There are No Universal Rules for Induction

John D. Norton[1]
Department of History and Philosophy of Science
Center for Philosophy of Science
University of Pittsburgh
http://www.pitt.edu/~jdnorton

In a material theory of induction, inductive inferences are warranted by facts that prevail locally. This approach, it is urged, is preferable to formal theories of induction in which the good inductive inferences are delineated as those conforming to universal schemas. An inductive inference problem concerning indeterministic, non-probabilistic systems in physics is posed and it is argued that Bayesians cannot responsibly analyze it, thereby demonstrating that the probability calculus is not the universal logic of induction.

1. The Material Theory of Induction

1.1 What It Is

In papers elsewhere (Norton, 2003, 2005) I have argued for a material theory of inductive inference. It urges that there are no universal rules of inductive inference; particular inductive inferences are not shown to be licit by demonstrating that they conform to universally applicable schemas, as is the case with deductive inference. Instead, individual inductive inferences are warranted by facts. Since these facts hold contingently, they warrant inferences only in limited domains. While some domains may be large and governed by their own well-defined logic, no single logic is warranted in all domains. The slogan is "all induction is local."

The principal burden of any logic is to separate the good from the bad inferences. Deductive logic provides the model of how a formal theory does this. Consider the deductive argument:

1. Either quantum particles in a singlet state are non-local or superluminal transmissions are possible.
2. Superluminal transmissions are not possible.
3. Therefore quantum particles in a singlet state are non-local.

Everyone sees immediately that this is a deductively valid argument. It is an instance of the universally applicable schema of disjunctive syllogism

1. A or B
2. Not-B
3. Therefore A

We see this without grasping the content of the propositions. One may have no understanding of a singlet state or non-locality. Nonetheless one can be assured of the validity of the argument.

Standard theories of inductive inference have long aspired to replicate this success. Maddeningly, these standard theories fall just a little short; and sometimes worse. The principal difficulty is that it is impossible to judge whether an inductive inference is good without knowing what the propositions say. Here are two formally identical inductive inferences:

1. This sample of bismuth melts at 271ºC.
2. Therefore all do.

1. The first day of the new millennium was 8ºC at noon in Pittsburgh.
2. Therefore all first days of new millennia in Pittsburgh will be so.

Whether the inferences are good depends on what the propositions say—their matter or material. The first is a good inference because of the fact that elements like Bismuth are generally, but not

assuredly, uniform in their properties. The second is a poor inference, since there is no corresponding fact of the uniformity of the weather. That is all we need to know to determine that the first inference is warranted and the second is not.

The natural response is that these examples are just too simple. State all the background facts properly and then the universal inductive logic that really governs the examples will come into view. This response expresses an optimistic hope that is not borne out by further efforts. When we do try to expose this universal background logic, things become worse. We may imagine that the world is governed by a principle of uniformity that holds for some aspects of the world but not others. But we will find no non-circular characterization of which are the favored aspects. We were better off with the simple idea that the first inference is warranted by a fact about elements and the second not warranted because the facts of weather are inhospitable.

1.2 A Psychical Example

The material theory urges that facts warrant inductions. Thus we should expect that two people who differ sufficiently in their assessment of the prevailing facts will also differ significantly in their judgments of the inductive import of the same evidence. This effect is illustrated in a discussion by Bertrand Russell of experimental evidence for personal survival of physical death. Russell (1957, 51-52) writes

> Psychical research professes to have actual scientific evidence of survival and
> undoubtedly its procedure is, in principle, scientifically correct. Evidence of this
> sort might be so overwhelming that no one with a scientific temper could reject it.
> The weight to be attached to the evidence, however, must depend upon the
> antecedent probability of the hypothesis of survival. There are always different
> ways of accounting for any set of phenomena, and of these we should prefer the one
> which is antecedently least improbable. Those who already think it likely that we
> survive death will be ready to view this theory as the best explanation of psychical
> phenomena. Those who, on other grounds, regard this theory as implausible will see
> for other explanations.

Some experimental report purports to favor survival. Those who believe that survival is possible will infer that the experiment does inductively support a survivalist outcome. Imagine that a psychic purports to learn information from a departed relative that only that relative could know. The believer will infer from that datum to the survival of the relative. Skeptics will infer to some form of trickery, perhaps "cold reading" or even collusion.

The facts we believe control the conclusion to which we infer inductively. The material theory asserts that this is the full analysis. Might a fuller analysis find that these inferences derive from application of some universal inductive schema? Russell's text mentions two terms, "explanation" and "probability." They immediately suggest the two presently, most favored universal accounts of inductive inference. Let me explain why each does no better than the material theory.

According to "inference to the best explanation" we should infer inductively to the best explanation of the evidence. The believer infers to the relative's survival as the best explanation of the origin of information conveyed by the psychic.

This is, in my view, a very imperfect explication of the inference. If we focus on the particular inference and all the pertinent facts, its compelling strength is quite evident. If we rise to the abstract description of it as an inference to the best explanation, its strength is obscured. We have lost essential detail. Our abstract description is merely a gloss that fails to capture what makes the specific inference successful at the strength it has. Moreover, inferring to the best explanation fails to provide a precise, universal inductive schema because our accounts of explanation still lack the precision expected in a general theory of inference. Contrast this with the case of disjunctive syllogism above. This effect is exactly reversed. In that example, identifying the schema of disjunctive syllogism in amongst the complications of the quantum theory was all that was needed to establish the deductive force of the inference. The quantum mechanical content can be ignored.

When Russell urges that "we should prefer the [hypothesis] which is antecedently least improbable," Bayesians will call up a familiar result in Bayesian epistemology. When two hypotheses entail the same evidence, the ratio of posteriors equals the ratio of priors. So we should prefer the hypothesis with the greater prior probability.

This is, in my view, an imperfect explication of the inferences. The notion of "probable" and "improbable" is metaphorical, as suggested by Russell's equating of "improbable" with

4

"implausible." One does not need the mathematics of probability theory to judge that we should prefer the antecedently most probable/plausible hypothesis if the two are equally adequate to the evidence.

If the term probability is used in its precise mathematical sense, then what precisely is the outcome space envisaged? And precisely which are the numerical values of the probabilities? Bayesians will here resist specifying the outcome space precisely and assigning definite numbers to the probabilities. They realize that any particular set assigned would be arbitrary. Instead they retreat to the assurance that whatever numbers they assign will lead to pretty much the same outcome. I read that reluctance as conceding the spurious precision in the assigning of definite numbers and that the inductive logic prevailing is actually a weaker logic that permits rankings only. I have suggested elsewhere (2007) that Bayesians would do well to find it.

Bayesians may also see in Russell's remarks a suggestion that the best explanation is just the antecedently most probable, so that the hypothesis that explains better is just the one with a greater prior. That proposal fails because explanation and explanatory strength is a relation between the hypothesis and the thing explained, whereas a prior probability is assigned to an hypothesis only. It is easy to find hypotheses that are very good explanations of some things and explanatory failures when it comes to others. Einstein's celebrated light quantum hypothesis of 1905 affords a Nobel Prize winning explanation of the photoelectric effect. Yet that same hypotheses is unable to explain the familiar interference effects of 19th century optics. A single prior probability cannot simultaneous reward and punish the hypothesis.

2. A Challenge to Bayesians

Here I present a challenge to anyone who thinks that the probability calculus is adequate as the universal logic of induction. It is a problem in inductive inference that I do not believe Bayesians can solve responsibly.

2.1 Indeterministic Systems in Physics

Indeterministic systems are those for which a full specification of their present state fails to fix their future states. In quantum theory a specification of the present state of the system generally only fixes the probabilities of different futures. In a more extreme form of indeterminism, the full specification of the present leaves the future undetermined and—the key

fact of importance here—our physical theories provide *no physical chances* for the different futures. They tell us only which futures are possible. Some recent examples arise in the supertask literature. See, for example, Alper at al. (2000); Norton, (1999).

An examples in Newtonian physics is "the dome." (See Norton, 2003a, §3; 2008a.) A point mass can slide frictionlessly over a dome with circular symmetry in a vertical gravitational field. Initially, the mass is motionless at the apex. See Figure 1. If the shape of the surface is chosen appropriately, Newton's equations admit many solutions. The mass may remain at rest indefinitely at the apex; or it may remain at rest for some arbitrary time T and then spontaneously accelerate in any radial direction. The spontaneous motion does not arise from some very slight perturbation, a miniscule wobble, say, that shakes the mass free at the moment of spontaneous excitation, time t=T. Nothing changes in the state of the dome. Newton's equations of motion just admit multiple solutions, one in which the mass remains at rest at times t>T and one in which it moves for t>T.
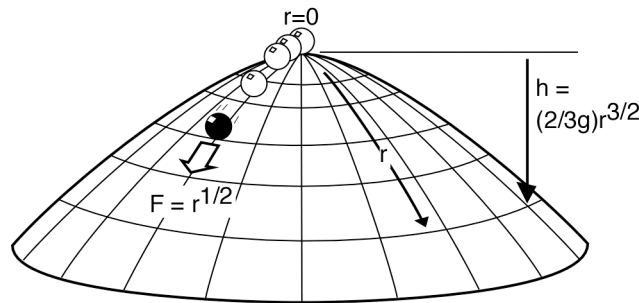

Figure 1. The dome: an indeterministic system[2]

A second example illustrates another sort of spontaneous excitation. An infinite sequence of identical masses and springs are laid out as mass-spring-mass-spring-… as shown in Figure 2, where the springs are governed by Hooke's law.

Figure 2. Masses and Springs

---

[2] Figure from Norton (2003a, §3).

If the masses are initially at rest and located so that the springs are unextended, then one possible future for the system is that it remains indefinitely quiescent. There is a second possibility. The system can become spontaneously excited at some time T>0. The first mass in the chain is accelerated after T=0 because of a faster acceleration of the second mass; and the second mass is accelerated after T=0 because of a faster acceleration of the third mass; and so on indefinitely. The spontaneous excitation depends essentially on the infinity of the masses; for, were there only n masses, there would be no additional (n+1)th mass to excite the nth mass and initiate the process. The computation is given in Norton (1999, 1269-71). The full specification of the physical system merely admits this spontaneous excitation as a possibility, but supplies *no probabilities* for the various possible futures.

2.2 The Inductive Inference Problem

We are presented with an indeterministic system such as the dome or the masses and springs in a quiescent state at time t=0. We know the full governing physics, so we know that spontaneous excitation is possible, and we know that no other perturbing influences will come into play. How much inductive support does this background information E accord to the hypothesis $H(T_1,T_2)$ that the system will spontaneously excite at a time t in the interval $T_1 \le t < T_2$?

2.3 A Model for the Solution

Consider the standard way of handling inductive inference problems concerning stochastic systems where we do have probabilities—physical chances—for the possible futures. We conform our degrees of belief to those physical chances; this informal idea is given more precise expression in Lewis' (1980) "principal principle." The law of radioactive decay asserts that a radioactive atom has a physical chance P(t) of decaying in time t where

$$P(t) = 1-\exp(-t/\tau). \tag{1}$$

The time constant τ is related to the element's half life by $t_{1/2} = \tau \ln 2$. So P(t) should also be our degree of belief that the atom decays within time t.

This procedure seems so straightforward as to require little justification. Once we accept that the quantum theory provides the full physical description of the system, it would be perverse

to distribute beliefs in a way that does not respect these physical chances and insist, say, that the evidence accords near certainty to the hypothesis that atom will decay in its half life. For to do that is to pretend to know more about the system than the physics tells us; it is to offer a corrective to the physics.

This idea that the physical facts prevailing should determine how we should infer inductively is the basic idea of the material theory of induction. Write

$$[H(0, T)|E]$$

as the degree of support provided by the evidence E that the atom is undecayed at time t=0 to the hypothesis H(T) that the atom decays over the following time $0 \le t < T$. Since the process is fully governed by the law (1), the material theory of induction enjoins us to set the degree of support equal to physical probability of decay

$$[H(0, T)|E] = P(T) = (1\text{-}\exp(\text{-}T/\tau)) \tag{2}$$

2.4 How NOT to Infer Inductively about Indeterministic Systems

Now consider the inductive inference problem posed in Section 2.2. We have

E: At t=0, the system is quiescent.

$H(T_1,T_2)$: The system excites at time t in the interval $[T_1,T_2)$, that is, in $T_1 \le t <$ T_2?

What is the degree of inductive support $[H(T_1,T_2)|E]$ accorded by the evidence E to the hypothesis $H(T_1,T_2)$?

Many cannot resist treating this problem just like radioactive decay. They propose we use the same rule

$$[H(0,T)|E] = P(T) = (1\text{-}\exp(\text{-}T/\tau))$$

The motivation is that the law of radioactive decay has an important property. It is the *unique* decay law that has "no memory." If the atom has not decayed after 1 time unit, or 5 time units, or 100 time units, then the probability of decay in the next unit of time is still the same. Speaking

metaphorically, the atom does not remember how long it has survived without decay, when it decides whether to decay in each new unit of time.[3]

The dome and masses and springs also have this "no memory" property. Whether the spontaneous excitation happens at some moment is independent of how long the system has been quiescent. So, if any probabilistic law is applicable to these systems, it is this one. However we cannot set our degrees of support [H|E] equal to probabilities governed by the same formula (2) of the law of radioactive decay. Any instance of this law has a time constant τ. That time constant exercises a powerful influence on the chances of the spontaneous excitation. Figure 3 displays graphs of P(T) for values of $\tau=0.1$, $\tau=1$ and $\tau=10$:
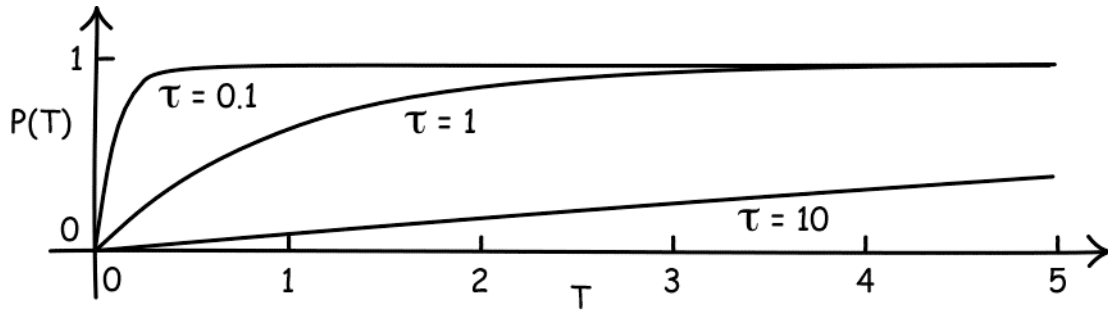

Figure 3. Decay with different time constants

A very small time constant makes the excitation very probable, virtually immediately; a very large time constant delays the excitation very probably, for a long time.[4] Nothing in the physics of the dome or masses and springs fixes a time constant or any time scale for the spontaneous excitation. The physics is completely silent on when the motion may happen. It just says "it's possible."

---

[3] To see this, consider Q(t) = 1–P(t) = exp(–t/τ), the probability of *no* decay in an initial time t. The probability of no decay in a subsequent time u given no prior decay is just

$$Q(t+u)/Q(t) = \exp(-(t+u)/\tau)/ \exp(-t/\tau) = \exp(-u/\tau) = Q(u),$$

which is just the probability of no decay in an initial time u.

[4] Setting $\tau=\infty$ gives us the case of P(T)=0 for all T, so that the mass never moves. That is we assign unit probability to the outcome of no motion ever.

So if we are to use the probabilistic formula, we must add a time scale. That is, we must pretend to know more than the full physical specification of the problem allows. Speaking metaphorically, Nature, in the guise of Newtonian physics, is unable to assign a time scale to the decay. If we assign one, we must know more than Nature. Our original goal was merely to reason inductively about a system. Yet we have ended up as physicists, proposing new physical properties that the system--by construction--does not have.[5]

There is a loophole. Statisticians sometimes use improper probability distributions—that is, ones that do not normalize to unity. An improper density with the "no memory" property is the uniform density $p(T) = dP(T)/dT$, shown in Figure 4. It assigns that the same small probability ε to each unit time interval.
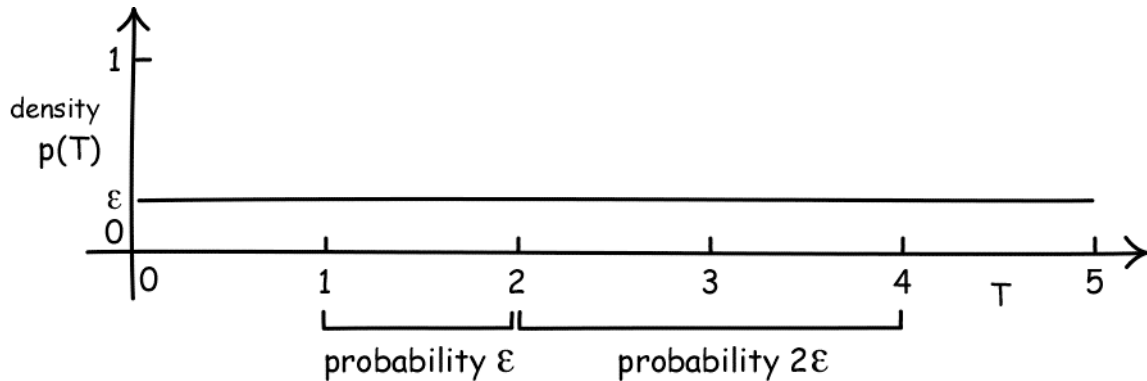

Figure 4. An Improper Probability Distribution

It is "improper" since the probability assigned to all the unit time intervals taken together does not sum to unity, as the probability calculus demands, but it is infinite.

Tempting as this improper distribution may be, it suffers the same problem as the proper distribution. It still adds physical properties. It entails that spontaneous motion in time t=1 to t=2

---

[5] Might a probabilistic analysis be possible if only we were given a little more data, such as the results of observation of several domes over some time? This strategy only makes sense if one does not accept the initial supposition that the Newtonian analysis gives the full physics of the dome. If one accepts that it does, then no catalog of outcomes will give any new, useful information for the inference problem. The situation is analogous to someone keeping detailed records of the outcomes of a roulette wheel's spins in the hope that some dependence between successive spins will be manifested. That strategy in a casino is futile if the wheel has been properly constructed so that there is no dependence.

has probability ε and spontaneous motion in the time interval t=2 to t=4 is 2ε. Motion in the one interval is twice as probable--no more no less--than motion in the other. But nothing in the physics licenses this precise judgment. All the physics says is that motion in each interval is "possible."[6] Once again, we have passed from being inductive logicians to being physicists, adding more physical properties to the system than Newtonian theory provides.

2.5 How to Infer Inductively about Indeterministic Systems

Something has gone very wrong. It is that we are trying to force the wrong inductive logic onto the indeterministic systems. How can we select the right one? The material theory of induction tells us that the prevailing material facts will fix the inductive logic, just as the probabilistic law of radioactive decay led us to a probabilistic logic for the radioactive decay of an atom. The physics of the indeterministic systems is more impoverished. So we should expect a more impoverished logic.

It is a mechanical exercise to read the relevant inductive logic from the physics. For radioactive decay, the chance of decay in time t=5τ is 0.99; so our degree of belief in that decay is 0.99. However the indeterministic physics of the dome and the masses and springs does not give us real-valued degrees. It just says that a spontaneous excitation at any time is possible; and that is all. It provides no degrees of possibility: not 50% possible, not 95% possible; and no comparative measures: not more possible, less possible, twice as possible. It just asserts what is possible; and by logical implication we can also know what is necessary and impossible. These three assignments, necessary, possible and impossible, become the three values of our inductive logic. The translation of the material facts in the physics to the inductive logic is illustrated in the table:

---

[6] A referee asked whether symmetry might warrant equal probabilities for equal intervals. The problem is that the resulting uniform distribution is not uniform under reparametrization, such as to inverse times. For more, see Norton (2008).

What the physics says:	What it induces in the inductive logic:
The present state does not fix the future (indeterminism). A future state is *necessary, possible or impossible*.	The inductive logic for the support [A|B] of A from B has three values: *nec, poss, imp*.
If the excitation happens in time [10,20), then it *necessarily* happens in [0,100).	[H(0,100) | H(10,20)] = *nec*
Excitation in any later time interval is *possible*, given E: the system is quiescent at t=0.	[H(0,10) | E] = [H(0,100) | E] = [H(10,20) | E] = … = *poss*
If the excitation happened in [0,10), it is *impossible* in [20,30).	[H(20,30) | H(0,10)] = *imp*

Table 1. Material Facts Dictate an Inductive Logic

The full logic is generated by the rules:

The complete inductive logic of the indeterministic systems is[7]

$$[A|B] = nec, \text{ if B entails A} \qquad (3)$$
$$= imp, \text{ if B entails not A}$$
$$= poss, \text{ otherwise}$$

2.6 Bayesian Response I: The Simulation Trick

There is a common Bayesian rejoinder. While the strengths [A\|B] of the inductive logic (3) are not conditional probability measures, we are just a few lines of mathematics away from them. Consider any probability measure at all that is adapted to the behaviors of the indeterministic systems through

$$P(A|B) = 1, \text{ if B entails A} \qquad (4)$$
$$= 0, \text{ if B entails not A}$$

and $\qquad 0 < P(A|B) < 1$ otherwise.

Every so adapted probability measure induces the logic (3) by

---

[7] This logic turns out to coincide with the distribution of complete ignorance as developed in Norton (2008).

$$[\ A|B\ ] \quad = nec, \ \text{if } P(A|B) = 1 \hspace{4cm} (5)$$
$$= imp, \ \text{if } P(A|B) = 0$$
$$= poss, \ \text{if } 0 < P(A|B) < 1$$

Obviously there are other ways to define the logic of Section 2.5 in terms of probability measures; finding them is simply a challenge to our ingenuity.

Has this sort of possibility shown us that the probability calculus is the One True and Universal Inductive Logic after all? It has not. The inductive logic of indeterministic systems is inherently non-additive. The degree of belief assigned to each of two mutually exclusive, contingent propositions is the *same* as the degree of belief assigned to their disjunction, if it is also contingent. If probability measures are to have meaning as a logic of induction, their additivity is their essence. We add the numerical probabilities of two mutually exclusive outcomes to find the probabilities of their disjunction.

What the exercise above shows is that we can take one sort of inductive logic, one of additive measures, and use it to simulate another, with the non-additive degrees (3). Since the additive measures of probability theory are now offered as devices for generating all other logics, they have ceased to be used a logic in their own right. They have been reduced to a useful adjunct tool in computation and are no more the Universal Inductive Logic than is the differential calculus.

Further, we can simulate the additive measures of probability theory by other adjunct tools. A trivial example is provided by complex valued, multiplicative measures M, for which we replace the additivity axiom of probability by a multiplication axiom: for mutually exclusive outcomes A and B, M(AvB) = M(A).M(B). These multiplicative measures can simulate additive measures through the formula P(A) = log Re(M(A)), so these multiplicative measures can replicate any result achievable with additive measures. However that fact in no way makes these new measures the Universal Logic of Induction.

2.7 Bayesian Response II: Subjectivism

The discussion so far has sought to present an intractable problem for Bayesians of all varieties. Perhaps subjective Bayesians specifically have an escape. They hold that probabilities may be assigned subjectively, initially, but that as we conditionalize on new evidence, the whim of our individual opinions will be overwhelmed by the weight of evidence. Why cannot a

subjective Bayesian assign a specific probability measure to the time of excitation? It merely represents that Bayesian's opinion and makes no pretense of being grounded in the facts. Why does that fail?

First there is a general problem with subjective Bayesianism, independent of this example. It changes the problem. Our original problem was to discern the bearing of evidence. That has been replaced by the problem of expressing one's opinion, in such a way that eventually the bearing of evidence will overwhelm opinion.

Second, in the context of the present example, the subjective project fails by its own standards. The subjective approach can only be relevant to an analysis of the bearing of evidence if there is no way to separate out mere opinion from the objective bearing of evidence in the probability measures; and if we have some reason to think that mere opinion will eventually be overwhelmed by the weight of evidence. Neither obtains.

Here, the separation of whim and warrant can be effected. The three-valued inductive logic (3) expresses precisely what the evidence warrants. In so far as the probability assigned goes beyond, it expresses mere opinion. The translation from probability measures to the three valued inductive logic (5) enables us to read precisely how it goes beyond. Any probability that lies strictly between 0 and 1 merely encodes the value *poss*. Anything more, including the specific numerical value, is opinion.

There will also be the familiar probabilistic dynamics as we conditionalize on new evidence. However these changing probabilities do not represent shifts in inductive warrant. The new evidence will be the observation of whether the excitation happens in successive time intervals. In so far as the resulting shifts in probability assignments leave the probabilities strictly between 0 and 1, they amount to pure shifts of opinion. If they force a probability assignment of 0 or 1, the shift is deductively generated, arising when the evidence deductively refutes or verifies the hypothesis.

3. Conclusion

The inductive logic of indeterministic systems such as the dome and the masses and springs is non-probabilistic. For that is the sort of logic that the material facts of the system require. Efforts to impose a probabilistic inductive logic on these systems either require us to presume facts about the systems that outstrip their full physical description; or to reduce the

probability measures to adjunct quantities in a simulation of an essentially different inductive logic.

One may doubt that the odd logic investigated has anything to do with our world, which, at least on the level of everyday experience, seems free of the indeterminism exemplified by the dome and the masses and springs. That is not the issue. Rather it is that these are perfectly well defined scenarios in which our ordinary modes of inference ought to be applicable. We have no hesitation in applying deductive inference to these indeterministic systems. Our deductive logics have universal scope and we fully expect the law of excluded middle to apply. If the probability calculus is the universal logic of induction, then we should expect it to apply to these indeterministic systems as well. It does not. Therefore the probability calculus is not the universal logic of induction.

There are systems for which the probability calculus provides a serviceable logic of induction; and there are systems for which it does not. As we proceed away from systems for which the probability calculus provides the appropriate inductive logic, we pass to systems that require different logics. The boundary may be sharp or it may be less well defined. The pressing problem for an inductive logician is to determine where this boundary lies and to be able to decide which side to place each new problem in inductive inference as it arises. The material theory of induction provides a direct solution to the problem: the facts prevailing in the domain of the problem fix the applicable inductive logic.

References

Alper, J. S., M. Bridger, J. Earman and J. D.Norton (2000), "What is a Newtonian System? The Failure of Energy Conservation and Determinism in Supertasks," *Synthese 124*: 281-293.

Lewis, D. (1980), "A Subjectivist's Guide to Objective Chance," in Richard C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*. Berkeley: University of California Press, pp. 263-93.

Norton, J. D. (2003), "A Material Theory of Induction," *Philosophy of Science 70*: 647-70.

Norton, J. D. (2003a), "Causation as Folk Science," *Philosophers' Imprint 3*(No.4); www.philosophersimprint.org/003004/; reprinted, H. Price and R. Corry (eds.), *Causation, Physics and the Constitution of Reality*. Oxford: Clarendon Press, 2007, Ch.2.

Norton, J. D. (2005), "A Little Survey of Induction," in P. Achinstein (ed.), *Scientific Evidence: Philosophical Theories and Applications*. Baltimore: Johns Hopkins University Press, pp. 9-34.

Norton, J. D. (2007), "Probability Disassembled," *British Journal for the Philosophy of Science*, *58*: 141-171.

Norton, J. D. (2008), "Ignorance and Indifference," *Philosophy of Science 75*: 45-68.

Norton, J. D. (2008a), "The Dome: An Unexpectedly Simple Failure of Determinism," *Philosophy of Science 75*(No. 5): 786-98.

Russell, B. (1957), *Why I am not a Christian*. George Allen and Unwin; repr. Touchstone.