The Metaphorical Constitution of AI

1. Introduction

Recent research into the genealogy of concepts central to statistics and artificial intelligence has uncovered the central role of metaphor in the birth of these concepts. The significance of metaphor is not of mere historical interest to statistics. It is critical to ethical debates throughout the history of statistics, from eugenics to the crisis of replicability in science to the harms of AI bias. This paper puts in dialogue, on the one hand, the critical role of metaphor within statistics and AI, one of the today's foremost ethical concerns, and on the other, the work of Paul Ricoeur, which delved deeply into both metaphor and ethics, to explore what we can learn from the history of metaphor within AI.

Metaphor comes into play in statistics in relation to the role of models. Models play very different roles within the work of different statisticians and engineers, depending on their approach to probability. Within the more descriptive approach to modeling uncertainty, the abstraction from reality achieved by models is only possible by framing features of reality as if they were the outcome of a random lottery. The use of the term *as* here indicates the crucial role of metaphor.

When we create a model of propensity to recidivate for use in judicial sentencing, and the features are crime type and number of previous crimes, we are momentarily ignoring all other features by assuming they are randomly distributed. But what does it mean to be randomly distributed? Contrary to claims by contemporary philosophers of statistics, randomness cannot be found in reality, either in the objective world of things, or in the mind. If it could, the modern idea of random variation would likely have appeared at other places or times. As it is, our notion of random variation was invented in 17th century Europe. Random variation can only be generated using artificial means – dice, wheels, etc. Our notion of random variation relies upon a metaphorical lottery – transferring the sense of a lottery to better understand other domains of human experience characterized by uncertainty - that was birthed by Jacob Bernoulli in the late 17th century.

This metaphor, like all metaphors, sheds new light on domains thanks to a transfer of meaning. In the case of probability, Fortuna, or fate, was previously held to play a central role in many areas of human affairs. Early Enlightenment thinkers were keen to ground human affairs not on overarching structures of fate and hierarchical orders of being from above, but on recognitions of the good sense of respectable persons, particularly those observed entering into contracts of chance, such as commercial contracts. By resolving the ambiguity of complex human affairs, good sense was held to be the province of respectable persons and not teachable – the original sense of probable referred to persons and not degrees of evidence. By replacing the concept of Fortuna with a metaphorical lottery, the early probabilists attempted to present the good sense of respectable persons for broader understanding and enlightenment. Probability provided a "calculus of good sense" in the words of Laplace.

From the outset, however, the life of the metaphorical lottery was in tension with its death. As will be seen below, the tensions encountered by classical probabilists with those who rejected their counsel prompted many probabilists to reify the metaphor of a lottery, reduce uncertainty to quantifiable doubt devoid of any ambiguity into the relevant features of a domain, and command certainty on the foundation of the objectivity of randomness. With the death of the metaphorical lottery came the rise of a narrower, more formalist and prescriptive approach to probability, and the ethical crises in its wake culminating in the harms of AI.

This history of metaphor within probability and statistics finds much in common with Ricoeur's theory of metaphor. Two questions, however, are raised for Ricoeur from this history. First, why does Ricoeur's work on metaphor, which stressed the generation, life and death of metaphor as a linear trajectory, not also account for a concurrent tension between living and dead metaphors in communities of practice? Second, why is metaphor not addressed in Ricoeur's ethics? Perhaps related to this second question, we can ask why a fairly obvious case of metaphor, the social contract, is criticized in Ricoeur's ethics for its non-literal meaning? This paper suggests preliminary answers to these questions, in the hopes of soliciting better answers from others.

2. Probability as a Metaphorical Lottery

Classical Probability: Grounding Social Progress in Universal Reasonableness

Probability in its classical incarnation was essentially a concept in tension between a broader, descriptive probability of the early probabilists that encompassed both qualitative ambiguity and quantifiable doubt, and a progressively narrower, formalist and prescriptive probability that reduced all error to random variation.

The critical backdrop to the 17th century emergence of classical probability, according to the pioneering work of Lorraine Daston, was the conviction that civic and commercial order comes not from a transcendent order reflected at different levels of being and enforced by the Church, but from the mutual recognition of men as reasonable through contracts.[1] The ambition of the early mathematicians of probability was not to formalize models free of human bias, but rather to uncover and describe in formal terms the unconscious intuitions of reasonable men. Such a recognition of the universal reasonableness of men, it was believed, would ensure a new, secure basis for social order free of the conflict and skepticism that defined the 17th century.[2]

This discovery was ultimately made, as Herbert Weisberg argues, by taking one such aleatory practice – the lottery – as a metaphor for other such practices[3]. Historians have long struggled to explain why the early probabilists were so consumed with problems of gambling. Weisberg demonstrates that it was the metaphorical role of the lottery that enabled classical probabilists to account for the reasonableness of men, all of whom approach an uncertain

---

[1] Lorraine Daston, *Classical Probability in the Enlightenment*, (Princeton University Press, 1988). Like most other advancements in the Scientific Revolutions of the 16th and 17th centuries in the mixed mathematical disciplines of mechanics, astronomy and optics, probability was not essentially a discovery based on new evidence, but a new account of existing evidence, in this case of chance events, that no longer appealed to an overarching natural order. As such, it was the ongoing hold on the early modern imagination of Fortuna as part of the natural order, even as late as Pascal in the 17th century, that held back the discovery of classical probability.

[2] The allure of method that captured Descartes and Bacon inspired Leibniz that controversy could be resolved through reasonable calculus "Let us calculate, Sir; and thus by taking to pen and ink, we should settle the question".

[3] Herbert I. Weisberg, *Willful Ignorance*, (John Wiley, 2014)

situation *as if* it were a lottery, in terms other than the role of fortune in the natural order. Specifically, according to classical probabilists beginning with Jacob Bernoulli, reasonableness reckons with chance by abstracting from our ambiguous intuitions a set of specific causes, which are assumed to function *as if* they were a lottery yielding odds for and against an outcome, odds which then dialectically turns back on and clarifies our intuition and judgment.

In Bernoulli's classic *Ars Conjectandi*, he seeks to understand "the dexterity of the physician" and "the prudence of the statesmen" by suggesting that "the work of all of these kinds of individuals depends upon *conjectures*, and every conjecture involves weighing complexions or combinations of causes". After suggesting that perhaps combinatorics can remedy "the defect in our minds" that results in "insufficient enumeration of cases", he then raises the following objection.

> *But here we come to a halt, for this can hardly ever be done. Indeed, it can hardly be done anywhere except in games of chance. The originators of these games took pains to make them equitable by arranging that the numbers of cases resulting in profit or loss be definite and known and that all the cases happen equally easily. But this by no means takes place with most other effects that depend on the operation of nature or on human will.*

The ambiguity around the operations of nature and the human will prevent any simple combinatorial calculation of possible outcomes of a situation. The solution to this dilemma, for Bernoulli, is to act *as if* the outcomes of economic and civic matters were the results of such games of chance.

> *Nevertheless, another way is open to us by which we may obtain what is sought. What cannot be ascertained a priori, may at least be found out a posteriori from the results many times observed in similar situations, since it should be presumed that something can happen or not happen in the future in as many cases as it was observed to happen or not to happen in similar circumstances in the past. If, for example, there once existed 300 people of the same age and body type as Titus now has, and you observed that two hundred of them died before the end of a decade, while the rest live longer, you could safely enough conclude that there are twice as many cases in which Titus also may die within a decade as there are cases in which he may live beyond a decade.*

So, while human affairs do not exhibit the equitable random variation that only exists in artificial games of chance, we can make assumptions about (1) the relative stability of the relevant operations of reality and (2) the lottery-like variation of specific features of reality in conformance with past observation, to disclose and make broadly available conjectural reasoning.

The Emergence of the Ideal of Objective Validity and Freedom from Bias

A tension within this classical account of probability has animated debates about the proper role of probability from its beginnings to the present-day. Does probability describe how people reason in the face of uncertainty, or prescribe how they should reason? When descriptive accounts of decision making clashed with the actual decisions made by most people, probabilists cast uncertainty no longer as sensible reasoning that is represented and refined by probabilistic abstraction, as if outcomes could be conceived as the result of a metaphorical lottery, but instead reduced uncertainty to mathematical probability. With the death of the metaphor of the lottery, random variation was reified. Uncertainty no longer encompassed ambiguity into the causal setup of a situation but was instead simply identified with random variation – a random lottery - as measured by mathematical probability.[4]

Thus began the emergence of the moral ideal animating statistics that today is paramount, the search for a universally valid model that is free of human bias and prejudice.[5] Whereas the first era of classical probability sought to inform judgments with models of formalized good sense, what is definitive of the 2nd era of probability, extending from roughly 1840 to the present-day, is the rejection of judgment and subjectivity itself.[6] The moral vision of enlightened judgments by the many was replaced by a moral vision of universally valid knowledge that eliminated subjective prejudice through mechanistic objectivity. As explained below, this formalist probability would lead to multiple ethical crises, from eugenics to the replicability crisis, and would inevitably cause more harm when applied to whole new practices in AI.

The mean became the chief object of statistical analysis. Rather than the rationality of the few as the model, statisticians turned to the irrationality of the many and made the average person the objective basis for statistics. The term 'statistics' itself arose in the late 1700s as the social mathematics of states, as the focus of analysis shifted from understanding individual decisions to understanding society.[7]

This shift was reflected in Poisson's revision of Bernoulli's law of large numbers to account not for a single underlying causal probability, but for numerous fluctuating underlying probabilities, whose effects were found to converge to a mean more quickly than those of a single probability. Unlike classical probability which assumed a mechanistic causal order that was probabilistically understood by people of good sense, 19th century probability saw lawlike

---

[4] From the beginning of classical accounts of probabilistic reasoning, there was a tension between descriptive accounts of reasonableness, on the one hand, and resistance to such accounts by social actors, on the other. Resistance was encountered from both those whose actions defied the accounts of the probabilists, and those who objected to the probabilist accounts of their good judgment. This was true of sellers of insurance and annuities as of gamblers, both of whom rejected models custom-made based on observation of their practices. Practitioners in areas involving risk and judgment viewed their expertise in terms of seasoned judgment of the individual case – not in terms of rules that applied en masse. This tension threatened the moral assumption of reasonableness as the basis of social order, a threat that led to a more concretized understanding of thinking-as-a-lottery that took on a more prescriptive role. This reification of the metaphorical lottery is the basis of the ongoing debate as to its location – whether probability exists in objective reality or subjectively in our minds.

[5] In the middle decades of the 18th century, a hope for the scourge of smallpox that killed 1/7th of Europe came in the form of an inoculation practiced in Turkey that appeared effective, but in the short-term was found to kill about 1 in 200 of its recipients. When Daniel Bernoulli submitted a paper in 1760 to the Paris Academy of Sciences using probability to combine the short-term and long-term risks of inoculation and of no inoculation, to his disappointment many people continued to choose the long-term risk.

[6] For the history of modern statistics from the 19th century see *The Taming of Chance*, Ian Hacking (Cambridge University Press, 2008) and Chapters 2-3 of *The Taming of Chance.*

[7] *Seeing Like a State*, James C. Scott (Yale U Press, 1999)

regularities as arising from disorder and a myriad of various causes. The regularities in births, marriages, crimes and suicides on display in the "avalanche of numbers" produced by new statistical offices across 19th century Europe pointed to a deep social order controlling what had been assumed to be random or anti-social activities. For advocates of statistics such as Quetelet, this deeper social reality, attested by the stability of mean values, is more real than the individuals counted, and he called for responsible government intrusion to cure anti-social maladies like crime and suicide now known to be under social control.

Quetelet was an astronomer and saw himself as bringing the mathematical study of physics to "social physics" which had heretofore been the backwater of a subjective calculus of reasonableness. He had learned the formula of Gauss governing the distribution of errors from true values in astronomical observations and applied it to social statistics. Quetelet and Francis Galton were deeply influenced by the application of the Gaussian distribution, known later as the normal distribution or bell curve, to society, and became committed to eugenics. Whereas Quetelet saw the average man as the type of a nation in comparison to which individuals were flawed, Galton saw the upper ends of the curve as the variation that was the hereditary source of genius. Challenged to explain why the offspring of exceptional people would ultimately revert back to the mean of their ancestors, Galton developed the methods of correlation and regression to attribute the variation of offspring in part to one's parents and in part to variation in the offspring.

It's critical to clarify the specific relationship between eugenics and statistics. At one extreme, one could argue that eugenics was a personal failing of great statisticians, but the personal failure doesn't implicate statistics. At the other extreme, one could argue that statistics is fatally flawed, as it was designed for eugenic aims and is only useful for similar projects of discrimination and control. The argument here takes a third route – statistics is beset by an internal tension, and it is the dominance of the formalist approach to probability, grounded on a reified notion of lottery-like random variation, that accounts for eugenics, the crisis of replicability in the sciences and the harms of AI.

Specifically, in the case of eugenics, by making random variation a part of objective reality, and no longer a convenient metaphor to describe and validate what others took to be the causal setup of their domains, statisticians could dispense with domain "experts" now dismissed as limited by subjective biases and reduce science to a selective abstraction of correlates with mean effects. The distribution of errors around such mean effects were no longer the result of ambiguity, which we would need domain experts to resolve, but random variation.

This reductive formalization of error allows the statistician to play a two-fold role: while claiming to be free from bias, they are now the ones selecting correlates whose main effects are privileged in scientific analysis, analysis that is shielded from criticism by claims to mechanistic objectivity. The two-fold role is what allows the biases of statisticians to influence scientific analysis, while also preventing them from seeing such biases under the guise of mathematical objectivity.

From a broader understanding of probability as a descriptive, dialectical process, there is no reason why main effects should be primary at all in the causal setup of a domain. This is known as the main effects fallacy, in which we first determine main effects, and then add on interaction effects of those main factors.[8] However, in social domains with complex interlocking causes we would expect interaction effects to be primary.

The seemingly objective reality of main effects with random errors convinced Galton that his mathematics of correlation and regression could explain a range of phenomena beyond biological inheritance. He called for an independent, objective mathematical statistics that could be applied to any field. Karl Pearson and R.A. Fisher, also concerned with eugenics, shared Galton's vision and together developed in the early decades of the 20th century much of the mathematical statistics that we know today. The mathematization of statistics over the first half of the 20th century, which dominates present-day statistics, thus emerged directly from the deeper reality of main effects and random error that eugenicists saw as underlying phenomenal appearances.

The main effects fallacy, premised on error being due to mathematically measurable random variation and not ambiguity, made statistics the vehicle for advancing eugenics. Pearson created the Chi Square significance test to measure the distance between two distributions, such as that between different racial groups, as conclusive evidence that the chosen racial correlates are the causal setup that explains variation in intelligence.[9]

The result was a field with an identity crisis. Statistics was a part of mathematics that was prior to, and thus superior to, any particular application, while at the same time its *raison d'etre* was its value to society and to decision making, and any such application of statistical theory required multiple judgments informed by domain expertise. When such judgments are not transparent, they are the means through which bias – whether that of the eugenicist, the scientist publishing non-replicable research, or the AI worker inflicting harm on marginalized communities – is concealed behind a façade of mechanical objectivity.

This identity crisis is on full display in the statistics of hypothesis testing. Significance tests such as p-values, f-values and R^2 values today use arbitrary thresholds of statistical significance to infer the presence of an effect in any object of study or to select from competing statistical models of an effect and have been applied across dozens of fields from agriculture and psychology to medicine and industrial acceptance sampling. This application of statistics has been criticized as having impeded as much as it advanced technical progress, by moving substantive domain expertise "out-of-the-loop" of work in favor of mechanical tests of statistical significance.[10]

[8] Weisberg, pp. 316-317

[9] Aubrey Clayton, "How Eugenics Shaped Statistics"

[10] "If you yourself deal in medicine or psychiatry or experimental psychology, …we would recommend that you focus on clinical significance. If you deal in complete life forms, environmental or ecological significance. If you deal in autopsies or crime or drugs, forensic or psychopharmacological significance. And so forth…An arbitrary and Fisherian notion of "statistical" significance should never occupy the center of scientific judgment." *The Cult of Statistical Significance: How the Standard Error Costs us Jobs, Justice and Lives*, by Deirdre McCloskey and Stephen Ziliak (Univ of Michigan, 2008), p. 20

This positioning of contemporary statistics, as both objectively free of judgment and socially beneficial when applied, even though applied statistics requires judgment, while untenable has become dominant. Nonetheless the broader understanding of probability as a dialectical process for resolving ambiguity, in turn refining the judgment of domain experts, has continued to animate a strand of statistical work, from Shewhart and Deming to Tukey.

This tension between two approaches to probability animates the field of AI and its responses to ethical concerns. The demand for fairness, in particular, is the scene of this tension. The response of formalist AI to fairness demands is to characterize them as independent of the sole concern of AI for accuracy, and thus as requiring a trade-off – between accuracy and fairness. When AI is viewed more broadly as a sociotechnical process, rather than as an algorithm, then demands for fairness and explainability are no longer framed as separate from the concern for accuracy; these demands actually point to the essential role of social context in creating accurate algorithms.

Fairness research is thus happening within the *algorithmic frame* of formalist AI. Within this frame, the relevant features are the output, the training data, and the relationship between them. Success is defined narrowly in terms of accuracy of the output in relation to the training data, and generalizability to new data from the same distribution. This artificial context comes with no concepts for expressing ethical notions such as fairness, and so fairness can only be expressed as a regulatory constraint on such an activity or an algorithmic constraint.

When faced with the question of regulatory constraints, AI researchers frequently appeal to the ethical vision of formalist AI as overcoming human limitations, an appeal that always begins with reducing human reasoning to algorithmic induction.

> *First, all decision-making – including that carried out by human beings – is ultimately algorithmic. The difference is that human decision-making is based on logic or behaviors that we struggle to precisely enunciate. If we humans had the ability to describe our own decision-making processes precisely enough, then we could in fact represent them as computer algorithms. So the choice is not whether to avoid using algorithms or not, but whether or not we should use precisely specified algorithms. All things being equal, we should prefer being precise about what we are doing.*[11]

However the premise that all decision-making, human or machine, is algorithmic, is precisely what is being challenged by demands for fairness and explainability, and in fact has been challenged through the history of probability and statistics.

Another reason AI researchers oppose regulatory constraints is because, having reduced human reasoning to algorithmic induction, it makes sense to believe we can similarly reduce notions of fairness, by mathematically formalizing fairness as a constraint on the algorithmic relationship between training data and the output.

Fairness, however, like many properties of real-world domains, is inherently ambiguous. The role of AI in judicial sentencing is thus not just to correct for human limitations of bias and

---

[11] *The Ethical Algorithm*, p. 191

limited memory, but also to further clarify what fairness means, a process that likely never will reach closure.

As the ethical concerns around AI have mounted, not all AI researchers within AI ethics have responded with mathematical formalizations of ethical concepts. An emerging critique from within AI diagnoses the underlying ethical problem of AI as algorithmic formalism[12] and the formalism trap[13].

3. Questions for Ricoeur

This history of metaphor within probability and statistics finds much in common with Ricoeur's theory of metaphor. The role of metaphor is not found in the mere substitution of words – in fact, lottery was only present as a background concept. Like Max Black's models, which Ricoeur describes as belonging "not to the logic of justification or proof, but to the logic of discovery", the metaphorical lottery of classical probability is productive of "scientific imagination".

Imagination is central to the birth of a metaphor for Ricoeur, a birth which reflects the stereoscopic vision of its author who sees a depth of meaning in a metaphorical term that exceeds that of its literal and new contexts. In fact, Ricoeur's theory of metaphor gives an account of the generation of metaphor that one does not find in semiotic theories of metaphor.

Two questions, however, are raised for Ricoeur from this history.

First, why does Ricoeur's work on metaphor, which stressed the generation, life and death of metaphor as a linear trajectory, not also account for a concurrent tension between living and dead metaphors?

Ricoeur's account of metaphor doesn't appear to rule out a concurrent tension between living and dead metaphors within some community of practice. Perhaps a reason for the absence of any such account lies in his greater concern to respond to the poststructuralism of Derrida, which he did with an emphatically linear account of the life and death of metaphor.

The unearthing of the role of metaphor 350 years ago in the birth of our notion of probability resembles the reanimation of dead metaphor by Derrida (*logos*) or by Heidegger (*alethia*). Ricoeur is concerned to oppose such reanimation as premised on words having a proper meaning rooted in their visible sense that contrasts with their metaphorical meaning operating as invisible metaphysical concepts, concepts that silently structure our worlds once their metaphorical origins fade away and die. For Ricoeur, "literal does not mean proper in the sense of originary, but simply current, 'usual'". "To revive dead metaphor is in no way to revive concepts", according to Ricoeur, because a revivified metaphor is really just a new metaphor.

This emphatically linear account the birth, life and death of metaphors is important for Ricoeur to avoid the denominational focus of metaphorical theories that see the potential to

---

[12] Green 2020
[13] Selbst 2019

reanimate dead metaphors. Any rebirth is really the employment of a term in a new context, a new narrative, not the unearthing of a hidden meaning within the word itself.

However, does this mean that a live and dead metaphor cannot be in tension at the same time? A living metaphor is created through a semantic impertinence between a word and its new narrative context, while its death occurs when the impertinence is gone and its new meaning is now simply its literal, non-metaphorical meaning. But presumably within a linguistic community this impertinence can be operative in some discourses and absent in others.

It seems to me that this fits well within Ricoeur's 'Little Ethics', as a dimension of the injunction to return to narrative unity when actualization of the good life breaks down. Ricoeur gives a hierarchical account of ethical reasoning that is anchored in a notion of the good life that constitutes the narrative unity of one's life plan, and that realizes this notion through moral norms. When such norms encounter conflict and contradiction with competing norms, the *phronimos* returns from the moral norm to the ethical aim so as to deliberatively reinterpret the singular situation. This openness of the *phronimos* contrasts with the one lacking in *phronesis* who does not follow the return path to one's notion of living well, but instead identifies with the moral norm itself.

Is metaphor not birthed by such tensions within a narrative? Aren't the hermeneutical circles within a narrative horizon the way that metaphors are generated?

This gives rise to the second question. Why is metaphor not addressed in Ricoeur's ethics? I confess to lacking a good answer to this question. Perhaps Ricoeur was taking for granted in *Oneself as Another* the reader's familiarity with the structure and dynamics of narrative, including the function of metaphor, as this had been the focus of decades of his writings at that point. What such a solution can't account for, however, is the curious case of Ricoeur's handling of the social contract in *Oneself as Another*. Ricoeur criticizes social contract theory, in particular that of Rawls, as anchored not in the narrative unity of living well, but in "a founding fiction".

The notion of a social contract, of course, is a fairly obvious case of metaphor. We don't actually enter into a contract with other members of society, but we do act *as if* we contracted with them. It was precisely because of the power of this metaphor that early moderns were convinced they could ground society on norms based on mutual consent and universal reasonableness, and not on a transcendent order reflected at different levels of being and enforced by the Church, a conviction that in turn, as we saw above, provoked the birth of probability with the metaphorical use of the lottery as making such reasonableness available broadly. In both cases, of probability and social contract theory, the literal concern was a topic of active thought and development – as is seen in the advances in combinatorial mathematics of games and in contract law in the 17th century – as was the metaphorical use in social contract theory and probability.

The conclusion I would suggest is that Ricoeur, like others seeking a hermeneutic revival of *phronesis* from Heidegger to Arendt and Gadamer, has a low regard for early modern thought, and sees his work as recovering classical insights around *phronesis* and *praxis* from the reductionism of the early moderns. If so, my view is that this bias against early modern thought fails to account for the appeal of early modern thought, an appeal which is evidenced by the successful use of metaphor to spread its more profound ideas.