Part 1

THE BASICS

Cambridge University Press
978-1-107-04230-8 - Pure Inductive Logic
Jeffrey Pari & Alena Vencovská
Excerpt
More information

Cambridge University Press
978-1-107-04230-8 - Pure Inductive Logic
Jeffrey Pari & Alena Vencovská
Excerpt
More information

Chapter 1

INTRODUCTION TO PURE INDUCTIVE LOGIC

Before a cricket match can begin the tradition is that the umpire tosses a coin and one of the captains calls, heads or tails, whilst the coin is in the air. If the captain gets it right s/he chooses which side opens the batting. There never seems to be an issue as to which captain actually makes this call (otherwise we would have to toss a coin and make a call to decide who makes the call, and in turn toss a coin and make a call to decide who makes that call and so on) since it seems clear that this procedure is fair. In other words both captains are giving equal probability to the coin landing heads as to it landing tails no matter which of them calls it. The obvious explanation for this is that both captains are, subconsciously perhaps, appealing to the *symmetry* of the situation.

At the same time they are, it seems, also tacitly making the assumption that all the other information they possess about the situation, for example the weather, the gender of the referee, even past successes at coin calling, is *irrelevant*, at least if it doesn't involve some specific knowledge about this particular coin or the umpires's ability to influence the outcome. Of course if we knew that on the last 8 occasions on which this particular umpire had tossed up this same coin the result had been heads we might well consider that that *was relevant.*

Forming beliefs, or subjective probabilities, in this way by considering symmetry, irrelevance, relevance, can be thought of as *logical* or *rational* inference. This is something different from statistical inference. The perceived fairness of the coin toss is clearly not based on the captains' knowledge of a long run of past tosses by the umpire which have favoured heads close to half the time. Indeed it is conceivable that this long run frequency might not give an average of close to half heads, maybe this coin is, contrary to appearances, biased. Nevertheless even if the captains knew that the coin was biased, provided that they also knew that the caller was not privy to which side of the coin was favoured, they would surely still consider the process as fair.

This illustrates another feature of probabilities that are inferred on logical grounds: they certainly need not agree with the long term frequency probability, if this even exists, and of course in many situations in which we form subjective probabilities no such probability does exist; for example when assigning odds in a horse race.

The aim of this monograph is to investigate this process of assigning logical, as opposed to statistical, probabilities by attempting to formulate the underlying notions, such as symmetry, irrelevance, relevance on which they appear to depend. Much has already been written by philosophers on these matters and doubtless much still remains to be said. Our approach here however will be that of mathematical, rather than philosophical, logicians. So instead of spending a significant time discussing these notions at length in the context of specific examples we shall largely consider ways in which they might be given a purely mathematical formulation and then devote our main effort to considering the mathematical and logical consequences which ensue.

In this way then we are proposing, or at least reviving since Rudolf Carnap had already introduced the notion in [14], an area of Mathematical Logic, *Pure Inductive Logic*, PIL for short.[1] It is not Philosophy as such but there are close connections. Firstly most of the logical, aka rational, principles we consider are motivated by philosophical considerations, frequently having an already established presence in the literature within that subject. Secondly we would hope that the mathematical results included here may feed back and contribute to the continuing debates within Philosophy itself, if only by clarifying that *if* you subscribe to A, B, C *then* you must, by dint of mathematical proof, accept D.

There is a parallel here with Set Theory. In that case we propose axioms based on our intuitions concerning the nature of sets and then investigate their consequences. These axioms have philosophical content and considering this is part of the picture but so also is drawing out their mathematical relationships and consequences. And as we go deeper into the subject we are led to propose or investigate axioms which initially might not have entered our minds, not least because we may well not have possessed the language or notions to even express them. And at the end of the day most of us would like to think that discoveries in Set Theory were telling us something about the universe of sets, or at least about possible universes of sets, and thus feeding back into the philosophical debate (and not simply generating yet more mathematics 'because it is there'!). Hopefully Pure Inductive Logic, PIL, will similarly tell us something about the universe of uncertain reasoning.

As far as the origins of PIL are concerned, whilst one may hark back to Keynes, Mill and even as far as Aristotle, in the more recent history of logical probability as we see it, W.E. Johnson's posthumous 1932 paper [58] in *Mind* was the first important contribution in the general spirit of what we

[1] For a gentle introduction to PIL see also [100].

are proposing here. It contains an initial assertion of mathematical conditions capturing intuitively attractive principles of uncertain reasoning and a derivation from these of what subsequently became known as Carnap's Continuum of Inductive Methods. Independently Carnap was to follow a similar line of enquiry in his [9], [12], which he developed further with [13], [15], [16], [17] into the subject he dubbed 'Inductive Logic'. Already in 1946 however N. Goodman's so called 'grue' paradox, see [35], [36] (to which Carnap responded with [10], [11]) threatened to capsize the whole venture by calling into question the very possibility of a purely logical basis for inductive inference[2]. Notwithstanding Carnap maintained his commitment to the idea of an Inductive Logic till his death in 1970 and to the present day his vision encourages a small but dedicated band of largely philosopher logicians to continue the original venture in a similar spirit, albeit in the ubiquitous shadow of 'grue'.

From the point of view of this text however 'grue' is no paradox at all, it is just the result of failing to make explicit all the assumptions that were being used. There is no isomorphism between premises involving grue and green (a point we will touch on again later in the footnote on page 177) because we have different background knowledge concerning grue and green etc. and it is precisely this knowledge which the paradox subsequently uses to announce a contradiction.[3] Indeed in his initial response to 'grue' Carnap had also stressed the importance of having all the assumptions up front from the start, what he called the 'Principle of Total Evidence', see [10, p138], [12, p211], known earlier as 'Bernoulli's Maxim', see [6, footnote 1, p215], [65, p76, p313].

Even so, 'grue' is relevant to this monograph in that it highlights a divergence within Carnap's Inductive Logic as to its focus or subject matter between *Pure Inductive Logic*, which is our interest in this monograph, and *Applied Inductive Logic*, which is the practical concern of many philosophers. The former was already outlined by Carnap in [14]; it aims to

[2]For the reader unfamiliar with this 'paradox' here is a pared down mathematician's version: Let *grue* stand for 'green before the 1st of next month, blue after'. Now consider the following statements:

All the emeralds I have ever seen have been green, so I should give high probability that any emerald I see next month will be green.

All the emeralds I have ever seen have been grue, so I should give high probability that any emerald I see next month will be grue.

The conclusion that advocates of this 'paradox' would have us conclude is that Carnap's hope of determining such probabilities by purely logical or rational considerations cannot succeed. For here are 'isomorphic' premises with different (contradictory even) conclusions so the conclusion cannot simply be a logical function of the available information.

[3]For example we learnt in school that emeralds are green and never heard anything about this possibly changing at some future date. In contrast if we had been talking here about UK Road Fund Licence discs changing to a new colour next January 1st there would have been a 'paradox' for the contrary reason!

study formal systems in the mathematical sense, devoid of explicit interpretation. Assumptions must be stated within the formal language and conclusions drawn only on the basis of explicitly given rules. On the other hand Applied Inductive Logic is intended as a tool, in particular, to sanction degrees of confirmation, within particular contexts. The language therein is interpreted and so carries with it knowledge and assumptions. What Goodman's Paradox points out is that applied in this fashion the conclusions of any such logic may be language dependent (see [136] for a straightforward amplification of this point), a stumbling block which has spawned a considerable literature, for example [131], [137], and which, within PIL, we thankfully avoid. In short then we might draw a parallel here with the aims and methods of Pure Mathematics as opposed to those of Applied Mathematics.

In the latter we begin with an immensely complicated real world situation, cut it down to manageable size by choosing what we consider to be the relevant variables and the relevant constraints, so ignoring a wealth of other information which we judge irrelevant, and then, *drawing on existing mathematical theories and apparatus*, we hopefully come up with some predictive or explicative formula. Similarly with Inductive Logic the applied arm has been largely concerned with proposing formulae in such contexts - prior probability functions, to provide answers. The value of these answers and the whole enterprize has been subject to near century long debate, some philosophers feeling that the project is fundamentally flawed. On the other hand it clearly finds new challenges with the advent of designing artificial reasoning agents. Be it as it may, PIL is not out to *prescribe* priors. Rather it is an investigation into the various notions of 'rationality' in the context of forming beliefs as probabilities. It is in this foundational sense that we hope this monograph may be of interest to philosophers and to the Artificial Intelligence community. Similarly to other mathematical theories, we would hope that it would serve to aid any researcher contemplating actual problems related to rational reasoning.

A rough plan of this monograph is as follows. In the early chapters we shall introduce the basic notation and general results about probability functions for predicate languages, as well as explaining what we see as the most attractive justification (de Finetti's Dutch Book argument) for identifying degrees of belief with probability. We will then investigate principles based on considerations of symmetry, relevance, irrelevance and analogy, amongst others, for *Unary Pure Inductive Logic*, that is for predicate languages with only unary relation symbols. This was the context for Inductive Logic in which Carnap et al worked and with only a very few exceptions it remained so until the end of the 20th century. In the second half of this monograph we will enlarge the framework to *Polyadic Pure Inductive Logic* and return to reconsider symmetry, relevance and irrelevance within this wider context.

The style of this monograph is mathematical. When introducing various purportedly rational principles we will generally give some fairly brief explanation of why they might be considered ‘rational’, in the sense that a rational agent should, or could, adhere to them, but there will not be an extended philosophical discussion.[4] It will not be our aim to convince the reader that each of them really is ‘rational’; indeed that would be difficult since in combination they are often inconsistent. We merely seek to show that they might be candidates for an expression of ‘rationality’, a term which we will therefore feel free to leave at an intuitive level. As these principles are introduced we will prove theorems relating them to each other and attempt to characterize the probability functions which satisfy them. Most proofs will be given in full although we sometimes import well known results from outside of Mathematical Logic itself. On a few occasions giving the proof of a theorem in detail would just be too extravagant and in that case we will refer the reader to the relevant paper and content ourselves instead by explaining the key ideas behind the proof. In any case it is our intention that this monograph will still be accessible even to someone who wishes to treat the proofs simply as the ‘small print’.

[4]There are a number of books which do provide extended philosophical discussions of some of the general principles we shall investigate, for example [12], [13], [15], [16], [17], [18], [22], [27], [29], [33], [37], [38], [40], [57], [72], [127], [135], [140], [148].

Cambridge University Press
978-1-107-04230-8 - Pure Inductive Logic
Jeffrey Pari & Alena Vencovská
Excerpt
More information

Chapter 2

CONTEXT

For the mathematical setting we need to make the formalism completely clear. Whilst there are various possible choices here the language which seems best for our study, and corresponds to most of the literature, including Carnap's, is where we work with a first order language L with variables $x_1, x_2, x_3, \ldots$, relation symbols $R_1, R_2, \ldots, R_q$, say of finite arities $r_1, r_2, \ldots, r_q$ respectively, and constants a_n for $n \in \mathbb{N}^+ = \{1, 2, 3, \ldots\}$, and no function symbols nor (in general) the equality symbol.[5] The intention here is that the a_i name all the individuals in some population though there is no prior assumption that they necessarily name different individuals. We identify L with the set $\{R_1, R_2, \ldots, R_q\}$.

Let SL denote the set of first order sentences of this language L and $QFSL$ the quantifier free sentences of this language. Similarly let FL, $QFFL$ denote respectively the set of formulae, quantifier free formulae, of L. We use θ, ϕ, ψ etc. for elements of FL and adopt throughout the convention that, unless otherwise stated, when *introducing*[6] a formula $\theta \in FL$ as $\theta(a_{i_1}, a_{i_2}, \ldots, a_{i_m}, x_{j_1}, x_{j_2}, \ldots, x_{j_m})$, or $\theta(\vec{a}, \vec{x})$, *all* the constant symbols (respectively free variables) in θ are amongst these a_i (x_j) and that they are distinct. In particular if we write a formula as $\theta(\vec{x})$ then it will be implicit that no constant symbols appear in θ. To avoid double subscripts we shall sometimes use $b_1, b_2, \ldots$ etc. in place of $a_{i_1}, a_{i_2}, \ldots$.

Let $\mathcal{T}L$ denote the set of structures for L with universe $\{a_1, a_2, a_3, \ldots\}$, with the obvious interpretation of the a_i as a_i itself. Notice that if $\Gamma \subseteq SL$ is consistent[7] and infinitely many of the constants a_i are not mentioned in any sentence in Γ then there is $M \in \mathcal{T}L$ such that $M \vDash \Gamma$. This follows since the countability of the language L means that Γ must have a countable model, and hence, by re-interpreting the constant symbols not mentioned in any sentence in Γ, a model in which every element of its universe is the interpretation of at least one of the constant symbols.

[5] We shall add equality in Chapter 37 but will omit function symbols throughout, that being a topic which, in our opinion, is still deserving of more investigation and thought.

[6] So this will not apply if we introduce a sentence as $\exists x\, \psi(x)$ and then pass to $\psi(a_n)$. In this case there may be other constants mentioned in $\psi(a_n)$.

[7] Because of the Completeness Theorem for the Predicate Calculus we shall use consistent and satisfiable interchangeably according to which seems most appropriate in the context.

To capture the underlying problem that PIL aims to address we can imagine an agent who inhabits some structure $M \in \mathcal{T}L$ but knows nothing about what is true in M. Then the problem is,

Q: *In this situation of zero knowledge, logically, or rationally, what belief should our agent give to a sentence* $\theta \in SL$ *being true in* M?

There are several terms in this question which need explaining. Firstly 'zero knowledge' means that the agent has no intended interpretation of the a_i nor the R_j. To mathematicians this seems a perfectly easy idea to accept; we already do it effortlessly when proving results about, say, an arbitrary group. In these cases all you can assume is the axioms and you are not permitted to bring in new facts because they happen to hold in some particular group you have in mind. Unfortunately outside of Mathematics this sometimes seems to be a particularly difficult idea to embrace and much confusion has found its way into the folklore as a result.[8]

In a way this is at the heart of the difference between the 'Pure Inductive Logic' proposed here as Mathematics and the 'Applied Inductive Logic' of Philosophy. For many philosophers would argue that in this latter the language is intended to carry with it an interpretation and that without it one is doing Pure Mathematics not Philosophy. It is the reason why Grue is a paradox in Philosophy and simply an invalid argument in Mathematics. Nevertheless, mathematicians or not, we all need to be on our guard against allowing interpretations to slip in subconsciously. Carnap himself was very well aware of this distinction, and the dangers presented by ignoring it, and spent some effort explaining it in [14]. Indeed in that paper he describes Inductive Logic as the study of the rational beliefs of just such a zero knowledge agent, a 'robot' as he terms it.

A second unexplained term is 'logical' and its synonym (as far as this text is concerned) 'rational'. In this case, as already mentioned, we shall offer no definition; they are to be taken as intuitive, something we recognize when we see it without actually being able to give it a definition. This will not be a great problem, for our purpose is to propose and mathematically investigate principles for which it is enough that we may simply *entertain* the idea that they are logical or rational. The situation parallels that of the intuitive notion of an 'effective process' in recursion theory, and similarly we may hope that our investigations will ultimately lead to a clearer understanding.

The third unexplained term above is 'belief'. For the present we shall identify belief, or more precisely degree of belief, with (subjective) probability and only later provide a justification, the Dutch Book Argument, for this identification. The main reason for proceeding in this way is that in order to give this argument in full we actually need to have already developed some of the apparatus of probability functions, a task we now move on to.

[8] See for example [106] and the issue of the representation dependence of maxent.