Probability as Instrumentalist Construction of Determinism

Florin Cojocariu

2025-06-17

A Response to Hume's Problem of Causation

Abstract

This essay argues that probability theory bridges pre-verbal pattern recognition (non-deterministic) and rational inference (deterministic), responding to Hume's challenge about causation. The central insight is that rationality requires deterministic structures to function, leading us to construct determinism even where none exists. Hume demonstrated that causality exemplifies this problem. Probability represents a sophisticated form of this construction—a mathematical reification of uncertainty that preserves the logical apparatus of rational discourse while acknowledging epistemic limitations. Moreover, the LLMs pattern recognition mechanism seems to be a new method to construct determinism, one that transcend probability.

Introduction

Hume's analysis in the *Inquiry Concerning Human Understanding* demonstrated that causal reasoning lacks rational foundation, yet we continue to reason causally and make successful predictions. This essay proposes that probability theory serves a crucial philosophical function: it satisfies rationality's structural need for determinism while honestly encoding our uncertainty.

The key insight is that rational discourse—with its logical operations, mathematical calculations, and precise predictions—requires deterministic objects to manipulate. When faced with genuine uncertainty, rationality doesn't abandon its project, but instead creates mathematical objects that can be treated deterministically. Probability assignments like $P(\text{rain}|\text{clouds}) = 0.7$ become facts we can reason about, calculate with, and build theories upon, even though they encode our epistemic limitations rather than metaphysical necessities.

This transformation turns epistemic uncertainty into formal mathematical structures that function as determinate objects. We preserve rational inference while avoiding the metaphysical overreach that Hume exposed.

This analysis reveals a broader trajectory in how rationality constructs formal structures: from naive causation, to sophisticated probability theory, to emerging frameworks for non-temporal pattern recognition exemplified in artificial intelligence systems.

Hume's Challenge and the Structure of Inference

The Problem of Inductive Reasoning

Hume showed that causal reasoning relies on the unjustifiable assumption that future instances will resemble past ones (*Inquiry*, §4.2). What we call causation reduces to observed "constant conjunction" projected onto future cases through psychological habit rather than logical necessity. This creates a fundamental problem: our most basic form of reasoning appears to lack rational foundation.

Levels of Cognitive Processing and Inference

The solution requires distinguishing different levels at which inference operates:

Pre-rational pattern recognition: Animals and humans unconsciously detect regularities in experience. This capacity requires no justification—it simply describes what cognitive systems naturally do when processing sensory input over time. Hume explicitly supports this view in his discussion of animal reasoning: "First, It seems evident, that animals, as well as men, learn many things from experience, and infer, that the same events will always follow from the same causes" (*Inquiry*, §9).

Crucially, Hume argues this learning occurs without rational deliberation: "It is impossible, that this inference of the animal can be founded on any process of argument or reasoning" (*Inquiry*, §9). He extends this point to humans: "Animals, therefore, are not guided in these inferences by reasoning: Neither are children: Neither are the generality of mankind, in their ordinary actions and conclusions" (*Inquiry*, §9). As Evans (2008) shows in dual-process theory, this "System 1" processing operates automatically and pre-consciously, forming the foundation for higher-level reasoning without itself requiring rational warrant.

Causal interpretation: Humans bring detected patterns to conscious reasoning and interpret them as *necessary connections*—"A must produce B." This interpretation imports metaphysical claims that cannot be rationally defended.

Hume identifies this as the core philosophical error: "There are no ideas, which occur in metaphysics, more obscure and uncertain, than those of power, force, energy or necessary connection" (*Inquiry*, §7). The problem is that we never actually observe the supposed causal powers: "In reality, there is no part of

matter, that does ever, by its sensible qualities, discover any power or energy, or give us ground to imagine, that it could produce any thing" (*Inquiry*, §7).

Even in familiar cases, the connection remains mysterious: "We know that, in fact, heat is a constant attendant of flame; but what is the connexion between them, we have no room so much as to conjecture or imagine" (*Inquiry*, §7). The idea of necessary connection arises not from observation but from mental habit: "This connexion, therefore, which we feel in the mind, this customary transition of the imagination from one object to its usual attendant, is the sentiment or impression from which we form the idea of power or necessary connection" (*Inquiry*, §7).

Probabilistic interpretation: There is an alternative (and better) interpretation that represents 'habit' strength mathematically without claiming necessity—"A produces B with probability P."

The key insight is that pattern recognition (or what Hume relegates to the "habit" theory) itself is not the problem. The difficulty arises when we try to translate detected patterns into claims about *logical necessity*. Hume captures this precisely: "All events seem entirely loose and separate. One event follows another, but we never can observe any tie between them. They seem conjoined, but never connected" (*Inquiry*, §7). The transition from observing regular conjunction to claiming necessary connection is the philosophical misstep that generates the Humean problem.

Moving from causal to probabilistic interpretation represents a more sophisticated strategy for preserving rational discourse despite Humean skepticism[1].

Probability as Reification of Indeterminism

The Demand for Determinism

Rational discourse operates through logical relations, mathematical operations, and systematic inferences—all of which require determinate objects. We cannot perform modus ponens on maybes, calculate with perhapses, or build scientific theories on vague hunches. This creates a fundamental tension: experience provides patterns without necessity, but rationality needs necessity (or something structurally equivalent) to function.

The causal interpretation represents rationality's first attempt to resolve this tension—simply declaring that observed patterns reflect necessary connections. Hume devastates this move by showing we never observe the supposed necessity, only regular succession.

---

[1] It can be noted that this is what saved Quantum Mechanics from its apparent contradictions.

The Instrumentalist Nature of Probability

Rather than abandoning rational discourse, we developed a more sophisticated construction. Probability theory transforms our epistemic limitations into mathematical objects with all the properties rationality requires:

- **Deterministic values**: $P(A) = 0.7$ is a precise number we can calculate with
- **Logical relations**: Probability spaces follow Boolean algebra
- **Systematic operations**: Bayes' theorem, marginalization, conditioning
- **Theoretical embedding**: Probabilities become facts in scientific theories

Consider how weather services treat "30% chance of rain" not as an expression of ignorance but as objective information to be broadcast, planned around, and evaluated for accuracy. The probability has become a determinate fact about the world, even though it encodes uncertainty.

The Bridge Function

This reification serves as a bridge by preserving what rationality needs (deterministic structures) while respecting what Hume showed (no observable necessity). We can engage in all the formal operations of rational discourse—calculation, inference, theory-building—without claiming to have detected metaphysical necessities in nature.

Probability transforms the problem: instead of trying to find determinism in the world (the causal strategy), we create deterministic representations of our indeterminate situation (the probabilistic strategy)[2]. This allows rational discourse to proceed without metaphysical overreach. It creates a hybrid form of inference that maintains mathematical precision while remaining grounded in pure, non causal, non-deterministic observation.

De Finetti's instrumentalist framework that [@barlowIntroduction-Finetti19371992] explicates, provides the philosophical foundation for this approach: just as de Finetti argued that probabilities function as degrees of belief rather than objective features of reality—famously declaring that 'probability does not exist' in any objective sense—the present argument extends this insight to show how probability assignments serve as mathematical tools that preserve rational discourse while honestly encoding our epistemic uncertainty about detected patterns.

---

[2]This approach differs fundamentally from probabilistic theories of causation (e.g., Suppes 1970; Kvart 1986, as discussed in [@ben-menahemCausationScience2018, p.7]) which attempt to analyze causal relations in probabilistic terms—arguing that causes raise the probability of their effects. Such theories still make metaphysical claims about causation itself and face their own paradoxes regarding spurious correlations and probability-lowering causes. By contrast, the present approach treats probability not as an analysis of causation but as an alternative mode of interpretation entirely—a formal structure we impose on detected patterns to enable rational discourse without any metaphysical commitments about causal relations in nature. We are not saying "causation is probabilistic" but rather "probability provides a way to mathematically represent patterns without invoking causation at all."

If probability serves as the mathematical reification of temporal uncertainty, this raises the question: what formal structures might serve non-temporal pattern recognition? Recent developments in AI systems suggest this is not merely theoretical speculation.

Probabilities and AI: Are LLMs Stochastic Parrots?

Having shown how probability theory imposes deterministic form on temporal uncertainty, we now consider a contrasting case. Large language models do not accumulate frequencies over time to build probabilistic expectations; instead, they encode an entire prompt all at once into a high-dimensional vector space, extracting patterns through simultaneous attention over every token. This parallel processing is not "temporal probability" in the usual sense—there is no sequence of past events whose frequencies are being tallied—but a non-temporal recognition of geometric configurations. In other words, whereas classical probability theory bridges inductive gaps by translating time-based uncertainty into precise numerical values, LLMs bridge the same gap by converting fuzzy, context-laden inputs into a single deterministic embedding that captures all relational patterns at once.

Current explanations of how LLMs function rely heavily on statistics and probabilities. The "most probable next word" theory is still the most widely adopted by the general public while in the academic papers on LLMs "statistics" and "probabilities" can be encountered at every step. [@benderDangersStochasticParrots2021]'s "stochastic parrot" critique helped to make this probabilistic vision almost the norm, to the point that any LLM answering questions about its own functioning, will present some sort of probabilistic or statistic view. This probabilistic paradigm exemplifies the pattern identified in our analysis of Hume: rationality creating mathematical structures to handle indeterminate phenomena. But what if this framework misses something fundamental about how these systems actually operate?

Human causal and probabilistic reasoning unfolds sequentially over time—we observe patterns, track frequencies, build temporal expectations. Our idea of probability is closely connected to the idea of repeated observation or measurement: something is repeated *in time*. LLMs, through the sequential behavior of their interfaces which print a word after another, created the illusion that they search "the most probable next word", which became the first main paradigm for understanding them. However, this is an illusion.

LLMs process patterns through weighted associations in high-dimensional spaces where relationships exist simultaneously rather than sequentially. As [@vaswaniAttentionAllYou2023] demonstrated in developing transformer architecture, attention mechanisms allow models to access all pattern relationships **at once** rather than processing them temporally.

To see this in concrete terms, imagine feeding an LLM the entire text of Shakespeare's *Hamlet* at once: the model's self-attention mechanism "computes" relationships among every pair of words **in parallel**, instantly "seeing" thematic links (e.g. "to be"　"not to be," "ghost"　"revenge") without traversing the play linearly. These attention-weighted associations are a simultaneous map of semantic and syntactic connections—rather than the time-based frequency counts underlying classical probability. In traditional inference, we build P(event) by tallying occurrences over successive trials; by contrast, the LLM's embedding space encodes the entire text's structure in one go and yields a deterministic representation that the decoder then linearizes. This non-temporal pattern capture has been demonstrated in the original Transformer paper, which shows how self-attention layers compute all pairwise token interactions in a single pass.

Modern LLMs actually operate in two distinct phases. First, upon receiving the entire prompt, self-attention layers compute representations for every token in parallel—analogous to instantly perceiving a cow as a whole image rather than piecemeal parts. In this encoding phase the model "decides" what content and arguments to produce, drawing on its high-dimensional pattern of associations. Only afterward does it enter an autoregressive decoding phase, linearizing that decision into a sequence of output tokens. Thus, the apparent sequentiality of LLM responses reflects the necessary formatting for human readers, whereas the core "inferential" work occurs non-temporally in the parallel encoding stage.

[@ferroneSymbolicDistributedDistributional2020] further show how these models create distributed representations where semantic relationships exist as geometric configurations in vector spaces, fundamentally different from sequential temporal processing. When we repeatedly flip a coin many times to prove that the probability for each side is 0.5, we need to enter each result in some sort of table so that we can compute probabilities after sufficient throws. But a LLM functions more like having the result table at once, with its internal pattern already clear.

This suggests that probability is the temporal special case of patterns—a formalism to describe patterns occurring in time.

Beyond Probabilities and Temporal Reasoning : A Broader Pattern

This progression from causation through probability toward pattern recognition reveals that probability theory may be one instance of a broader class of formal bridges between immediate perception and rational discourse. If probability serves temporal reasoning by mathematically encoding uncertainty over time, perhaps other formal structures could serve different types of pattern recognition. Recent developments in large language models illuminate this possibility by revealing forms of pattern recognition that operate through simultaneous rather than sequential processing.

Consider the difference between building probability estimates through temporal observation (flipping coins over time) versus apprehending patterns immediately (seeing balance in a spatial array of 10,000 simultaneous coin flips displayed as a 100 by 100 matrix, where each cell is either black (a "head" result) or white(a "tail" result). In the temporal case, we require probability theory as mathematical scaffolding to handle uncertain accumulation over time. In the simultaneous case, we might perceive pattern directly without numerical calculation—the overall "noisiness" or "balance" of the visual display (its grayness) embodies the probabilistic information without mathematical abstraction.[3]

This distinction suggests that probability might be one instance of a broader class of formal structures that enable rational discourse about different types of patterns. Just as we created probability to handle temporal uncertainty, i.e. *temporal patterns*, we may need new mathematical frameworks to handle simultaneous, non-temporal, high-dimensional patterns. This remains a promissory note for future investigation, but it suggests the philosophical strategy explored here may have applications beyond traditional causal reasoning.[4]

Some More Objections and Limitations

The Persistence of Inductive Assumptions

Critics might argue that probabilistic reasoning still assumes past frequencies guide future expectations. The claim that $P(B|A) = 0.8$ based on observed frequencies requires believing future instances will resemble past ones—exactly what Hume showed cannot be rationally justified: "If there be any suspicion that the course of nature may change, and that the past may be no rule for the future, all experience becomes useless, and can give rise to no inference or conclusion" (*Inquiry*, §4).

Response: The probabilistic interpretation acknowledges this limitation explicitly. Rather than claiming justified knowledge about future frequencies, it represents the current state of our pattern detection systems. We're not making metaphysical claims about the world, but describing our cognitive situation.

The Normative-Descriptive Gap

Understanding the psychological mechanism of pattern recognition does not address whether we *should* rely on such mechanisms. Hume's challenge operates at the level of rational justification, not psychological description.

---

[3]This insight is actually developed in a new paper, linking into my actual research on pattern recognition; a draft is available on request.

[4]One of the most intriguing ideas following from this approach is that what we call "reasoning" or "inference" is just the special, temporal case of something more general that becomes apparent in the parallel, non-temporal nature of pattern recognition in LLMs. There is some sort of "reasoning", only it is not in its formalized, temporal form, but in something closer to what any living being is practicing daily: pattern recognition.

Response: This objection assumes that rational justification must be foundational rather than pragmatic. The success of probabilistic reasoning in prediction and control may provide sufficient warrant without requiring a priori certainty.

Causation in Special Theory of Relativity

In [@ben-menahemCausationScience2018], the author argues that:

> According to the special theory of relativity (STR), the temporal relations between events are only well-defined in regions of spacetime charted by light signals representing (and limiting) the possibility of causal interaction. When events are separated by space-like distances, there can be no causal interaction between them, and consequently, their temporal order is not invariant, but varies with the coordinate system. Rather than being reducible to spatiotemporal relations, causality now appears to be the basis for the very structure of spacetime. Causal relations are thus at least as fundamental as temporal relations, and arguably (as suggested, for example, in Reichenbach 1956), conceptually prior to temporal relations.

While this seems to bring causation back, I argue the opposite is the case. The relativity of simultaneity in STR further illustrates how causal interpretation depends on theoretical frameworks. Without absolute simultaneity, we cannot even identify which events could potentially be cause and effect without first accepting the theory's constraints on causal possibility.[5] This reinforces that causation functions as a theoretical organizing principle rather than an observable relation. That is to say in modern physics causation is rather constructed[6] than discovered.

Conclusion

Probability theory serves as a crucial philosophical bridge between unconscious pattern recognition and rational inference. This preserves the bridge function identified earlier—enabling rational discourse while acknowledging epistemic limitations.

---

[5]A simple mental experiment can be imagined to show it: a fire cracker explodes in London at midnight and another in New York 1 milliseconds later (suppose we have synchronized clocks). The interval is sufficiently small so light cannot reach from the London event to the New York event before the NY event. Depending on their speed, position and direction, two observers may not agree on the real order of events. In fact causation is only possible for events inside the light-cone (separated in space so that light can travel from one to another in a shorter time), but for an observer who knows nothing about STR this is not apparent. The conclusion here is that causation is a feature of the theory, not of reality and ,yes, certain theories like STR are explicetly causal.

[6][@ben-menahemCausationScience2018] introduces the notion of causal constraints, one that better integrates into the modern physics theories and better avoids this 'construction' argument.

This reveals something deep about the nature of rationality itself: it requires deterministic structures to function, and when reality doesn't provide them, rationality constructs them (hence all what we call "instrumentalism"). The movement from causal to probabilistic interpretation represents an evolution in this constructive project—from naive metaphysical claims about necessity to sophisticated mathematical encodings of uncertainty.

This philosophical strategy finds remarkable parallel in the development of quantum mechanics. Classical physics assumed ontological determinism—complete predictability given perfect knowledge of initial conditions. Quantum mechanics forced physicists to confront fundamental uncertainty, yet rather than abandoning systematic reasoning, they developed mathematical frameworks to work with irreducible uncertainty itself. The wave function represents precisely this transformation: taking epistemic limitations about measurement and encoding them in mathematical structures that become the foundation for physical predictions.

Rather than abandoning systematic reasoning, physicists developed mathematical frameworks that treat uncertainty itself as the fundamental reality. The wave function performs precisely the transformation identified here: it takes our epistemic limitations about measurement and encodes them in mathematical structures that become the foundation for physical predictions. Probability amplitudes become determinate mathematical objects we can calculate with, even though they represent our fundamental inability to simultaneously know position and momentum.

Bibliography