to reasons against the action were the reasons real (the absent causes) and, on the other hand, courses of reasoning in which the agent would not be sensitive to similar reasons were they real (the absent non-causes). Philosophy is in some places Humpty-Dumptyish, and metaphysicians are legally free to talk as they want, including saying that if in deciding to do something you would be sensitive to a reason, were you to have it, a reason that you do not in fact have, then the absence of that reason is a cause of what you do. I don't think such talk helps anything, and in science, where absences are ceteris paribus clauses or shorthands for unknown (or boring) positive details, it's silly.

Absences as causes necessitate recourse to "a normal and non-deviant way," she argues, because the absence of a reason could be a cause of an effect because, were the reason to be present, that would cause some external process (Sartorio likes examples with miraculous neuroscientists standing ready to intervene) to prevent the effect, and so the agent would be "sensitive" to the absence of the reason.

Ever since it became abundantly clear that we are biological and physical machines, not just our bodies, as Descartes allowed, but the whole of us, as Helmholtz allowed, philosophers doing "moral psychology" have tried to reconcile us to the loss of the Thomistic/Cartesian fancy. The plain fact seems to be that we do not have anything of the kind that Aquinas and Descartes claimed for us. So live with it. Daniel Dennett (*Elbow Room*) assures us that we should be content, even happy with our state; it gives us everything we could want. He is wrong. We could want not to be *like that*, and most of us do. The *that*is a machine whose workings are determined—or at least caused—by forces that antedated us. The *that* is a person who has as a zygote or neonate been implanted with a device that determines her subsequent responses to her environment. We do not want to be *like that* even if nature did the implanting. To be in human bondage, and know it, is one of the metaphysical agonies.

One compatibilist response to the metaphysical agony is that it pines for an incoherence, that there could not *thinkably* be a system of the kind Descartes and Aquinas claimed us to be. But of course there could. We have perfectly clear mathematical theories of non-deterministic automata, whose transitions between states (Hilary Putnam once thought of them as mental states) are neither determined nor probabilistic. The other compatibilist response is Orwellian, meaning changing the language. I think Sartorio's response is of the Orwellian kind, but tempered. She says she has the intuition that if the human machine is formed by nature, well, its actions can be free. She doesn't offer a survey of others' opinions. Bless her, she elaborates only on the condition that her intuition is correct.

There remains the serious scientific project of how consciousness, and deliberation happen, and how they came about, and the sociological, anthropological project of understanding the conditions under which various communities claim free agency and when they do not, and how those conditions (which have evidently changed) come about as a social process, and perhaps the moral project of consoling those who agonize for the loss of free will, but there doesn't remain anything metaphysical to do about freedom of the will. Nothing, at least, of value.

Recent Books on Causation II, Douglas Kutach, Causation

Douglas Kutach, *Causation*, Polity Press, 2014

This, too, is an introductory book, but a good one. The author mixes in historical sources with a wide ranging, and generally accurate and informative exposition of contemporary (i.e, since 1946) accounts of the metaphysics of causation. It has some sensible questions for readers. I would use it as a textbook, with some apologies to the students. What apologies?

1. Like most other discussions of the metaphysics of causality, Kutach appeals to what we think we know for motivation, examples and counterexamples, but there is not the least hint of how causes can be, and are, discovered.
2. While the book is less mathophobic than most philosophy texts, it is not always mathematically competent, doesn't use what it does develop well, and presents mathematical examples that will be unenlightening or worse to most students.

a. Early on "linearity" is discussed a propos of causal relations, but the author clearly doesn't mean linearity. It is not clear what he means. Monotonicity perhaps, or non-interaction.
b. Having introduced conditioning and independence and the common cause principle, there is a rather opaque discussion of Reichenbach's attempt to define the direction of time by open versus closed "conjunctive forks" but the author fails to note that closed forks become open when common causes are conditioned on. One question asks students to describe a graphical causal model with a specific probability feature, which would have been straightforward if the reader had been given an illustration of how graphical causal models are parameterized to yield probability relations, but that did not happen.
c. As an example of uncertain extensions of familiar cases, students are referred to transfinite arithmetic. Some help.

3. Some the exposition could be more attractive, notably the explanations of token versus type, singular versus general. Distinctions (never mind notation) from formal logic are suppressed everywhere, even when they would help. The presentation of determinism is unclear and inadequate.
4. Metaphysical discussions of causality inevitably make claims about what people would say without any consideration of what people do say. The extensive psychological literature on causal judgement, some of which has interesting theories, is entirely ignored.

5. And sometimes the author says exactly the opposite of what he means—slip of the keyboard?

Ok, nothing is perfect, there could be better textbooks, but this one is usable, which is to say, given the alternatives, outstanding.

Posted by Unknown at 4:34 PM No comments:
Email ThisBlogThis!Share to TwitterShare to FacebookShare to Pinterest

Saturday, August 13, 2016

Recent Books on Causation, from the Really, Horribly Bad to the So-So to the Pretty Good

There is a bunch of books on causation recently. I expect to review them all here in due time. At least one is so bad that it does not deserve reviewing, let alone having been published, but at least there should be a warning somewhere. So here.

I. The Worst: Stephen Mumford and Rani Lill Anjum, *Causation, A Very Short Introduction,* Oxford, 2013

Causation is meant to be a quick introductory text surveying contemporary and historical views of causation. For an astute reader, it would be very quick, stopping at, say, page 12. Should in misplaced charity that reader venture on, she would find chapters badly organized, missing their targets (e.g, "finding causes" is reduced to an uninformative mention of randomized, controlled trials), historically uninformed, and terribly referenced. But, as I say, any reader on cortical alert would throw the book away around page 12. There, the authors address Russell's early argument that causes cannot be fundamental because causes are asymmetrical and the fundamental laws of physics are symmetrical equations.

Russell is wrong they say, because "equations have at least some directionality." Here is their argument:

"We say that 2 + 2 = 4, for instance, which is to say that each side is of equal sum. But is is less obvious that 4 = 2 + 2 insofar as 4 can also be the sum of 1 + 3. The point is that 2 +2 can equal only one sum 4, whereas 4 can be the sum of several combinations (2 and 2, 1 and 3, 10 minus 6, and so on). And in this respect there is at least some asymmetry." (pp 12-13)

Somewhere, in Norway or Nottingham, the transitivity of equality, and Russell's point, was missed.

Then, in nice condescension, the authors write that

"Second, Russell's account was based on his understanding of the physics of 1913. There have been a number of attempts by physicists to put asymmetry back into physical theory. One such notion is entropy, which an irreversible thermodynamic property."

The last clause of the last sentence is a bit of nonsense, --it's not the property that is irreversible, it's changes in it, but more importantly the idea of entropy, and the word, had been in physics for about 50 years when Russell wrote.   In 1913, Russell didn't understand the physics of 1913, and neither, apparently, did the authors in 2013.

Posted by Unknown at 12:55 PM No comments:
Email ThisBlogThis!Share to TwitterShare to FacebookShare to Pinterest

Wednesday, January 13, 2016

The ;Nonsense of "The Stone"

The New York Times occasional philosophy column, *The Stone*, has built a reputation for unilluminating heat, slovenly inference and wanton accusations. Almost any column would do as an example. I will take a recent reflexive example, "When Philosophy Lost its Way" in the January 11, 2016 Times.

First, what way did philosophy lose? The high moral ground, for one thing, say the Texan authors. Philosophers of yesteryear (before the 19th century) showed integrity and selflessness. Our contemporaries by and large do not. The study of philosophy, in yesteryear, elevated those who pursued it. Of old, philosophers were concerned with human functions and purposes. Now they are not. Philosophy was a quasi-priesthood, a vocation. Now it's just a job. Philosophy of old was spread among the professions, the idle rich, etc. Now it's confined to philosophy professors.

Second, how did philosophy lose its way? It became part of the university. That removed philosophers from "modern life." (I wonder where the philosophy professors live who don't: pay taxes, have illnesses, worry for their children, hold political views, fall in and out of love, get divorced, give to charities, etc. Maybe it's North Texas.) In the good old days, lots of people with different interests were philosophers, but after the 19th century they all became academics. lost their virtue and their connection with human concerns. That's the story.

Unlike the Texas philosophers, I am loathe to defame the integrity or selflessness of contemporary philosophers. I have met a few really vile ones, but mostly they have seemed pretty ordinary folk on moral dimensions. But I am not so sure that philosophers of old were selfless and notably different in integrity from their contemporaries. It reads to me as if the Texans have been taking The Apology as the common standard of philosophers before philosophers became professors. Was Aristotle, who left a contentious democracy to educate the mad son of a monarch, selfless? Was Plato, the Athenian aristocrat, selfless? Moving up, what was selfless about Leibniz—did he sacrifice himself in some way for others? Few characters in intellectual history seem less selfless or charitable than Hobbes and Newton, who saw personally to the mutilation of coin clippers. Integrity (and courage)? You won't find it uncompromised in Locke, who contributed (albeit on tolerance) to the Fundamental Constitution of Carolina, an oligarchy ruling over indentured servants that violated both letter and spirit of Locke's 2nd treatise—which treatise Locke made sure not to publish while he lived.

There are lots of examples of 20th century philosophers who acted with selflessness and integrity. Bertrand Russell, who went to prison over his opposition to World War I; David Malament, who did the same over his opposition to the Vietnam War; Paul Oppenheim and Carl Hempel, who helped Jews out of Germany during the Third Reich; Albert Camus, who was part of the French underground. Philosophers not engaged with modern life? Read Philip Kitcher, read Daniel Dennett's more recent works, read just about anything by Peter Singer. Are there no 20th century philosophers who were not professors? Alan Turing was one of the most influential philosophical writers of the 20thcentury—among other things of course. He held an academic position only in the last years of his life. Camus was a journalist. Paul Oppenheim was a businessman. John von Neumann, who stimulated both the philosophy of quantum theory and computation, was a mathematician. Russell spent most of his career outside of the academy. Lawrence Krauss, a physicist, is a metaphysician as well.

What is true is that as universities spread and secularized, a lot more people became "philosophers" and a lot of them are very ordinary people with ordinary minds. The same is true of lots of disciplines I expect, say physics.

What is the author's remedy? Simple: philosophers should get out of universities. The authors teach at the University of North Texas.

Posted by Unknown at 10:02 AM No comments:
Email ThisBlogThis!Share to TwitterShare to FacebookShare to Pinterest

Causal Decision Theory and Conditioning: a Primer

Standard Savage decision theory as well as Richard Jeffrey's alternative, address a normative problem for an odd doxastic condition. an agent fully believes:

- a set of all of the available, mutually exclusive actions;
- a set of exhaustive and mutually exclusive possible states of the world;
- a set of consequences—outcomes—of each possible state of the world/action pair.

and the agent:

- Has coherent degrees of belief in the possible states of the world;
- Has utilities (or in Jeffrey's version, desirabiities) for the outcomes.

The normative question is which action the agent ought to take. The answer offered is the action, or one of them, that maximizes the expected utility, where the expectation is with respect to the degrees of belief in the states of the world.

From an ideal Bayesian perspective, what is essential is the distinction between actions and outcomes and their costs or values. The ideal Bayesian knows which actions have the maximal expected utility. The states of the world are gratuitous. Followers of Savage, or Jeffrey's in effect assume the agent only obtains the expected utilities by calculating them using the specified states of the world and probabilities of outcomes, given the various possible states of the world and actions.

What is odd is that no epistemological problem is considered about how an agent knows, or could know, or rationally assess, the possible states of the world and their probabilities, the possible actions, or the probabilities of outcomes effected by alternative actions in the several possible states of the world. a thorough subjectivist such as Jeffrey would answer these questions: all that is relevant are the agent's degrees of belief about actions, states of the world,and outcomes and their desirabilities. Epistemology reduces to observing, Bayesian updating, and rather trivial computation. Be that as it may, or may not, causal decision theory considers two kinds of complications.

1. The agent believes that the action chosen will influence the state of the world.
2. The agent believes that the state of the world will influence the action chosen;

This is already a conceptual expansion for the agent, to include causal relations and probabilities of actions.
In case 1, how should the agent take account of the belief that the choice of action will be influenced by the state of the world? For simplicity, first assume the outcome is a

deterministic function of the action, a, and the state, s, of the world, and the utility is U(o(a, s) where o is some function actions and states.

Proposal 1: Calculate the expected utility for each action as the sum over states of the world of the utility of each action in that state of the world multiplied by the probability of that state of the world given the action:

(1) $Exp(U(a)) = \Sigma_s U(o(a,s)) Prob(s \mid a)$

In Savage theory the last factor on the right hand side of (1) and (2) is just Prob(s)

One "partition question" concerns whether the action that maximizes utility is the same depending on how the set of states is "partitioned." Let S be a variable that ranges over some finite set of values, s1,…,sn. a coarsening of S is a set S1 = {{s1 v..v sk}, {sk +1 v …v sm},….{sm v…v sn}}, etc. a refinement is the inverse.

Coarsening can change the probability of an outcome on an action. Let S = {s1, s2, s3} and suppose S' is a coarsening of S to {(s1 v s2), s3}. For all outcomes o and actions a, let o and a be independent conditional on s1 and likewise on s2 and s3, but S not be independent of A. Then for any outcome in O:

P(O | a, (s1 v s2)) = P(O, | (a,s1 v a,s2) = P (O, a, (s1 v s2)) / (P(a,s1 v a,s2)) =

P((O, a, s1) v P(O, a, s2)) / (P(a,s1 v a,s2)) =

P(O, a, s1) + P(O, a, s2) / ((P(a,s1) +P(a,s2)) =

[P(O | a, s1) P(a, s1) + P(O | a, s2)] / ((P(a,s1) +P(a,s2)) =

[P(O | s1) P(a, s1) + P(O | s2) P(a, s2)] / ((P(a,s1) +P(a,s2)) =

(P(a) [P(O | s1) P(s1 | a) + P(O | s2) P(s2 | a)]) / (P(a)((P(s1 | a) +P(s2) | a)) =

[P(O | s1) P(s1 | a) + P(O | s2) P(s2 | a)] / (P(s1 | a) + P(s2) | a))

The probability distribution of 0 given the state s1 v s2 in S' varies as the conditional probabilities of s1 and, respectively, of s2 vary with the value of A they are conditioned on, and O and A are not independent in S' but they are independent—by assumption—in S.

For case 2, the results and the argument are similar. The general point is an old one, Yule's (on the mixture of records).

The partitioning problem does not apply to Savage's theory—it makes no difference how the range of possible state values are cut up into new coarsened variables.

So decision theory when the actions influence the states or the states influence the actions is up in the air—the right decision depends on the right way to characterize the states. Various writers, Lewis, Skyrms, Woodruff and others, have proposed vague or ad hoc or infeasible solutions. Lewis proposed to chose the most specific "causally relevant" partition, which I take to mean the finest partition for which there is a difference in elements of the partition in the probabilities of outcomes conditional on actions. Skyrms objects that this is often unknowable, and proposes an intricate set of alternative conditions, which Woodruff generalizes. The general strategy is to embed the problem in a logical theory of conditonals, and entwine it with accounts of "chance"and relations of chance and degrees of belief, e.g., the principal principle. The general point is hard to extract.

When states influence actions Meek and Glymour propose that there are two theories. One simply calculates the expected values of the outcomes on various actions as with Jeffrey's decision theory, the other assumes that a decisive act is done with freedom of the will, represented as an exogenous variable, that breaks the influence of the state on the act.

Appealing as the second story may be to our convictions about our own acts as we do them, or deliberate on what to do, it is of no avail when the actions influence the states, not vice-versa. For that case, one either knows the total effect of an action on the outcome, or one doesn't, and if one doesn't, there is nothing for it except to know what the states are that make a difference. One would think serious philosophy would have focused then on means to acquire such knowledge. One would be wrong.

Monday, August 25, 2014

Review of Philosophy of Science, 81, July, 2014

This issue of *Philosophy of Science* contains some good, some bad, some odd. It gives evidence that methodology in philosophy of science is pretty much in the doldrums or worse, while good work is being done producing economic models for various ends.

Brian Skyrms, Grades of Inductive Skepticism

Reject.

This is a very brief rehash of some history of probability, coupled with some remarks on ergodic probabilities, remarks that go nowhere. The piece seems oddly trivial and unworthy of its distinguished author. One has to wonder why it was published—or submitted. Hypothesis: The author is eminent and a colleague of the editors. That sort of thing has happened before in *Philosophy of Science*, although not that I can think of under the current editors. But one of the things colleagues should do for one another is discourage the publication of stuff that is trivial or bad in other ways.

Ben Jantzen,

Piecewise versus Total Support: How to Deal with Background Information in I

Accept.

Likelihood has an apparent problem. Suppose you are weighing hypotheses h1 and h2. You know b. You learn e. Should you compare h1 and h2 by

p(e | h1, b) / p(e | h2, b) or by p(e, b | h1) / p(e, b | h2)?

Which hypothesis is preferred may not always be the same on the two comparisons. Jantzen makes the sensible suggestion that which to use depends on whether you are asking about the extra support e gives to h1 versus h2 in a context in which b is known, or whether you are asking about the total support. Jantzen's point is not subtle, but the paper is well done and the examples (especially about fishing with nets with holes too large) are illuminating.

Which reminds me of a deeper problem with likelihood ideas that seem not to be much explored: *likelihood doctrine seems to imply instrumentalism*.

Likelihood arguments are used not just to compare hypotheses but to endorse hypotheses, e.g., via maximum likelihood inference. Consider two principles:

1. Hypotheses addressing a body of data should be preferred according to the likelihood they give to that data.
2. A hypothesis should not be endorsed if it is known that there are other hypotheses that are preferred or indifferent to it by criterion 1 above, especially not if there is a method to find such alternatives .

If the data is finite, the hypothesis just stating the evidence has maximum likelihood. So some additional principle is required if likelihood methodology is to yield anything more than data reports. The hypothesis space must somehow be restricted.

Try this:

3. Only hypotheses that make predictions beyond the data are to be considered.

So suppose there are data e1…en and consider some new experiment or observation e not in the data but for which "serious" hypotheses explaining e1…en gives some probability to the outcomes. Let the outcomes be binary for simplicity and so h gives the probability to be is P(e | h). Consider the hypotheses:

e1&…&en & $\text{argmax}_{<h,>}$ (P(e | h) if $\text{argmax}_{<h,>}$ (P(e | h) > $\text{argmax}_{<h,>}$ (P(~e | h) ,and e1&…&en & $\text{argmax}_{<h,>}$ (P(~e | h) otherwise

This hypothesis meets condition 3 and gives e (or ~e) a likelihood at least as great as any alternative hypothesis.

Ok, try this:

> 4. Only hypotheses that make an infinity of predictions are to be considered.

But the stupid pet trick above can be done infinitely many times. So try this

> 5. The hypotheses must be finitely axiomatizable.

Still won't do, as (I think) an easy adaptation of) the proof in http://www.jstor.org/stable/41427286 shows.

Lina Jansson

Causal Theories of Explanation and the Challenge of Explanatory Disagreement

Reject

Both the thesis and the argument of this paper are either opaque or weird; it is difficult to see the warrant for publishing. Her stalking horses are "causal accounts of explanation." On Streven's account, causal asymmetry is why X explains Y rather than the other way round—Dan Hausman had that idea earlier; on Woodward's account, X causes Y but Y does not cause X implies that a manipulation of X changes a manipulation of Y, but not vice versa. So far as I know, neither of them claim that all explanations are causal explanations. But a lot of them are.

Jansson's argument seems to be as follows:

Leibniz held that Newton's gravitational theory was not a causal explanation, because causal explanations require mechanisms and no mechanism was given for gravitational attraction. She reads Newton as "causally agnostic" about his laws, which seems to me a very long reach. He was agnostic (publicly) about the mechanisms that produce the laws, but not that the laws imply causal regularities: drop a ball and that will, ceteris paribus, cause it to take up a sequence of positions at times in accordance with the law of gravity. But suppose, for argument, she is right, then what is the argument?

She writes: "Put simply, the problem of understanding this debate from a causal explanatory perspective stems from the reluctance, on both sides, to take there to be a straightforward causal explanation given by the theory." And, a sine qua non of a correct account of explanation is that it be able to "understand the debate. "

There is this oddity about universal gravitation and causation. If I drop a ball it causes the ball to fall, the ball's falling influences the motion of Mars (instantaneously on Newton's theory), and the change in the motion of Mars influences the course of the ball, also instantaneously. Immediate feedback loop. But Mars influence doesn't determine the position of the ball after I drop it, and the position of the ball after I drop it doesn't cause my dropping it.

Anyway, her point is different. Here is the form of the argument.

Accounts S and W say Newtonian gravitational theory is causal.
Neither the creator of the theory nor its most prominent critic unequivocally said it was causal.

Therefore accounts S and W are false (or inadequate, or something).

Parallels.

A: Chemical changes involve the combination or releases of substances made up of elements.

Lavoisier said combustion involves combination with oxygen.
Priestley said combustion involves the release of phlogiston

Therefore A is false.

The theory of probability specifies measures satisfying Kolmogoroff's axioms.

Bayesians say probability is opinion.
Frequentists say probability is frequency

Therefore the theory of probability is false.

Jansson's "methodology" assumes that concepts of causation and explanation never change, and that historical figures are always articulate, and never make errors of judgement in the application of a concept, and that if some historical figure would only apply a concept under restrictive circumstances (e.g., no action at a distance), an account of the concept must agree with that judgement or posit a new concept. Individuation of concepts is a vague and arbitrary matter—are there the concept of causality, Leibniz's concept of causality, Newton's concept of causality, etc.? On her view, so far as I can see, for every sentence about causal relations, general or specific, about which some scientists sometime have disagreed, two new concepts will be needed. Not much to be learned from that.

Robert Batterman and Colin Rice
Minimal Models
Revise and resubmit

Another essay on explanation (will philosophers of science *ever* let up on this) whose exact point is difficult to identify.

"We have argued that there is a class of explanatory models that are explanatory for reasons that have largely been ignored in the literature. These reasons involve telling a story that is focused on demonstrating why details do not matter. Unlike mechanist, causal, or difference-making accounts, this story does not require minimally accurate mirroring of model and target system.

We call these explanations *minimal model explanations* and have given a detailed account of two examples from physics and biology. Indeed, minimal model explanations are likely common in many scientific disciplines, given that we are often interested in explaining macroscale patterns that range over extremely diverse systems. In such instances, a minimal model explanation will often provide the deeper understanding we are after. Furthermore, the account provided here shows us why scientists are able to use models that are only caricatures to explain the behavior of real systems."

The idea seems to be that there are theories that find features and relations among them that entail phenomenological regularities, no matter the rest of the features of a system, and no matter whether the features in question are exactly exemplified in a system. There are two examples, one from fluid dynamics, the other Fisher's opaque explanation of the 1:1 sex ratio in many species based on the equal effort required to raise males or female offspring, but the differential average reproductive return to raising males if females are in excess or raising females if males are in excess. I don't understand the fluid dynamics model, and Fisher's requires a lot of extra assumptions and ceteris paribus clauses to go through, (grant the equal cost of rearing male and female offspring but imagine that one male can fertilize many females and there is a predator that prefers males exclusively) but never mind.

What I don't understand about this paper is why most theories in the physical sciences don't satisfy B and C's criteria for a minimal model. Thermodynamics? The details of the molecular constitution of a system are largely ignored. Relativity? It doesn't matter whether the system is made of wood or iron, the Lorentz tranformations still hold; it doesn't matter how the light is generated, its velocity is still the same. Newtonian celestial mechanics? Doesn't matter that Jupiter is made of gas, Mercury of rock, and Pluto of ice, still the same planetary motions. Even theories that probe into the internal structure of a system are minimal with respect to some other theories. Dalton appealed only to masses of elemental particles—that, and a few assumptions yields the law of definite proportions. Berzelius added electrical forces between atoms, which were gratuitous for deriving definite proportions.

What is not clear in this paper is how B & C intend to distinguish between minimal models and almost every theory that shows a set of features, individual or aggregate, or approximations to such features, and related laws, of a kind of system suffice for phenomenological relations. That is what physical theories generally do. Their fluid flow example almost suggests that all that is required is an algorithm that generates the phenomena from (perhaps) measurable features a system. So, considering that example, the authors might have asked: when is an algorithm for generating the phenomena an explanation of the phenomena? They did not.

Dean Peters

What Elements of Successful Scientific Theories Are the Correct Targets for "Selective" Scientific Realism?

Revise and resubmit

Peters' essay is useful in two respects. First, it treats the question in the title as turning on this: what parts of the data confirm what parts of a theory? That adds a little structure to the philosophical discussions of realism. And, second, it provides a succinct critical review of bad proposals to answer the question. Peters' has his own answer, which is not obviously useful. Here it is:

"So, to pick out the essential elements of the theory under the ESSA, start with a subtheory consisting of statements of its most basic confirmed empirical consequences or perhaps its confirmed phenomenological laws. These, after all, are the parts of a theory that even empiricists agree we should be "realists" about. Further propositions are added to this subtheory by a recursive procedure. Consider any theoretical posit not in the subtheory. If it entails more propositions in the subtheory than are required to construct it, tag it as confirmed under the unification criterion, and so add it to the subtheory. Otherwise, leave it out. When there are no more theoretical posits to consider in this way, the subtheory contains the essential elements of the original theory."

The proposal as developed is insubstantial: "Consider any theoretical posit not in the subtheory. If it entails more propositions in the subtheory than are required to construct it" – what does "required to construct it" mean?

In criticizing other proposals, Peters appeals to logical consequences, and proceeds with a distinguished set of "posits"—i.e., axioms. Hold him to the same standard. Theories can be axiomatized in an infinity of ways. We need an account of the invariance of the result of the procedure—whatever it is—over different axiomatizations, or an account of "natural axiomatizations" and warrant for using them exclusively. The work of Ken Gemes and Gerhard Schurz is relevant here. So it seems to me that Peters has an idea—conceivably ultimately a good idea—that he did not do the work to make good on.

Roger DeLanghe

A unified model of the division of cognitive labor

Accept

This is a very nice essay providing a simple economic model in which there are balancing incentives for scientists to adopt and contribute to an existing theory or to propose a new one. Lots that might be done to expand the picture for more realism, and it would be nice if those pursuing Kitcher's original idea assembled some relevant data.

Marius Stan

Unity for Kant's Natural Philosophy

I have no opinion about this essay, which is on how Kant might have sought, although he did not, synthetic a priori grounds for Euler's torque law. Nor do I see why anyone should care. Clearly, some do.

Carlos Santana

Ambiguity in Cooperative Signaling
Accept

This well argued and lucid essay shows that there is a model in which agents with ambiguous signaling (under replicator dynamics) invade a population of unambiguous signalers, but not vice-versa. Despite the considerable empirical evidence the author (a graduate student at Penn) gives for the insufficiency of other explanations of the frequency of ambiguity in human and animal communication, I am worried by the following thought. The evolution of language—or at least signaling-- we expect to have gone from the very ambiguous to the more precise. That is what syntactic structure and an expanded lexicon afford. So if signaling by ambiguous strategies cannot be invaded by signaling by "standard" (i.e., perfectly precise) strategies, how did more precise, if still ambiguous in some respects, signaling systems evolve? It strikes me that the author may have proved the wrong result.

Philosophy Reviewed

Friday, June 22, 2018

Remarks on Constructive Empiricism and on Nora Boyd, "Evidence Enriched," Philosophy of Science, 85, 201

Counting the Deer in Princeton

Remarks on Constructive Empiricism and on Nora Boyd, "Evidence Enriched," *Philosophy of Science*, 85, 2018

Once upon a time, philosophers thought that scientific theories are collections of statements about the world. The statements have logical connections that could be studied mathematically by the idealization of formal languages, and the statements have semantic relations that could be studied mathematically by the idealization of model theory, supplemented by various accounts of how terms in the language or mathematical objects in the models relate to things one can see, hear or touch. Then along came constructive empiricism, which kept the idealized models but did away entirely with the formalized language and the logical relations it characterized and said little about how mathematical objects in the models relate to things one can see, hear or touch.

Rather belatedly, two difficulties with constructive empiricism were noticed. The first was, indeed, how the models relate to things we can see, hear or touch, a matter that is, after all, at the heart of empiricism. The answer given is so odd that one might have thought the author was just kidding. The idea is that the theorist has a mathematical data model, and either that model can be embedded in a model of the theory or it cannot be. Van Fraassen considers a theory T of the growth of the deer population in Princeton, and the theorist's data model, a graph of the variation of the deer population over time. He writes: "Since this is *my* representation of the deer population growth, there is *for me*no difference between the question whether T fits the graph and the question whether T fits the deer population growth" (256). The question of whether the mathematical model describes the actual deer population (not *for me*, but in fact) does not arise; it is not even sensible.

Suppose we ask a scientist how the curve of deer population growth in Princeton was obtained, and we are told "For each of several years, I counted the number of hoof marks in Princeton and divided by 4.'" We advise the scientist that his curve may be a severe overcount, since the same deer makes many more than 4 hoof marks. The scientist replies that there is no point to the challenges. If the critics have a different theory, construct their own data model. Constructive empiricism, after all.

Suppose a group of physicists launch a mass spectrometer aboard a satellite to record ion concentrations above the atmosphere. They fail to calibrate the instrument before launch, with the result that it returns values in wild disagreement with previous measurements. (This really happened with the Swedish Freya satellite.) Would the scientists use the data anyway to try to publish a new estimate of ion concentrations? Would referees and a journal editor not care? Of course they would care, and what the scientists actually published was a procedure for calibrating the spectrometer in-flight.

No one who takes science seriously can take seriously this constructive empiricist account of how data and theory meet. Nora Boyd does. Her essay focuses on facts familiar to anyone who has read almost any scientific paper: scientific data typically are accompanied by ancillary information that records the provenance of the measurements: what instruments were used, how they were calibrated and shielded, what resolutions of space or time or other variables were obtained, how were the data censored, or clustered or transformed, what statistical procedures were used, how were the units selected for measurement or treatment, where and when the measurements were made, whether the study was blinded or double-blinded, etc. This sort of information is typically given in the body of scientific reports or in supplementary material or in documents attached to databanks.

Framing her story as an extension of Van Fraassen's, she claims the value of such ancillary information is twofold: it helps multiple data sources to be used for related problems or investigations or arguments and it "breaks underdetermination." I agree it does the first, but not in a way that is accommodated by constructive empiricism. I doubt it does the second in any sense except that of allowing further tests of a theory or theories; if some other theory can account for all of the same possible evidence—Quine's sense of underdetermination—combining data sets won't distinguish them. But the main thing such information does is something she ignores, something to which van Fraassen seems to think there is no point: it gives assurances that the measurements have not been made by a process that disqualifies them as premises in the assessment of a theory or theories because the measurements are not faithful to the quantities claimed to be measured; and it provides information to investigate whether such assurances are unwarranted. On constructive empiricist grounds, there is no point to such assurances and no point to arguments that quantities have been mismeasured, or to arguments that data treatments destroyed information, or to objections that in view the provenance of the data the wrong statistical procedures were used, or that the experimental design leaves open alternative explanations of the data whose possibility better designs would have eliminated etc. Boyd misses all of that, perhaps because once science is cast in a constructive empiricist framework, faithfulness to the phenomena, truth, is not the point.

Boyd's suggestion that ancillary information helps in the proper use of multiple data sets for a question, or the same data set for multiple problems is of course correct, but it is unintelligible in the constructive empiricist framework. And that is the second belatedly noticed problem with constructive empiricism. On the old-fashioned view, language provides linkages between models. Language makes the connections that a relation in one model is the same relation as in another model. As Hans Halvorson points out, there is no such connection in constructive empiricism, only so many disconnected models, so many monads. A theory that constrains quantities conditionally, Newtonian dynamics for example, has many models under different conditions. One would like to say that the force holding the planets in their orbits is the same as the force acting on pendula, and indeed Newton says just that. On the constructive empiricist reconstruction, these are just different models of the theory, and nothing identifies the property *acceleration*, in one model with the property, *acceleration*, in another. On the old-fashioned philosophy of science that is one of the services of language. Boyd tell me (private communication) that she does not endorse this part of "constructive empiricism," and she does refer to "minimal empiricism."

Minimal empiricism turns out to be bad wine in new bottles. Citing van Fraassen, she says data are acquired to a theoretical purpose, to support, or not, a particular theory, and data are empirical only with respect to such a purpose. Being empirical for a purpose is just what has been called, since longtime, being relevant to a theory or hypothesis. So what determines that relevance? No answer. If I collect data on the spread of California poppies is that relevant to a hypothesis about the acceleration of the universe? Is it if I *say* that is its purpose? Of course, there is no theory of relevance in "constructive empiricism" either. If a theory combines dynamics for the universe with dynamics for the spread of poppies, and someone's "data model" for poppies fits into it, is that evidence for the dynamics I postulate for the universe?

Boyd is a new Ph.D from Pitt HPS, and it is not fair to take her to task. Who then? Pitt HPS. They take smart young people and make them, well, without a sense of what it is personally to discover something worth discovering, even the development of an actually new idea. As Pitt HPS goes, so goes philosophy of science in America, pretty much.

Posted by Unknown at 2:36 PM No comments:
Email ThisBlogThis!Share to TwitterShare to FacebookShare to Pinterest
Tuesday, August 16, 2016

Recent Books on Causation III: Carolina Sartorio, Causation and Free Will, Oxford, 2016

Carolina Sartorio, *Causation and Free Will*, Oxford, 2016

The styles of philosophy change. Spinoza gave us axioms, from which it was patent his "theorems" did not follow. Hobbes, and Locke and Hume gave as long essays. Berkeley and Hume, dialogues. Nowadays, philosophical style is more often like a video game with unspoken rules: the reader is told the author has a goal, followed by example, counterexample, perplex after perplex, which the author dispatches one after another, like so many arcade mopes, with occasional reverses to revive the dead and kill them again. Double tap. And, then, finally, the reader reaches The Theory. Or not. Ellery Ells' endlessly annoying *Probabilistic Causality* is like that, and so, less endlessly—hers is a short, dense book--is Carolina Sartorio's *Causation and Free Will*. You can't say Ells didn't think hard about his topic, he did, and so evidently has Sartorio, but you can say that both of them, and a lot of other philosophers, could have made reading and understanding a lot easier by laying cards on the table to begin with. At least her syntax is not contrived to hide banality beneath bafflement.

Shelled and peeled, the story is this: an action is done freely by a person if (and I suppose only if) the person caused the action via a sequence of events that included, as actual causes, rational (given the person's desires and beliefs) reasons for the act and absences of reasons not to do it, absences, again, as actual causes in "a normal, non-deviant way." (p. 135).

How can absences of reasons be causes, you ask. Easy, you ate ice cream because you did not have a reason not to of the kind "I am allergic to ice cream" because you are not allergic to ice cream and you know it. So the absence of that reason was a cause of your eating ice cream. In the vernacular, we allow absences as causes all the time: my tomato plants died because I didn't water them. Of course, if metaphysicians take the vernacular literally and allow absences as causes then they will have an infinity of them in every case: my plants died because Barack Obama did not water them, and so on. Sartorio is content with that, and presumably content with an infinity of such ghost causes accompanying every cause that actually happens. Essentially, every ceteris paribus clause becomes an infinity of actual but non-actual (because absent) causes.

Absences as causes might seem gratuitous in her story. They are there because she wants to distinguish, on the one hand, between courses of action in which the agent would be sensitive to reasons against the action were the reasons real (the absent causes) and, on the other hand, courses of reasoning in which the agent would not be sensitive to similar reasons were they real (the absent non-causes). Philosophy is in some places Humpty-Dumptyish, and metaphysicians are legally free to talk as they want, including saying that if in deciding to do something you would be sensitive to a reason, were you to have it, a reason that you do not in fact have, then the absence of that reason is a cause of what you do. I don't think such talk helps anything, and in science, where absences are ceteris paribus clauses or shorthands for unknown (or boring) positive details, it's silly.

Absences as causes necessitate recourse to "a normal and non-deviant way," she argues, because the absence of a reason could be a cause of an effect because, were the reason to be present, that would cause some external process (Sartorio likes examples with miraculous neuroscientists standing ready to intervene) to prevent the effect, and so the agent would be "sensitive" to the absence of the reason.

Ever since it became abundantly clear that we are biological and physical machines, not just our bodies, as Descartes allowed, but the whole of us, as Helmholtz allowed, philosophers doing "moral psychology" have tried to reconcile us to the loss of the Thomistic/Cartesian fancy. The plain fact seems to be that we do not have anything of the kind that Aquinas and Descartes claimed for us. So live with it. Daniel Dennett (*Elbow Room*) assures us that we should be content, even happy with our state; it gives us everything we could want. He is wrong. We could want not to be *like that*, and most of us do. The *that*is a machine whose workings are determined—or at least caused—by forces that antedated us. The *that* is a person who has as a zygote or neonate been implanted with a device that determines her subsequent responses to her environment. We do not want to be *like that* even if nature did the implanting. To be in human bondage, and know it, is one of the metaphysical agonies.

One compatibilist response to the metaphysical agony is that it pines for an incoherence, that there could not *thinkably* be a system of the kind Descartes and Aquinas claimed us to be. But of course there could. We have perfectly clear mathematical theories of non-deterministic automata, whose transitions between states (Hilary Putnam once thought of them as mental states) are neither determined nor probabilistic. The other compatibilist response is Orwellian, meaning changing the language. I think Sartorio's response is of the Orwellian kind, but tempered. She says she has the intuition that if the human machine is formed by nature, well, its actions can be free. She doesn't offer a survey of others' opinions. Bless her, she elaborates only on the condition that her intuition is correct.

There remains the serious scientific project of how consciousness, and deliberation happen, and how they came about, and the sociological, anthropological project of understanding the conditions under which various communities claim free agency and when they do not, and how those conditions (which have evidently changed) come about as a social process, and perhaps the moral project of consoling those who agonize for the loss of free will, but there doesn't remain anything metaphysical to do about freedom of the will. Nothing, at least, of value.

Monday, August 15, 2016

Recent Books on Causation II, Douglas Kutach, Causation

Douglas Kutach, *Causation*, Polity Press, 2014

This, too, is an introductory book, but a good one. The author mixes in historical sources with a wide ranging, and generally accurate and informative exposition of contemporary (i.e, since 1946) accounts of the metaphysics of causation. It has some sensible questions for readers. I would use it as a textbook, with some apologies to the students. What apologies?

1. Like most other discussions of the metaphysics of causality, Kutach appeals to what we think we know for motivation, examples and counterexamples, but there is not the least hint of how causes can be, and are, discovered.
2. While the book is less mathophobic than most philosophy texts, it is not always mathematically competent, doesn't use what it does develop well, and presents mathematical examples that will be unenlightening or worse to most students.

a. Early on "linearity" is discussed a propos of causal relations, but the author clearly doesn't mean linearity. It is not clear what he means. Monotonicity perhaps, or non-interaction.

b. Having introduced conditioning and independence and the common cause principle, there is a rather opaque discussion of Reichenbach's attempt to define the direction of time by open versus closed "conjunctive forks" but the author fails to note that closed forks become open when common causes are conditioned on. One question asks students to describe a graphical causal model with a specific probability feature, which would have been straightforward if the reader had been given an illustration of how graphical causal models are parameterized to yield probability relations, but that did not happen.

c. As an example of uncertain extensions of familiar cases, students are referred to transfinite arithmetic. Some help.

3. Some the exposition could be more attractive, notably the explanations of token versus type, singular versus general. Distinctions (never mind notation) from formal logic are suppressed everywhere, even when they would help. The presentation of determinism is unclear and inadequate.
4. Metaphysical discussions of causality inevitably make claims about what people would say without any consideration of what people do say. The extensive psychological literature on causal judgement, some of which has interesting theories, is entirely ignored.
5. And sometimes the author says exactly the opposite of what he means—slip of the keyboard?

Ok, nothing is perfect, there could be better textbooks, but this one is usable, which is to say, given the alternatives, outstanding.

Saturday, August 13, 2016

Recent Books on Causation, from the Really, Horribly Bad to the So-So to the Pretty Good

There is a bunch of books on causation recently. I expect to review them all here in due time. At least one is so bad that it does not deserve reviewing, let alone having been published, but at least there should be a warning somewhere. So here.

I. The Worst: Stephen Mumford and Rani Lill Anjum, *Causation, A Very Short Introduction,* Oxford, 2013

Causation is meant to be a quick introductory text surveying contemporary and historical views of causation. For an astute reader, it would be very quick, stopping at, say, page 12. Should in misplaced charity that reader venture on, she would find chapters badly organized, missing their targets (e.g, "finding causes" is reduced to an uninformative mention of randomized, controlled trials), historically uninformed, and terribly referenced. But, as I say, any reader on cortical alert would throw the book away around page 12. There, the authors address Russell's early argument that causes cannot be fundamental because causes are asymmetrical and the fundamental laws of physics are symmetrical equations.

Russell is wrong they say, because "equations have at least some directionality." Here is their argument:

"We say that 2 + 2 = 4, for instance, which is to say that each side is of equal sum. But is is less obvious that 4 = 2 + 2 insofar as 4 can also be the sum of 1 + 3. The point is that 2 +2 can equal only one sum 4, whereas 4 can be the sum of several combinations (2 and 2, 1 and 3, 10 minus 6, and so on). And in this respect there is at least some asymmetry." (pp 12-13)

Somewhere, in Norway or Nottingham, the transitivity of equality, and Russell's point, was missed.

Then, in nice condescension, the authors write that

"Second, Russell's account was based on his understanding of the physics of 1913. There have been a number of attempts by physicists to put asymmetry back into physical theory. One such notion is entropy, which an irreversible thermodynamic property."

The last clause of the last sentence is a bit of nonsense, --it's not the property that is irreversible, it's changes in it, but more importantly the idea of entropy, and the word, had been in physics for about 50 years when Russell wrote.   In 1913, Russell didn't understand the physics of 1913, and neither, apparently, did the authors in 2013.

Posted by Unknown at 12:55 PM No comments:
Email ThisBlogThis!Share to TwitterShare to FacebookShare to Pinterest

Wednesday, January 13, 2016

The ;Nonsense of "The Stone"

The New York Times occasional philosophy column, *The Stone*, has built a reputation for unilluminating heat, slovenly inference and wanton accusations. Almost any column would do as an example. I will take a recent reflexive example, "When Philosophy Lost its Way" in the January 11, 2016 Times.

First, what way did philosophy lose? The high moral ground, for one thing, say the Texan authors. Philosophers of yesteryear (before the 19th century) showed integrity and selflessness. Our contemporaries by and large do not. The study of philosophy, in yesteryear, elevated those who pursued it. Of old, philosophers were concerned with human functions and purposes. Now they are not. Philosophy was a quasi-priesthood, a vocation. Now it's just a job. Philosophy of old was spread among the professions, the idle rich, etc. Now it's confined to philosophy professors.

Second, how did philosophy lose its way? It became part of the university. That removed philosophers from "modern life." (I wonder where the philosophy professors live who don't: pay taxes, have illnesses, worry for their children, hold political views, fall in and out of love, get divorced, give to charities, etc. Maybe it's North Texas.) In the good old days, lots of people with different interests were philosophers, but after the 19th century they all became academics. lost their virtue and their connection with human concerns. That's the story.

Unlike the Texas philosophers, I am loathe to defame the integrity or selflessness of contemporary philosophers. I have met a few really vile ones, but mostly they have seemed pretty ordinary folk on moral dimensions. But I am not so sure that philosophers of old were selfless and notably different in integrity from their contemporaries. It reads to me as if the Texans have been taking The Apology as the common standard of philosophers before philosophers became professors. Was Aristotle, who left a contentious democracy to educate the mad son of a monarch, selfless? Was Plato, the Athenian aristocrat, selfless? Moving up, what was selfless about Leibniz—did he sacrifice himself in some way for others? Few characters in intellectual history seem less selfless or charitable than Hobbes and Newton, who saw personally to the mutilation of coin clippers. Integrity (and courage)? You won't find it uncompromised in Locke, who contributed (albeit on tolerance) to the Fundamental Constitution of Carolina, an oligarchy ruling over indentured servants that violated both letter and spirit of Locke's 2nd treatise—which treatise Locke made sure not to publish while he lived.

There are lots of examples of 20th century philosophers who acted with selflessness and integrity. Bertrand Russell, who went to prison over his opposition to World War I; David Malament, who did the same over his opposition to the Vietnam War; Paul Oppenheim and Carl Hempel, who helped Jews out of Germany during the Third Reich; Albert Camus, who was part of the French underground. Philosophers not engaged with modern life? Read Philip Kitcher, read Daniel Dennett's more recent works, read just about anything by Peter Singer. Are there no 20th century philosophers who were not professors? Alan Turing was one of the most influential philosophical writers of the 20thcentury—among other things of course. He held an academic position only in the last years of his life. Camus was a journalist. Paul Oppenheim was a businessman. John von Neumann, who stimulated both the philosophy of quantum theory and computation, was a mathematician. Russell spent most of his career outside of the academy. Lawrence Krauss, a physicist, is a metaphysician as well.

What is true is that as universities spread and secularized, a lot more people became "philosophers" and a lot of them are very ordinary people with ordinary minds. The same is true of lots of disciplines I expect, say physics.

What is the author's remedy? Simple: philosophers should get out of universities. The authors teach at the University of North Texas.

Causal Decision Theory and Conditioning: a Primer

Standard Savage decision theory as well as Richard Jeffrey's alternative, address a normative problem for an odd doxastic condition. an agent fully believes:

- a set of all of the available, mutually exclusive actions;
- a set of exhaustive and mutually exclusive possible states of the world;
- a set of consequences—outcomes—of each possible state of the world/action pair.

and the agent:

- Has coherent degrees of belief in the possible states of the world;
- Has utilities (or in Jeffrey's version, desirabiities) for the outcomes.

The normative question is which action the agent ought to take. The answer offered is the action, or one of them, that maximizes the expected utility, where the expectation is with respect to the degrees of belief in the states of the world.

From an ideal Bayesian perspective, what is essential is the distinction between actions and outcomes and their costs or values. The ideal Bayesian knows which actions have the maximal expected utility. The states of the world are gratuitous. Followers of Savage, or Jeffrey's in effect assume the agent only obtains the expected utilities by calculating them using the specified states of the world and probabilities of outcomes, given the various possible states of the world and actions.

What is odd is that no epistemological problem is considered about how an agent knows, or could know, or rationally assess, the possible states of the world and their probabilities, the possible actions, or the probabilities of outcomes effected by alternative actions in the several possible states of the world. a thorough subjectivist such as Jeffrey would answer these questions: all that is relevant are the agent's degrees of belief about actions, states of the world,and outcomes and their desirabilities. Epistemology reduces to observing, Bayesian updating, and rather trivial computation. Be that as it may, or may not, causal decision theory considers two kinds of complications.

1. The agent believes that the action chosen will influence the state of the world.
2. The agent believes that the state of the world will influence the action chosen;

This is already a conceptual expansion for the agent, to include causal relations and probabilities of actions.
In case 1, how should the agent take account of the belief that the choice of action will be influenced by the state of the world? For simplicity, first assume the outcome is a deterministic function of the action, a, and the state, s, of the world, and the utility is U(o(a, s) where o is some function actions and states.

Proposal 1: Calculate the expected utility for each action as the sum over states of the world of the utility of each action in that state of the world multiplied by the probability of that state of the world given the action:

(1) $Exp(U(a)) = \Sigma_s U(o(a,s)) Prob(s \mid a)$

In Savage theory the last factor on the right hand side of (1) and (2) is just Prob(s)

One “partition question” concerns whether the action that maximizes utility is the same depending on how the set of states is “partitioned.” Let S be a variable that ranges over some finite set of values, s1,…,sn. a coarsening of S is a set S1 = {{s1 v..v sk}, {sk +1 v …v sm},….{sm v…v sn}}, etc. a refinement is the inverse.

Coarsening can change the probability of an outcome on an action. Let S = {s1, s2, s3} and suppose S’ is a coarsening of S to {(s1 v s2), s3}. For all outcomes o and actions a, let o and a be independent conditional on s1 and likewise on s2 and s3, but S not be independent of A. Then for any outcome in O:

P(O | a, (s1 v s2)) = P(O, | (a,s1 v a,s2) = P (O, a, (s1 v s2)) / (P(a,s1 v a,s2)) =

P((O, a, s1) v P(O, a, s2)) / (P(a,s1 v a,s2)) =

P(O, a, s1) + P(O, a, s2) / ((P(a,s1) +P(a,s2)) =

[P(O | a, s1) P(a, s1) + P(O | a, s2)] / ((P(a,s1) +P(a,s2)) =

[P(O | s1) P(a, s1) + P(O | s2) P(a, s2)] / ((P(a,s1) +P(a,s2)) =

(P(a) [P(O | s1) P(s1 | a) + P(O | s2) P(s2 | a)]]) / (P(a)((P(s1 | a) +P(s2) | a)) =

[P(O | s1) P(s1 | a) + P(O | s2) P(s2 | a)] / (P(s1 | a) + P(s2) | a))

The probability distribution of 0 given the state s1 v s2 in S’ varies as the conditional probabilities of s1 and, respectively, of s2 vary with the value of A they are conditioned on, and O and A are not independent in S’ but they are independent—by assumption—in S.

For case 2, the results and the argument are similar. The general point is an old one, Yule’s (on the mixture of records).

The partitioning problem does not apply to Savage’s theory—it makes no difference how the range of possible state values are cut up into new coarsened variables.

So decision theory when the actions influence the states or the states influence the actions is up in the air—the right decision depends on the right way to characterize the states. Various writers, Lewis, Skyrms, Woodruff and others, have proposed vague or ad hoc or infeasible solutions. Lewis proposed to chose the most specific “causally relevant” partition, which I take to mean the finest partition for which there is a difference in elements of the partition in the probabilities of outcomes conditional on actions. Skyrms objects that this is often unknowable, and proposes an intricate set of alternative conditions, which Woodruff generalizes. The general strategy is to embed the problem in a logical theory of conditonals, and entwine it with accounts of “chance”and relations of chance and degrees of belief, e.g., the principal principle. The general point is hard to extract.

When states influence actions Meek and Glymour propose that there are two theories. One simply calculates the expected values of the outcomes on various actions as with Jeffrey's decision theory, the other assumes that a decisive act is done with freedom of the will, represented as an exogenous variable, that breaks the influence of the state on the act.

Appealing as the second story may be to our convictions about our own acts as we do them, or deliberate on what to do, it is of no avail when the actions influence the states, not vice-versa. For that case, one either knows the total effect of an action on the outcome, or one doesn't, and if one doesn't, there is nothing for it except to know what the states are that make a difference. One would think serious philosophy would have focused then on means to acquire such knowledge. One would be wrong.

Monday, August 25, 2014

Review of Philosophy of Science, 81, July, 2014

This issue of *Philosophy of Science* contains some good, some bad, some odd. It gives evidence that methodology in philosophy of science is pretty much in the doldrums or worse, while good work is being done producing economic models for various ends.

Brian Skyrms, Grades of Inductive Skepticism

Reject.

This is a very brief rehash of some history of probability, coupled with some remarks on ergodic probabilities, remarks that go nowhere. The piece seems oddly trivial and unworthy of its distinguished author. One has to wonder why it was published—or submitted. Hypothesis: The author is eminent and a colleague of the editors. That sort of thing has happened before in *Philosophy of Science*, although not that I can think of under the current editors. But one of the things colleagues should do for one another is discourage the publication of stuff that is trivial or bad in other ways.

Ben Jantzen,

Piecewise versus Total Support: How to Deal with Background Information in I

Accept.

Likelihood has an apparent problem. Suppose you are weighing hypotheses h1 and h2. You know b. You learn e. Should you compare h1 and h2 by p(e | h1, b) / p(e | h2, b) or by p(e, b | h1) / p(e, b | h2)?

Which hypothesis is preferred may not always be the same on the two comparisons. Jantzen makes the sensible suggestion that which to use depends on whether you are asking about the extra support e gives to h1 versus h2 in a context in which b is known, or whether you are asking about the total support. Jantzen's point is not subtle, but the paper is well done and the examples (especially about fishing with nets with holes too large) are illuminating.

Which reminds me of a deeper problem with likelihood ideas that seem not to be much explored: *likelihood doctrine seems to imply instrumentalism*.

Likelihood arguments are used not just to compare hypotheses but to endorse hypotheses, e.g., via maximum likelihood inference. Consider two principles:

1. Hypotheses addressing a body of data should be preferred according to the likelihood they give to that data.
2. A hypothesis should not be endorsed if it is known that there are other hypotheses that are preferred or indifferent to it by criterion 1 above, especially not if there is a method to find such alternatives .

If the data is finite, the hypothesis just stating the evidence has maximum likelihood. So some additional principle is required if likelihood methodology is to yield anything more than data reports. The hypothesis space must somehow be restricted.

Try this:

3. Only hypotheses that make predictions beyond the data are to be considered.

So suppose there are data e1…en and consider some new experiment or observation e not in the data but for which "serious" hypotheses explaining e1…en gives some probability to the outcomes. Let the outcomes be binary for simplicity and so h gives the probability to be is P(e | h). Consider the hypotheses:

e1&…&en & argmax$_{<h,>}$ (P(e | h) if argmax$_{<h,>}$ (P(e | h) > argmax$_{<h,>}$ (P(~e | h) ,and e1&…&en & argmax$_{<h,>}$ (P(~e | h) otherwise

This hypothesis meets condition 3 and gives e (or ~e) a likelihood at least as great as any alternative hypothesis.

Ok, try this:

4. Only hypotheses that make an infinity of predictions are to be considered.

But the stupid pet trick above can be done infinitely many times. So try this

5. The hypotheses must be finitely axiomatizable.

Still won't do, as (I think) an easy adaptation of) the proof in http://www.jstor.org/stable/41427286 shows.

Lina Jansson

Causal Theories of Explanation and the Challenge of Explanatory Disagreement

Reject

Both the thesis and the argument of this paper are either opaque or weird; it is difficult to see the warrant for publishing. Her stalking horses are "causal accounts of explanation." On Streven's account, causal asymmetry is why X explains Y rather than the other way round—Dan Hausman had that idea earlier; on Woodward's account, X causes Y but Y does not cause X implies that a manipulation of X changes a manipulation of Y, but not vice versa. So far as I know, neither of them claim that all explanations are causal explanations. But a lot of them are.

Jansson's argument seems to be as follows:

Leibniz held that Newton's gravitational theory was not a causal explanation, because causal explanations require mechanisms and no mechanism was given for gravitational attraction. She reads Newton as "causally agnostic" about his laws, which seems to me a very long reach. He was agnostic (publicly) about the mechanisms that produce the laws, but not that the laws imply causal regularities: drop a ball and that will, ceteris paribus, cause it to take up a sequence of positions at times in accordance with the law of gravity. But suppose, for argument, she is right, then what is the argument?

She writes: "Put simply, the problem of understanding this debate from a causal explanatory perspective stems from the reluctance, on both sides, to take there to be a straightforward causal explanation given by the theory." And, a sine qua non of a correct account of explanation is that it be able to "understand the debate. "

There is this oddity about universal gravitation and causation. If I drop a ball it causes the ball to fall, the ball's falling influences the motion of Mars (instantaneously on Newton's theory), and the change in the motion of Mars influences the course of the ball, also instantaneously. Immediate feedback loop. But Mars influence doesn't determine the position of the ball after I drop it, and the position of the ball after I drop it doesn't cause my dropping it.

Anyway, her point is different. Here is the form of the argument.

Accounts S and W say Newtonian gravitational theory is causal.

Neither the creator of the theory nor its most prominent critic unequivocally said it was causal.

Therefore accounts S and W are false (or inadequate, or something).

Parallels.

A: Chemical changes involve the combination or releases of substances made up of elements.

Lavoisier said combustion involves combination with oxygen.
Priestley said combustion involves the release of phlogiston

Therefore A is false.

The theory of probability specifies measures satisfying Kolmogoroff's axioms.

Bayesians say probability is opinion.
Frequentists say probability is frequency

Therefore the theory of probability is false.

Jansson's "methodology" assumes that concepts of causation and explanation never change, and that historical figures are always articulate, and never make errors of judgement in the application of a concept, and that if some historical figure would only apply a concept under restrictive circumstances (e.g., no action at a distance), an account of the concept must agree with that judgement or posit a new concept. Individuation of concepts is a vague and arbitrary matter—are there the concept of causality, Leibniz's concept of causality, Newton's concept of causality, etc.? On her view, so far as I can see, for every sentence about causal relations, general or specific, about which some scientists sometime have disagreed, two new concepts will be needed. Not much to be learned from that.

Robert Batterman and Colin Rice
Minimal Models
Revise and resubmit
Another essay on explanation (will philosophers of science *ever* let up on this) whose exact point is difficult to identify.
"We have argued that there is a class of explanatory models that are explanatory for reasons that have largely been ignored in the literature. These reasons involve telling a story that is focused on demonstrating why details do not matter. Unlike mechanist, causal, or difference-making accounts, this story does not require minimally accurate mirroring of model and target system.
We call these explanations *minimal model explanations* and have given a detailed account of two examples from physics and biology. Indeed, minimal model explanations are likely common in many scientific disciplines, given that we are often interested in explaining macroscale patterns that range over extremely diverse systems. In such instances, a minimal model explanation will often provide the deeper understanding we are after. Furthermore, the account provided here shows us why scientists are able to use models that are only caricatures to explain the behavior of real systems."

The idea seems to be that there are theories that find features and relations among them that entail phenomenological regularities, no matter the rest of the features of a system, and no matter whether the features in question are exactly exemplified in a system. There are two examples, one from fluid dynamics, the other Fisher's opaque explanation of the 1:1 sex ratio in many species based on the equal effort required to raise males or female offspring, but the differential average reproductive return to raising males if females are in excess or raising females if males are in excess. I don't understand the fluid dynamics model, and Fisher's requires a lot of extra assumptions and ceteris paribus clauses to go through, (grant the equal cost of rearing male and female offspring but imagine that one male can fertilize many females and there is a predator that prefers males exclusively) but never mind.

What I don't understand about this paper is why most theories in the physical sciences don't satisfy B and C's criteria for a minimal model. Thermodynamics? The details of the molecular constitution of a system are largely ignored. Relativity? It doesn't matter whether the system is made of wood or iron, the Lorentz tranformations still hold; it doesn't matter how the light is generated, its velocity is still the same. Newtonian celestial mechanics? Doesn't matter that Jupiter is made of gas, Mercury of rock, and Pluto of ice, still the same planetary motions. Even theories that probe into the internal structure of a system are minimal with respect to some other theories. Dalton appealed only to masses of elemental particles—that, and a few assumptions yields the law of definite proportions. Berzelius added electrical forces between atoms, which were gratuitous for deriving definite proportions.

What is not clear in this paper is how B & C intend to distinguish between minimal models and almost every theory that shows a set of features, individual or aggregate, or approximations to such features, and related laws, of a kind of system suffice for phenomenological relations. That is what physical theories generally do. Their fluid flow example almost suggests that all that is required is an algorithm that generates the phenomena from (perhaps) measurable features a system. So, considering that example, the authors might have asked: when is an algorithm for generating the phenomena an explanation of the phenomena? They did not.

Dean Peters

What Elements of Successful Scientific Theories Are the Correct Targets for "Selective" Scientific Realism?

Revise and resubmit

Peters' essay is useful in two respects. First, it treats the question in the title as turning on this: what parts of the data confirm what parts of a theory? That adds a little structure to the philosophical discussions of realism. And, second, it provides a succinct critical review of bad proposals to answer the question. Peters' has his own answer, which is not obviously useful. Here it is:

"So, to pick out the essential elements of the theory under the ESSA, start with a subtheory consisting of statements of its most basic confirmed empirical consequences or perhaps its confirmed phenomenological laws. These, after all, are the parts of a theory that even empiricists agree we should be "realists" about. Further propositions are added to this subtheory by a recursive procedure. Consider any theoretical posit not in the subtheory. If it entails more propositions in the subtheory than are required to construct it, tag it as confirmed under the unification criterion, and so add it to the subtheory. Otherwise, leave it out. When there are no more theoretical posits to consider in this way, the subtheory contains the essential elements of the original theory."

The proposal as developed is insubstantial: "Consider any theoretical posit not in the subtheory. If it entails more propositions in the subtheory than are required to construct it" – what does "required to construct it" mean?

In criticizing other proposals, Peters appeals to logical consequences, and proceeds with a distinguished set of "posits"—i.e., axioms. Hold him to the same standard. Theories can be axiomatized in an infinity of ways. We need an account of the invariance of the result of the procedure—whatever it is—over different axiomatizations, or an account of "natural axiomatizations" and warrant for using them exclusively. The work of Ken Gemes and Gerhard Schurz is relevant here. So it seems to me that Peters has an idea—conceivably ultimately a good idea—that he did not do the work to make good on.

Roger DeLanghe

A unified model of the division of cognitive labor

Accept

This is a very nice essay providing a simple economic model in which there are balancing incentives for scientists to adopt and contribute to an existing theory or to propose a new one. Lots that might be done to expand the picture for more realism, and it would be nice if those pursuing Kitcher's original idea assembled some relevant data.

Marius Stan

Unity for Kant's Natural Philosophy

I have no opinion about this essay, which is on how Kant might have sought, although he did not, synthetic a priori grounds for Euler's torque law. Nor do I see why anyone should care. Clearly, some do.

Carlos Santana

Ambiguity in Cooperative Signaling
Accept

This well argued and lucid essay shows that there is a model in which agents with ambiguous signaling (under replicator dynamics) invade a population of unambiguous signalers, but not vice-versa. Despite the considerable empirical evidence the author (a graduate student at Penn) gives for the insufficiency of other explanations of the frequency of ambiguity in human and animal communication, I am worried by the following thought. The evolution of language—or at least signaling-- we expect to have gone from the very ambiguous to the more precise. That is what syntactic structure and an expanded lexicon afford. So if signaling by ambiguous strategies cannot be invaded by signaling by "standard" (i.e., perfectly precise) strategies, how did more precise, if still ambiguous in some respects, signaling systems evolve? It strikes me that the author may have proved the wrong result.

Philosophy Reviewed

Saturday, August 16, 2014

The Fortress of Metaethics: Reviews of Thomas Scanlon, What We Owe to Each Other and Being Realistic about Reasons.

Metaethics is about what ethical claims mean, how they can be "justified," and how ethical reasoning ought to be conducted. Thomas Scanlon's writing on metaethics has become a verbal icon for the enterprise. Scanlon now has two books elaborating his views, *What We Owe to Each Other* and *Being Realistic about Reasons*. The first was reviewed with applause in literary venues where philosophy is seldom seen, and one can only expect the same of the second. Each book is a theoretical disappointment—no, the second is a disaster--the first from lacunae the second tries to fill, the second from the filling.[1]

In *What We Owe to Each Other* Scanlon's stalking horse is utilitarianism. The many variants of utilitarianism share this much: they are voting theories, and so is Scanlon's alternative, absent some key features. In utilitarianism, every sentient creature, or at least every human, has a stock of interests. Properly scaled, those are voting stocks. Anyone's action predictably affects some of the interests, or the well-being, of some creatures. An action is permissible only if, among the alternative available actions, it maximizes some aggregate of the interests of all who may be affected by any of the available actions, so far as the actor can estimate. That is the utilitarian schematic for how the affected stocks are to be weighed in moral assessment of action. The specifics are contentious. How are pains and pleasures and sorrows and joys to be compared across persons, let alone persons and other animals, so that they can be aggregated? That issue aside, suppose we have a number for each human state or well or mal being that takes account the diverse interests of each human being. Should one try to maximize the total, the average, the median well-being over all persons? Minimize the variance? Maximize something under a constraint (e.g., a bound on the variance, or the Difference Principle)? Count only changes to an individual's interests that are above a certain threshold? Order interests by type, higher type trumping lower, as one might read Mill to suggest? The utilitarian literature from Bentham on has this virtue: it takes these questions seriously and offers answers, not the same answers, of course. These are philosophers after all.

Scanlon's theory is one piece of a different voting theory. For any action one might do, the deliberator and others may have reasons, conscious or not, why it should or should not be done. The best of those reasons is decisive. Reasons are not to be aggregated; the best reason wins, no matter how few people have it. Which reasons count?

Scanlon offers only this in general: We should seek to act only in accord with principles that other people could not reasonably reject if they too sought principles others could not reasonably reject. A reason has moral bearing only if it is an application of such a principle.

One catch to Scanlon's general schema for moral principles is "reasonably," which in this context is particularly flabby. We have pretty clear notions of reasonable views in mathematics, less definite notions in science, but in ethics "unreasonable" is more a slur than a methodological complaint. Does Scanlon imagine a negotiation about which principles meet his criteria, or a social survey to find nearly universal principles (they would be few), or does he imagine principles that someone sincerely thinks others ought to share? In the latter case, utilitarian principles have a vote. No one thinks that the only reasons I can have for or against an action must be exclusively about me. Indeed, if by the proper weighting of principles and reasons utilitarian principles and reasons are the best, then Scanlon is committed to utilitarianism, one step back.

So what then are the principles of ordering of reasons or principles, and why? Scanlon does not say, only, examples aside, that the best reason wins, winner take all. What then individuates reasons: if your action will impoverish me and cause me an illness besides, is that one reason or two? Scanlon doesn't say, but in his voting scheme it matters. The basic elements of the theory are unspecified: by what clear principles are reasons to be weighed, and why, what counts as a reason, how are reasons individuated? Almost all you might want in a theory is unspecified. One would not want election rules so vaguely posted.

The book has any number of appeals to quotidian examples where one would not consider global consequences, or global good, or aggregate harms and benefits. There are old saws: If a broadcasting engineer is painfully and continuously injured by continuing shock during the broadcast of a sports game—say Argentina versus Germany in the World Cup--and the only way to stop his agony is to interrupt the electric current that is necessary to broadcast the game, shouldn't that be done—aren't the enormous sum of annoyances to the passions of millions of soccer fans outweighed by the one engineer's acute pain? But what would be said in this kind of case by versions of utilitarianism in which what counts has thresholds, or versions in which there are layers of goods and bads that trump others? Should we not consider the number of people who might be killed in mad rages and riots were the television to go off in the midst of the game? In any case these examples can be bought cheap by either side. Being murdered is worse than not having children. If you met a person with a remarkably infectious, uncurable disease that renders all who catch it permanently sterile but is otherwise asymptomatic, and you could not capture him and isolate him before he spread it to the entire human population, would you kill him if you could?

If the aim is to articulate standards for correct norms, why trifle with mundane examples, unless one is doing moral sociology from an armchair? That question takes us to Scanlon's more recent book, a defense of putting weight on such examples, and, more broadly, a defense of metaethics as a serious enterprise. Scanlon does not fill all the gaps his first book left, but he addresses two major ones. The claims of *Being Realistic about Reasons*come down to these: there are true normative principles that are not just about the best means to ends, and there is a method for discovering them. What is the argument?

Scanlon claims that there are various "domains" of inquiry. Each domain has its own standards for seeking the truth in its domain. Mathematics has its methods, empirical science has its methods, and, lo, normative matters, moral matters in particular, have theirs. One domain may not contradict, or, presumably, undermine, the methods and conclusions of another. They are separate empires, contractually at peace. This last I call Scanlon's Rule. He continues with implausible parallels between set theory (Scanlon was a logician in his youth) and normative reasoning. Just as set theorists may debate and disagree about, say, Zorn's Lemma, so metaethicists may debate and disagree fine points of the correct normative principles. Can one seriously think that the reasoning in ethics and metaethics has the rigor of mathematics? I can't, and I doubt Scanlon does. Scanlon's thesis is that they share the same style, the same form viewed with sufficient abstraction. So, with sufficient abstraction, do science and Cargo Cults. The intellectual legitimacy of metaethics needs a better bolster.

The crucial point is Scanlon's Rule. Scanlon's Rule is pure defense, a paper wall to keep out critics. Mathematics and logic may be immune to contradiction from physics (although Hilary Putnam once thought otherwise, and presumably Mill would have allowed the possibility, and certainly the *status* of geometry has been altered by physics—and the status of metaethics is what is at stake here), but ethics is not immune to contradiction from other "domains." Is religion a domain? Theological reasoning is more like metaethics than is set theory, and theology most definitely intrudes on ethics and on metaethics. Empirical science may not directly contradict normative claims, but it can surely undermine them. Once upon a time it was widely thought that there are particularly evil people, sorcerers and witches, who had made contracts with the most evil entity, Satan, and should be killed. Science has convinced the civilized that there are no witches and no sorcerers and no Satan. Once upon a time, it was thought that living beings have a superphysical constitution, that their chemicals are not the ordinary, “inorganic” stuff, and that living beings possess an unphysical “vital force” that guides evolution. Science has convinced us (Tom Nagel perhaps aside) otherwise. Science bids fair to do the same with the Will, and Autonomy, and Agency, and as, and if that more fully comes about, the idea of true, moral principles will go the way of true principles of witchcraft.

So what about the methods of ethics? Scanlon’s is “reflective equilibrium,” I think first proposed in Rawl’s essay “Outline of a decision procedure for ethics.” Rawls imagined a panel of moral experts (much of his essay is about the qualifications for membership) who report on the moral statuses of sundry actions. The ethical theorist takes their pronouncements as data—putative moral facts—and attempts to form a general theory that accounts for them. Rawls allowed that on reflection one might reject a few of the experts’ decisions if accounting for them required excessive complexity in the theory, arbitrary exceptions and so on. The procedure came to be called “reflective equilibrium.” That is Scanlon’s method, with the panel of experts replaced by one’s own judgements and the judgements of those whose ethical perspicuity one respects. His explanation of the reliability of its data sounds very much like Descartes “clear and distinct ideas”:

“In order for something to count as a considered judgment… It is necessary also that it should be something that seems to me to be clearly true *when I am thinking about the matter under good conditions for arriving at judgments of the kind in question*.” Scanlon, T. M. (2014-01-06). *Being Realistic about Reasons* (p. 82). Oxford University Press, USA. Kindle Edition. (Italics are Scanlon’s)

I have no doubt that some very thoughtful people, no doubt Scanlon himself, form their moral views in this way. I even think it’s a good way. But I have no doubt, either, that in many other “domains” something similar is often followed. It is general and vague enough to characterize both the process of Islamic jurisprudence and the quasi Bayesian process often at work in science in which data are thought to come with probabilities of error and the sufficiently low posterior of a datum conditional on a hypothesis of sufficiently high probability is reason to reject the datum, or even to reject an entire set of measurements. But these examples are exactly the problem. In Bayesian statistics one can prove that under specified, general assumptions, application of Bayes rule converges to the truth. One can do the same for modifications, perhaps even considered variants of the one I suggest off-hand above. In statistical estimation and machine learning (the latter of which Rawls, in keeping with the opinion of his time, announced was impossible) proofs are given that under very general assumptions search methods converge on the truth, and methods are provided for testing the assumptions. Nothing like that can be done for Islamic jurisprudence, and nothing like that can be done for Rawl’s decision procedure for ethics or Scanlon’s variant. That one can apply a vaguely specified procedure in a domain is no argument, no evidence, that the procedure finds the truth in that domain, or that there is any truth there to be found.

The least attention to the world shows that the range of considered moral judgements is incompatible with any unified theory of morality. Scanlon will have to discard many of the moral judgements of most people in the world. He should, but he should not claim that in doing so he is exercising a method for finding truth. Scanlon has only two responses. Those who want to (and do) crucify Christians and behead Jewish journalists and do other atrocities are not in “good conditions” for such judgements; and that ethics and metaethics have their own standards for concluding what is true--outside standards and alien practices, however common, are irrelevant.

There is plenty of work for metaethics to do: systematizing vague strategies of inference—Nozick’s efforts were a good start[i]--finding and recognizing contradictions, figuring out how principles apply in morally difficult cases, contrasting misweighings of moral importance, finding agreements and disagreements in clarified moral perspectives, tempering ethical demands to human capacities, and so on, all without Scanlon’s truth claims. Scanlon’s redoubt is a parochial fortress, impenetrable to the forces of science or to the objections of the world outside and its domains.

[1] I write this and what follows with some regret, since Scanlon was my closest friend when we were colleagues. This blog may lose me a lot of friends.

[i] Nozick, Robert, "Moral Complications and Moral Structures" (1968). *Natural Law Forum.*Paper 137. http://scholarship.law.nd.edu/nd_naturallaw_forum/137

Monday, August 11, 2014

Low Bars: Reviews of Four Semi-Recent Books in Philosophy of Science

I take a dyspeptic look at four recent books, one of which, Kyle Stanford's *Exceeding Our Grasp*, has previously been reviewed with praise in several places. And now, a second, Paul Churchland's *Plato's Camera*, reviewed with praise in *Mind and Machines.*

Paul Churchland, *Plato's Camera*, MIT Press, 2012

As his title indicates, Paul Churchland is a man of big metaphors. He is a man of big ambitions as well, not for himself but for his theory. He thinks that that neuroscience will provide—and is well on the way to providing --a complete logic and philosophy of science. Academic philosophers have missed the boat, or the bandwagon, whichever metaphor you prefer. Neuroscience provides "a competing conception of cognitive activity, an alternative to the "sentential" or "propositional attitude" model that has dominated philosophy for the past 2,500 years." (14) "these spaces [of synaptic weights and patterns of neural activation] specify a set of 'nomically possible worlds…these spaces hold the key to a novel account of both the semantics and the epistemology of modal statements, and of counterfactual and subjective conditionals." (18). "Notably, and despite its primacy, that synapse-adjusting space-shaping process is almost wholly ignored by the traditions of academic epistemology, even into these early years of our third millennium." (13)

A little potted history will put Churchland's book in context. The great philosophers joined theories of mind with theories of method for acquiring true beliefs. For Leibniz and Hobbes and even Hume, logic was the algebra by which the mind constructs complex concepts, or ideas, from simpler ones. George Boole realized that whatever the laws of thought may be, they are not in necessary agreement with the laws of logic. People make errors, and some people make them systematically. Logic, semantics, causality, probability have their relations, the mind has its relations, and the twain shall sometimes, but not always, meet.

Sparked by Ramon y Cahal's discovery of the axon-dendrite structure of neural connections, suggesting that the nerve cell is an information processing unit and the synaptic connection is a channel, in the last quarter of the 19th century avante-garde speculation turned to how the distribution of "excitation" and its transfer among cells might produce consciousness, thought and emotion. Connectionist neuropsychology was born in the writings of Cahal, Sigmund Exner and, yes, Sigmund Freud. Exner, like Freud, was as an assistant to the materialist physiologist Ernst von Brucke, and Freud's neuropsychological speculations from1895 elaborate (one might say exaggerate) lines suggested in Exner's 1891 *Entwurf zu einer physiologischen Erklärung der psychischen Erscheinungen*, both inspired in a general way by Hermann von Helmholtz, with whom Freud once proposed to study. In Freud's *Entwurf einer Psychologie*—still in print in English translation as *Project for a Scientific Psychology*--the neurons are activated by stimuli from the sense organs, or by chemical sources internal to the body. Neurons pass activation to those they are connected with in the face of some resistance, which is reduced by consecutive passage (an idea now called, with historical injustice, the "Hebb synapse") and eventually produce a motor response. Depending on the internal and external stimuli that result from motion, a feedback process occurs which eventuates in a semi-stable collection of facilitations among nerve cells that constitute our general knowledge of the world—what Freud called the "reality principle." The particular neural activations of memory and momentary experience occur within those learned constraints captured by the facilitations. Logic, the subject–predicate logic Freud had learned from Franz Brentano–is at once created (as thought) and realized (as model) by the synaptic connections.

That is pretty much Churchland's theory. There are modern twists, of course—Cajal and Exner and Freud had no computers with which to do simulations or make analogies, and they had a different data set—and Churchland has all sorts of terminological elaborations. But, other than a review of connectionist computing and some modern neurobiology, and of course a host of new metaphors—"sculpting the space" of activation connections and so on, what is new in Churchland's book? What he says: "a novel account of both the semantics and the epistemology of modal statements, and of counterfactual and subjunctive conditionals" as well as a novel account of synonymy and an explanation of scientific discovery and intertheoretical reduction and more. In sum, Churchland shares the aim of the Great Philosophers to produce a unified account of mind, meaning and method, but this time founded on the neuroscience of neural processes rather than on Hume's introspective science of impressions and ideas or Kant's a priori concepts.

Historians and philosophers of science have written reams about how Darwin came to the view that species formed and evolved by spontaneous variation and natural selection, what knowledge and arguments and hypotheses he had available when he embarked on the voyage of the Beagle, what he was convinced of by what he saw in those passages, what the collections and notes with which he returned taught him, what influences his subsequent reading and conversation and correspondence bore. Churchland's explanation of Darwin's discovery can be Bowdlerized but not summarized:

"The causal origins of Darvin's explanatory epiphany resided in the pecular modulations, of his normal perceptual and imaginative processes, induced by the novel contextual information brought to those processes via his descending or recurrent axonal pathways…A purely feed forward network, once its synaptic weights have been fixed, is doomed to respond to the same sensory inputs with unchanging and uniquely appropriate cognitive outputs…A trained network with a recurrent architecture, by contrast, is entirely capable of responding to one and the same sensory input in a variety of very different ways..As those states meander, they provide an ever changing cognitive context into which the same sensory subject-matter, on different occasions, is constrained to arrive. Mostly, those contextual variations make only a small and local difference in the brain's subsequent processing of that repeated sensory input. But occasionally they can make a large and lasting difference. Once Darwin had seen the now-famous diversity of finch-types specific to the *environmentally* diverse Galapagos Islands as being historically and causally analogous to the diversity of dog-types specific to the selectionally diverse dog-breeding kennels of Europe, he would never see or think of the overall diversity of biological forms in quite the same way again. And what gave Darwin's conceptual reinterpretation here the lasting impact that it had on him was precisely the extraordinary explanatory power that it provided…The Platonic camera that was Darwin's brain had redeployed one of its existing 'cognitive lenses' so as to provide a systematically novel mode of conceptualization where issues of biological history were concerned." (191-200).

A lot has gone wrong here. How the output (the realization of explanatory power) "sculpts the space" of neural connectivities anew is unexplained. The "recurrent neural network" and "descending axonal pathways" stuff has nothing to do specifically with Darwin. It could as well be said of the epiphanies of Newton or Einstein or the fantasies of Erich van Dalen. When Churchland wants actually to engage Darwin, he has to step out of the neurological generalities and into the actual history, and he has to appeal to a notion, "extraordinary explanatory power" taken from old-fashioned philosophy of science. And that is because he knows nothing specific about what neural processes took place in Darwin, and nothing about what neural processes constitute the realization of explanatory power, or what about the neural processes themselves distinguishes genius from crank from paranoid. He is not to blame for that, but it shows the impotence of his framework for elucidating much of anything about scientific discovery, let alone for providing guidance to it.

It is the same everywhere with Churchland. He is not to be faulted for want of theoretical ambition. Take the question of inter-theoretic reduction. After whipping off criticisms—the quality of which I have not space to pursue--of various accounts, Churchland offers this:

"A more general framework, G, successfully reduces a distinct target framework, T, if and only if the conceptual map G, or some part of it, subsumes the conceptual map T, at least roughly… More specifically
(a) the high-dimensional configuration of prototype-positions and prototype-trajectories with in the sculpted neuronal-activation space that constitutes T (a conceptual map of some abstract feature-domain) is (b) roughly homomorphic with
(c) some substructure or lower-dimensional projection of the high dimensional configuation of prototype-positions and proto-type trajectories within the sculpted neuronal activation space that constitues G (a conceptual map of some more extensive abstract feature-domain.)" (210-211).

Good. Now does statistical mechanics reduce thermodynamics? Does quantum theory reduce classical mechanics? Or what? Consult prototype positions in sculpted neuronal activation space. I will skip the details of Churchland's account of "homorphisms between sub-structures of configurations of prototype-positions and proto-type trajectories." Suffice that is an ill-defined attempt at a little mathematics, so odd as perhaps to have been whimsical.

About meaning relations, the general idea seems to be that one thinks counterfactually or hypothetically by activating patterns that are neither sensory responses nor exact reproductions of previous activation patterns—not memories, which, less the 'activiation patterns' is precisely Hume's account. Nothing particular is established, and we are left to wonder what constraints on our meandering activations incline us to think that if, necessarily if p then q, then if necessarily p then necessarily q. What distinguishes the hypothetical from the counterfactual, the entertained from the believed, the supposition from the plan, the wish from the fear from the doubt from the conviction--is unexplained, and it seems doubtful that Churchland can do better than Hume on imagination.

When it comes down to it, Churchland does not want to explain propositional attitudes, he wants to do away with them. Some reasons are given in his argument against one propositional attitude, the analysis of knowledge as true, justified belief. He notes the usual Gettier problems but that is not what bothers him. We, and infants and animals, have he says, a-linguistic knowledge. Beliefs are attitudes to propositions and truth is a property of sentences, so to attribute them to much of what we know and other animals know is a category mistake. And so, for much of what is known but is not, or cannot, be said, justification is impossible and to ask for it is likewise a kind of category error.

There is something to this, but only a little. There is implicit knowledge, exhibited in capacities, which someone can have and yet have no awareness of, no thought of. The psychologist evoking the capacity can generally state what her subject implicitly knows. She may even claim to know in a general way how the subject came to know it, and so find it justified and true. Whether such implicit knowledge is a belief of the knower is the hard question. Churchland would I think say not; Freud, who lived on the premise of unconscious beliefs, would have had no trouble allowing it. We have thoughts we never formulate in language—we can think we see a familiar face in a crowd and automatically look again, testing the thought before it takes, even to ourselves, a linguistic form. Evidence of a-linguistic thought is all around anyone who lives with dogs or cats or even a closely watched cow. But I do not see why such thoughts cannot be believed or had with surprise or fear by those entities that have them, why they cannot be the objects of the very attitudes that philosophers call propositional. There is generally a proposition that approximately expresses them even if their possessor cannot formulate it. However this may be, it remains that our thoughts are not on a par. There is a difference between formulating a plan, an intention, and entertaining a possibility, and Churchland's framework has no place for it. Perhaps one could be made, but for that one would have to want to allow something very much like propositional attitudes.

On technical points, the book is a mixture. Lots of things are explained vividly and correctly, some not so much. For example, recurrent networks have a problem with long term memory. A class of algorithms Churchland does not discuss, Long Short Term Memory (S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997) do better. He is a bit weak on biology. Churchland dismisses innateness hypotheses on the grounds that genes would have to specify synaptic connections, and there are billions of those and only 30,000 or so genes. He forgets (I know he forgets, because once I told him) that a person's liver cells and neurons have the same genes but very different forms and functions--cellular form, function and location involve gene expression, and it isn't just one gene-one expression, one protein, one synaptic connection. The combinatorics are enormous. He writes metaphorically of "sculpting activation space" but fails to note that nerve connections are physically pruned—literally destroyed--from infancy to maturity. Remarkably, the book entirely ignores the growing neuropsychological research on predicting an agent's environment from indirect measurements of brain physiology—the very work that comes closest to realizing Churchland's vision.

The real problem with Churchland's book is too long an arm, a lengthy overreach. One can grant the general Cajal-Exner-Freud connectionist framework. It provides a theoretical position from which to do research and that research is prospering. A few professional philosophers have contributed, Stephen Quartz for example with fMRI experiments, and Joseph Ramsey with improvements in fMRI methodology. But decorating the framing assumptions of scientific research in neuroscience with metaphors, accounts of computer simulations, and vacuous applications neither helps with our problems in philosophy of science nor contributes to methods for effectively carrying out that research.

P. Kyle Stanford, *Exceeding Our Grasp*, Oxford University Press, 2012

Banality, Nelson Goodman once said, is the price of success in philosophy. Here is a banality: One cannot think of everything, and if a truth is something one cannot think of, then one will not believe that truth.

That is the fundamental substance of Stanford's thesis, elaborated with brief discussions of some of the philosophy of science literature on theoretical equivalence, underdetermination, and confirmation, and with a more extended discussion of examples in the history of science. More elaborately, the thesis is that historical scientists did not, and could not, think of the alternatives to their theories that later explained their evidence in different ways; so, too, our contemporaies are unable to think of such alternatives that may lurk in Plato's heaven. Hence we should not believe our current theories. The conclusion does not follow. Perhaps one *ought* to believe, of the hypotheses one can conceive and analyze, those best supported by current evidence. The general agnostic will never believe the truth; those who believe on their best evidence and available conceptions at least have a shot. Even so little strategic reflection is not to be found in Stanford's essay.

Much of Stanford's philosophical argument is negative: there are no general characterizations of theoretical equivalence even assuming a definite space of possible data; there are no general theories of what parts of a theory are confirmed by what data. One could apply his argument reflexively: there may be possible characterizations of such relations that have not been thought of, in which case perhaps we should be agnostic about being agnostic about our theories. I don't know if agnosticism is transitive. The rest of his argument consists of historical discussions about what various scientists thought that turned out to be wrong, for example what they thought were the indisputable parts of their theories. Here the absence of any normative theory in the book collides with the historical exegesis: why should we think that various historical figures, Maxwell, for example, were right about what they thought were the indubitable, or best confirmed, aspects of their theories? More than that, Stanford's histories neglect historical stability. Two centuries later, the atomic weight of oxygen is still greater than the atomic weight of hydrogen.

Logic is also neglected in Stanford's effort to make novelty out of banality. Stanford's discussion of Craig's theorem, for example, is odd. He takes it as establishing that a theory has a perfectly observationally equivalent instrumentalist ghost, and of no further significance for theoretical equivalence. But what the theorem establishes is that if there is a recursively enumerable linguistic characterization of the possible data for a theory, then there is an infinity of theories that entail the same possible data. Under mild assumptions, there is an infinity of finitely axiomatizable, logically inequivalent such theories, and there is no logically weakest finitely presentable theory.

Some years ago I attended lectures by a prominent philosopher and by the late Allen Newell. The prominent philosopher went on for two lectures to the effect that some features of cognition are "hard wired" and others not. Having enough of this, Newell asked what the philosopher's laboratory had discovered about which cognitive features are "hard-wired." Flustered, the philosopher appealed to "division of labor" between philosophy and psychology. To which Newell observed privately that if that was the philosophers' labor, psychologists could do it themselves, thank you. And there is the trouble with Stanford's book. It is a lazy effort. If there are theories we cannot think of, or have not thought of, in some domain, and surely in many domains there are a great many, by all means help us find ways to survey and assess them. That is what machine learning is about. Stanford has nothing to say. If we need a reliable means to assign credit or blame among the many claims entailed by a theory, seek for one. Stanford has nothing to say. The main thing he has to say you knew before opening his book.

Sandra Mitchell, *Unsimple Truths*, University of Chicago Press, 2012

Sandara Mitchell's book is more shadow than smoke. Try to catch some definite, original content is like grasping a shadow, but the shadow is always there, moving with your grasp. Mitchell rightly observes that contemporary science proceeds across different "levels," that many relations are not additive (she says not "linear"), that many phenomena, especially biological and social phenomena, have multiple causes, and that much of contemporary science is addressed to finding regularities that are contingent, or impermanent, or not general (she doesn't distinguish these) . One wonders for whom this is news. No one I know. No doubt she gets around more.

She argues for "emergence" rather than "reduction" and proclaims a "new epistemology": *integrated pluralism*. One might hope that this is the definite, original part, but it turns out not to be so.

Epistemology comes in two phases: analyses: "S knows that P" and such; and method: how S can come to know that P, and such. There is no concrete thought in this book on either score that is helpful, either to philosophy or to science. Modern systems biology and neuropsychology have lots of problems about "high dimensional, low-sample size" data. She has nothing to offer. Social epidemiology has a hoard of problems about measurement, sampling and statistical inference. She has nothing to offer. Cancer has complex interactive causes hard to establish, and so do lots of social and cognitive phenomena. She observes that there are problems, but has nothing helpful to offer.

Mitchell's discussion of emergence and reduction is a bit bewildering. On the one hand, she allows that no one seriously thinks we are actually going to deduce social patterns from facts about fundamental particles—and if some should try, let them go to it but don't pay them. So there is no methodological issue, only a metaphysical one. On the other hand, she does not dispute that, at the basis of nature, it's physics. She isn't arguing for any transcendent powers. So what's left? Apparently only this: one language can't express everything, so no language for physics can express everything. Something will be left out. She offers no candidates for the omitted, but suppose she were right. Suppose for any physical theory there are aspects of the physical world that theory does not capture—not even logically, let alone practically. Proving, rather than merely asserting, as much would be an impressive achievement merely as a theoretical exercise, but what's the point for "integrative pluralism"? I see no implication whatever for the conduct of science. Whether we think there is a theory of everything is possible or not, the scientific community will still measure the large and the small, try to separate phenomena into multiple aspects, look for mechanisms and try to separate their components, suffer with interaction, with the limits of predictability, computational complexity and the rest. Makes no difference to any of it whether the language of physics is finally complete or finally completable.

To judge from the blurb on the book jacket, scientists may like reading this stuff, but if so that can only be because it is an aid to their vanity, not to their science.

Bill Harper, *Isaac Newton's Scientific Method*, Oxford University Press, 2012.

Much of this book is about another, Books I and III of the Principia. Harper details, almost lovingly, the theorems from Book I and how they are used in the argument for universal gravitation in Book III, and on that account the book is worth reading—with a copy of the Principia to hand. But the question of Harper's book is : *What was Newton's method?* It was more than theorems.

Any reader of the first pages of Book III should get the general idea of Newton's argument. Starting with Kepler's laws and using theorems of Book I that are consequences of the three laws of motion, Newton proves that for each primary in the solar system with a satellite, there *exists* an inverse square force attracting the satellite to its primary. He then shows that the motion of the moon can be approximately accounted for the combination of two such forces, one directed to the sun and one directed to the Earth. He then engages in a hypothetical, or suppositional exercise, counting the acceleration the moon would have at the surface of the Earth. Using experiments with pendulums, he shows that the acceleration of the bob is independent of the mass and equals the suppositional acceleration of the moon at the Earth's surface, and infers that the acceleration produced in one body by another is proportional to the mass of the acting body and independent of the mass of the body acted upon. Applying his rules of reasoning, he identifies the force of the Earth on the moon with terrestrial gravity, and likewise the forces that solar system primaries exert on their satellites, and concludes that gravitational force is universal.

There are lots of details, many of which Harper carefully goes through. But that leaves open the question at issue, what is the general form of Newton's method? Newton expresses the same themes of "general induction from the phenomena" at the end of the Opticks but we still want a general, precise account of the method, whatever it is. How would we apply it or recognize it in other cases? I essayed an account I called bootstrapping to which various philosophers have offered objections I will not consider here. Others, Jon Dorling for example, have offered reconstructions. Harper discusses mine and rejects it citing the various criticisms without further assessment. That's ok, but what we should expect is an alternative. Harper's only suggestion is that Newton's hypotheses are "subjunctive." We are left to wonder how that helps. Is Newton's method "subjunctive bootstrapping," whatever that is, and, to engage the subjunctive, what *would* that be and how *could* we recognize it or apply it in other cases?

Harper resorts to vagaries, the substance of which is ostensive: Newton's method is *like that*. We *should* expect more from philosophical explication than demonstratives.

-