View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Philsci-Archive

THE META-INDUCTIVE JUSTIFICATION OF INDUCTION: THE POOL OF STRATEGIES

TOM F. STERKENBURG

Abstract. This paper poses a challenge to Schurz's proposed meta-inductive justification of induction. It is argued that Schurz's argument requires a notion of optimality that can deal with an expanding pool of prediction strategies.

1. Introduction

Schurz (2008; 2009; most recently, 2018; 20xx) proposes a justification of induction based on *meta-induction*, induction at the level of competing methods of inference. The argument proceeds in two steps. First, there is the *analytical* justification of meta-inductive strategies in the setting of sequential prediction. This consists in mathematical results on these strategies' optimality, as established in the machine learning branch of *prediction with expert advice* (see Cesa-Bianchi and Lugosi, 2006; Vovk, 2001). Second, there is the *empirical* observation that *object-induction*, induction at the level of events, has been most successful so far. Hence, the argument goes, the optimal meta-inductive strategy favors the object-inductive strategy, thus justifying it.

Schurz's proposal is a refinement of Reichenbach's attempted *pragmatic justification* or *vindication* of induction (see Salmon, 1967, 52ff, 85ff). The fundamental idea underlying both is that the aim for *reliability*, guaranteed success, may be replaced for *optimality*, guaranteed success *whenever some method would be successful*. This weaker aim is still reasonable, because the cases in which *no* method can be successful are in an obvious sense not so interesting—in those cases there is simply nothing we could do. And, importantly, this weaker aim looks more feasible: while it appears impossible to design a single inductive method that can take into account everything nature could possibly do (this is in a sense the original problem of induction, see Howson, 2000), it looks more feasible to design a single method that tracks *what we could possibly do*. Thus Schurz (2018, 3895) proclaims that "optimality justifications constitute new foundations for foundation-oriented epistemology."

The obvious qualm is whether the aim of a truly general optimality is really more feasible. This qualm finds a sharp expression in the question *what class of methods* we should actually require optimality *for*. In this paper, I investigate this question within the context of Schurz's argument. My conclusion will be that the argument needs an optimality that covers *expanding* pools of strategies, which suggests that things may not be easier, after all.

The plan of the paper is as follows. First, I will briefly describe the presupposed framework of sequential prediction (sect. 2) and the structure of Schurz's argument

Date: October 30, 2018.

(sect. 3). (This is based on a much more detailed reconstruction of the argument elsewhere, Sterkenburg, 20xx.) In order to constitute an actual justification for object-induction, the conclusion of the argument also needs us to accept that the *optimality* of the meta-inductive strategy amounts to a *justification* for it. The question whether this is really so then prompts us to have a closer look at the pool of prediction strategies assumed.

I start with the objection due to Arnold (2010) that the optimality results that Schurz relies on are restricted to finite pools of strategies (sect. 4). I point out why Schurz's argument, to go through at all, *must* presuppose a finite pool; but I argue that it does not *need* an infinite pool to yield the desired justification: it only needs optimality relative to the (necessarily finitely many) actually proposed alternatives. However, I then argue (sect. 5) that this does involve something more: it needs a notion of optimality that is robust against *new* strategies being proposed over time.

2. The framework of prediction

2.1. **The framework of sequential prediction.** We define a prediction game as a triple $(\boldsymbol{y}^\omega, \Pi, \ell)$ of a *history* $\boldsymbol{y}^\omega$, a pool Π of *prediction strategies*, and a *loss function* ℓ.

A *history* $\boldsymbol{y}^\omega$ is an infinite sequence of *events*. Events are identified with values in some set Val of possible values. Write y_n for the n-th element of $\boldsymbol{y}^\omega$, or the event in *round* n of the game ($n \in \mathbb{N}^{>0}$).

Predictions are elements in some set $\mathrm{Val}_{\mathrm{pred}}$. A *prediction strategy* P, an element of the pool Π, specifies in each round n a prediction $\mathrm{pred}_n(P)$ about the next event.

In this paper, I will restrict attention to the central class of *probabilistic prediction games*. In these games we assume binary events, $\mathrm{Val} = \{0, 1\}$, and predictions that are probabilities (for the next event being 1, say), $\mathrm{Val}_{\mathrm{pred}} = [0, 1]$.

Strategy P, when making prediction $\mathrm{pred}_n(P) = \mathrm{pred} \in \mathrm{Val}_{\mathrm{pred}}$ for round n, suffers, when the outcome is revealed to be $y_n \in \mathrm{Val}$, a certain *loss* $\ell(\mathrm{pred}, y_n)$. That is, a loss function $\ell : \mathrm{Val}_{\mathrm{pred}} \times \mathrm{Val} \to [0, \infty)$ quantifies how much a prediction was off in light of the actual outcome.

A basic example is the *absolute* loss function, defined by $\ell_{\mathrm{abs}}(\mathrm{pred}, y) = |\mathrm{pred} - y|$. Another loss function, prominent, among other things, for its strong connection to *Bayesian* prediction (sect. 3.1 below), is the *logarithmic* or *log*-loss function defined by

$$\ell_{\log}(\mathrm{pred}, y) = \begin{cases} -\ln(1 - \mathrm{pred}) & \text{if } y = 0 \\ -\ln \mathrm{pred} & \text{if } y = 1 \end{cases}.$$

For given loss function, the *cumulative* loss of P by the conclusion of round n is the sum $\mathrm{Loss}_n(P) := \sum_{i=1}^{n} \ell(\mathrm{pred}_n(P), y_n)$. The *loss rate* $\mathrm{loss}_n(P)$ of P by n is the average $\mathrm{Loss}_n(P)/n$ of its losses up to n.

2.2. **The goal: an optimal strategy.** Given a pool Π of prediction strategies, we aim to design a *meta-inductive* strategy MI that, having access to the predictions of all the other strategies, predicts in such a way that it is *optimal* with respect to Π. That is, by following MI we will *always* do about as good as, in hindsight, we possibly could have done—given that the strategies in Π represent what we could have done. Here 'always' means: on *every* single history of events.

What it means for a meta-inductive strategy to be 'about as successful' as any other strategy we make precise in terms of the divergence between MI's loss rate and the quantity $\text{minloss}_n := \min_{P \in \Pi \cup \{\text{MI}\}} \text{loss}_n(P)$, the minimum loss rate among all the strategies (including MI itself) by round n. Specifically, we seek a function f, that depends on n and inevitably also on the size $K := |\Pi|$ of the pool of strategies, such that for all rounds n,

$$\text{loss}_n(\text{wMI}) \leq \text{minloss}_n + f(n, K). \tag{1}$$

A minimal requirement is that f is such that it entails *long-run convergence*,

$$\lim_{n \to \infty} (\text{loss}_n(\text{MI}) - \text{minloss}_n) = 0, \tag{2}$$

for which it at least needs to be decreasing in n. But as we will see below, there actually exist prediction algorithms that achieve bounds (1) for f that decrease in n at a very fast rate, giving strong *short-run* guarantees.

3. The argument

3.1. Step one: the analytical optimality of meta-induction.

A general type of meta-inductive strategy is the *weighted-average* strategy waMI, specified by

$$\text{pred}_{n+1}(\text{waMI}) := \sum_{P \in \Pi} w_n(P) \cdot \text{pred}_{n+1}(P). \tag{3}$$

Here the *weight function* w_n assigns a weight to each strategy P based on its past success.

An important example of a weighted-average strategy in the probabilistic binary prediction game, for the particular choice of the log-loss function, is the *Bayesian strategy* BayMI, that updates its weights via Bayes's rule. It is given by

$$w_n(P) = \frac{w_0(P) \cdot \exp(-\text{Loss}_n(P))}{Z}, \tag{4}$$

with normalization term $Z = \sum_{P \in \Pi} w_0(P) \cdot \exp(-\text{Loss}_n(P))$. Here w_0 is some prior probability assignment or *initial weight function* over Π. With a *uniform* initial weight assignment, where $w_0(P) = 1/K$ for each $P \in \Pi$, assignment (5) simplifies to

$$w_n(P) = \frac{\exp(-\text{Loss}(P))}{Z}, \tag{5}$$

so that the weights depend on the strategies' performance only.

Now one can derive that BayMI, for the log-loss function, satisfies, for each $P \in \Pi$,

$$\text{Loss}_n(\text{BayMI}) \leq -\ln w_0(P) + \text{Loss}_n(P). \tag{6}$$

Choosing again a uniform w_0, this translates in the short-run optimality bound

$$\text{loss}_n(\text{BayMI}) \leq \text{minloss}_n + \frac{\ln K}{n}. \tag{7}$$

That is, for this game we can achieve bound (1) with f of order $1/n$. What is more, it turns out to be possible, for a wider class of loss functions, to design strategies that explicitly mimic the Bayesian strategy for the log-loss function, for *these* loss functions, in order to achieve a similar bound. Thus for these so-called *mixable* loss functions, which include the quadratic loss function, there also exist meta-inductive strategies with bounds of order $1/n$. These are the strongest possible bounds for any game; but for an even wider class of loss functions, that also includes the absolute loss function, it is still possible to define meta-inductive strategies—specifically, *exponentially-weighted* strategies that can also be seen as generalizations of the Bayesian strategy—with bounds of order $1/\sqrt{n}$.

Taking stock, we have that for a wide class of games there exist meta-inductive strategies that are optimal in a very strong sense. Moreover, these optimal strategies predict by combining weighted predictions of all the other strategies in the pool, where the weights depend on these strategies' attractiveness or past performance—and in the case of uniform weights, on their past performance *only*. In particular, the strategy in the pool that so far has been performing best receives the largest weight: it is in that sense that we say that the meta-inductive strategy *favors* the most successful strategies so far. Thus the first step of Schurz's argument is that

(A) The meta-inductive strategy MI, that at each point in time favors strategies to the extent of their relative success so far, is an optimal method.

3.2. **Step two: the empirical success of object-induction.** The second step is the empirical observation that "so far object-induction has turned out to be the most successful prediction strategy" (Schurz, 2008, 304).

In (20xx), I argued that the relevant perspective here is to view the object-inductive or *scientific* method as competing with a number of alternative *non*scientific methods. Importantly, for Schurz's argument it is not necessary to further specify what this scientific method actually consists in, the notorious problem of description (see, e.g., Lipton, 2004). It is enough to recognize that there is something like the scientific procedure, that we wish to find justification for; and, plausibly, that its predictions have been highly successful so far, at least more successful than those of nonscientific alternatives. Thus the second step of Schurz's argument is that

(E) As a matter of empirical fact, the object-inductive strategy OI, that we identify with the scientific method (and that we imagine to be in competition with various proposed nonscientific methods), has been, at this point in time, the most successful prediction strategy (among the pool Π of all of these competing strategies).

3.3. **Conclusion: meta-induction favors object-induction.** From (A) and (E) it follows that

(C) The meta-inductive strategy MI for the pool Π of OI and its nonscientific competing strategies, an optimal strategy for Π, favors most, at this point in time, the object-inductive strategy OI.

In (20xx), I noted that, for (C) to yield the desired justification of OI, we also need to say that an optimal strategy favoring OI actually amounts to a justification for it. The discussion of this step brings out an important limitation of the argument: it cannot provide a justification for the object-inductive *strategy* (for *always* sticking with object-induction), but at best—though this would still be an important result—a justification for sticking with the object-inductive prediction *for now* (thus allowing for the possibility that in the more distant future it will no longer be a good strategy to follow).

Furthermore (ibid.), I noted that we would also still need to argue that the optimality of the meta-inductive strategy actually amounts to a justification for following it. I allowed that the notion of optimality is sufficiently strong that it does—*given* that the pool of strategies is appropriate, a proper rendition of all we could possibly do. We will now investigate whether this is truly so.

4. The restriction to a finite pool

Arnold (2010) points out that the analytical justification of meta-induction does not extend to pools of *infinitely* many strategies, and suggests that this is a problem for Schurz's proposal.

4.1. **The impossibility result.**

Arnold's observation, his impossibility theorem 3 (ibid., 589), comes down to the following. For every strategy MI we might propose, nature can construct an adversarial history that makes it fail maximally: in each round, it can choose $y = 0$ precisely if our strategy's pred > 0.5. Then our strategy's total loss grows linearly, and its loss rate $\mathrm{loss}_n(\mathrm{MI})$ never goes to 0. However, for a rich enough infinite pool of strategies, say a pool that includes all *computable* strategies, there exists for *every* finite history (including every finite initial segment of the adversarial history we are constructing), some strategy that has managed to predict this history *perfectly*. That is, $\mathrm{minloss}_n = 0$ for every round n. Hence our strategy's loss rate does not converge to the best strategy's.

This shows that optimality is impossible to achieve in the general case of infinite pools of strategies: at least for sufficiently rich such pools, we can for any given strategy construct a history that refutes its optimality.

4.2. **Universal but non-uniform optimality.**

Arnold writes (ibid., 592, emphasis mine), "If only a finite number of prediction strategies are taken into account, then we exclude the overwhelming majority of *possible* prediction strategies from the game right from the beginning." Arnold's suggestion that Schurz's argument needs a notion of optimality relative to infinite pools of strategies thus appears to be motivated by a more definite demand: the argument would need a notion of optimality relative to the infinite pool of *all possible strategies*. The argument would need an optimality that is no longer relative but truly *universal*.

The general move from reliability to optimality actually makes universality look genuinely more feasible. Namely, it seems reasonable to "take into consideration only those prediction strategies that can be described by an algorithm" (ibid.; see Sterkenburg, 2018 for more details). While there seems little justification for limiting possible histories to computable sequences of events, it does seem reasonable to limit the methods of prediction we could possibly devise to the computable ones. Rather than the continuum of all possible histories, we then only need to consider the vastly more restricted pool of computable prediction strategies.

Of course, this is also still a countably infinite number, and so optimality in the original sense is ruled out by the impossiblity result above. However, we can still attain a weaker, *non-uniform* optimality for countably infinite pools.

Consider again the Bayesian strategy in the log-loss game, and the bound (6) on its cumulative loss: notably, this bound holds just as well for an initial weight assignment over a countably *infinite* pool of strategies. Thus even in case of an infinite pool Π, we can still derive, parallel to (7), that for every strategy $P \in \Pi$, for all n,

$$\mathrm{loss}_n(\mathrm{BayMI}) - \mathrm{loss}_n(P) \leq \frac{-\ln w_0(P)}{n}, \tag{8}$$

so that in particular, for all $P \in \Pi$,

$$\lim_{n\to\infty} \left(\mathrm{loss}_n(\mathrm{BayMI}) - \mathrm{loss}_n(P)\right) \leq 0. \tag{9}$$

The crux here is that (8) depends on the given predictor, specifically, on the initial weight $w_0(P)$ assigned to it. Now in case of only a *finite* number of strategies, it is possible to *uniformly* assign each the *same* initial weight, and we can derive a uniform bound (1), and consequently uniform convergence (2). But in case of an infinite number of strategies, this is obviously impossible: we are forced to give some non-uniform initial weight assignment. Consequently, the convergence (9) is non-uniform: we are *not* guaranteed to eventually match the success of all strategies in the pool, *at the same time*. We are guaranteed, for any given strategy in Π, to eventually match the success of this strategy, but by the time we do other strategies might always still be way ahead of us; this is the reason why this bound is consistent with the impossibility result of sect. 4.1 above.

But it is still something—given a pool Π, and in the absence of stronger guarantees, one can argue there is some justification for sticking to a non-uniformly optimal strategy, for sticking to a strategy that is guaranteed for any selected strategy from Π to eventually match this strategy's success. Granted this, there is certainly some justification for sticking to a strategy that is guaranteed for *any possible strategy* to eventually match this strategy's success—for a *universally* non-uniformly optimal strategy.

Now if we identify all possible strategies with the computable ones, then the Bayesian strategy over all computable strategies, that is non-uniformly optimal relative to all computable strategies, would be universally non-uniformly optimal. Could this then be a strategy that meets Arnold's demand?

Unfortunately, it cannot, and the reason is that this Bayesian meta-inductive strategy is no longer computable itself. This follows from a diagonal argument that goes back to Putnam (1963), an impossibility argument that is actually very similar to that of sect. 4.1 above. What it means is that, on our earlier restriction of the possible strategies to the computable ones, the candidate optimal strategy is actually no longer a proper strategy; nor is any optimal strategy for the pool of computable strategies. This quandary holds with great generality: it is not restricted to the log-loss function, and we cannot escape it by looking for weaker computability constraints (Sterkenburg, 2018). Thus Arnold's demand is, indeed, unrealizable, even on a weaker notion of non-uniform optimality: there cannot be a universally optimal prediction strategy.

4.3. **Infinite pools in Schurz's argument.** On reading Arnold's presentation, one gets the impression that Schurz's proposed justification of induction boils down to the description of an optimal strategy. In contrast to "[m]ost of the proposed solutions to the problem of induction [that] tried to prove the reliability of the inductive procedure," he writes, "Schurz, following Reichenbach, merely tries to show the optimality of a specific inductive strategy" (2010, 585). With the understanding that this must be *universal* optimality, such a project, we just discussed, is indeed doomed to fail.

But Schurz's actual argument is more subtle than that. As explained in sect. 3 above, the argument seeks to justify object-induction, a strategy that is presumably not optimal itself. The meta-inductive strategy, optimal relative to all of OI's competitors, only comes in to confer justification to OI. Now even if one insists that OI's competitors are *all possible* strategies, things might still look better for Schurz's actual argument—perhaps, for instance, it is not so important here that a universally optimal strategy cannot actually be a proper (i.e., computable) strategy itself? But we can save ourselves the trouble of going into this: unfortunately, there is a more direct reason why the pool of competitors *must* be finite, for Schurz's actual argument to work.

To see this, we return to the observation in sect. 4.2 above that the optimality bound (7), and in general a bound (3) for any meta-inductive strategy, must involve an initial weight assignment w_0. It is only with a uniform prior, which is only possible for a finite pool, that the initial weights all cancel out and P's weights in later rounds depend on its success *only*. Thus for a countably infinite pool of strategies, a meta-inductive strategy *must* express some prior preference for some strategies above others, that works through in the posterior weights.

But this is devastating to Schurz's argument. Conclusion (C) follows from step (E) if the optimal strategy at this point in time favors the most successful strategy, OI. In case of a finite pool of strategies, where the weights are only determined by the success, it does. But in case of an infinite pool, the meta-inductive strategy only favors OI at this point of time *if it assigned strategy* OI *a sufficiently high initial weight*. The meta-inductivist will *not* favor OI, even if OI has been the most successful strategy, if it assigned OI too low an initial weight (and conversely, it *would* favor OI, even if OI had *not* been successful at all, if this were compensated by a high enough initial weight). In short, the meta-inductive justification of object-induction would have to presuposse a sufficiently strong prior preference for object-induction, and this would render it an obviously circular argument.

4.4. **The finite pool of actually proposed alternatives.** Thus in the end Arnold is right to worry that Schurz's argument is not compatible with an infinite pool of competing strategies: indeed it is not. This leads us to "the philosophical question whether an optimality result demonstrated for a finite number of prediction strategies might suffice to answer the problem of induction" (ibid., 585).

Schurz writes, "I make the realistic assumption that [the meta-inductive strategy] has finite computational means, whence I restrict my investigation to prediction games with finitely many strategies" (2009, 206; also see 2008, 284). More precisely, Schurz (2018, 3891) offers in defense of the limitation to finite pools an

> *Argument from cognitive finiteness*: Epistemic subjects are assumed to be finite beings. Finite beings can simultaneously access (and compare) only finitely many methods of finite complexity. Therefore the optimality justification of meta-induction is not affected by the finiteness restriction.

The second statement is not strictly true, though, and anyway does not entail the desired conclusion. It is not strictly true, because finite computational means are consistent with weighing over an infinite enumeration of strategies. (We can plausibly only give probabilistic—real-valued—predictions up to some finite accuracy, and since we also have to give decreasing weights to the strategies in the enumeration, there are in each round only finitely many strategies that can have an impact; yet this is different from stipulating a finite pool from the start.) But more importantly, as Arnold noted already, the observation that a meta-inductive strategy can only deal with finitely many strategies falls short of a justification for this restriction: in itself, this "merely amounts to admitting that under this 'realistic assumption' [the meta-inductive strategy] simply cannot always perform optimally" (2010, 592).

Arnold continues, "[a]s there is no logical contradiction involved in the assumption of an infinite number of alternative strategies, the only grounds on which it could be defended are empirical" (ibid.). These are exactly the grounds, I will now argue, on which it *can* be defended, in the context of Schurz's actual argument.

Again, Schurz's argument is not to identify a universally optimal strategy, optimal among the infinity of all possible strategies; it is to justify object-induction, from the empirical observation (E) that object-induction has been most successful so far. Most successful among what? Certainly *not* among all possible strategies—we can probably conceive, in hindsight, of strategies that would have been more successful still. No: object-induction has been most successful, so far, among *all actually proposed alternative strategies*. The relevant empirical observation (E) is that object-induction has been most successful among the various actually proposed nonscientific strategies—of necessity a *finite* number of strategies.

It is in this sense that Schurz (2018, 3891) is surely right when he, after his initial and unconvincing defense, adds that "[i]n any case, the problem of choosing among finitely many competing methods captures the most important part of the induction problem." Now the problem is to give a good reason for sticking to OI, rather than turning to one of its contestant strategies; and the hope, again, is to derive such a noncircular reason with the help of the optimality of a meta-inductive strategy, that by (E) favors it. But then it seems enough to have an optimality relative to *this same pool* of all actually proposed strategies. The pool of all actually proposed strategies seems to properly represent all we could have done, and so an optimal strategy for this pool would be justified.

Unfortunately, there is still a crucial sense in which this optimality falls short of including all we could have done.

5. The restriction to a fixed pool

The basic intuition, again, behind the optimality of the meta-inductivist over the pool of all proposed strategies, is the one going back to Reichenbach: for every possible history, and for every alternative strategy proposed, if this strategy is successful, the meta-inductivist will mimic it and be successful, too. Thus Schurz (2008, 304) concludes by once more evoking this intuition to answer the obvious skeptical reservation: "how can it *ever* be possible to prove that a strategy is optimal with respect to *every* other accessible strategy in *every* possible world—without assuming anything about the nature of alternative strategies and possible worlds?" To understand how this is possible, Schurz answers, one should note that "meta-induction has an unlimited learning ability: whenever this strategy is confronted with a so far better method, it will learn from it and reproduce its success" (2018, 3892).

There is, however, a clear sense in which *this is not true*: namely, when the meta-inductivist is confronted with a *new* strategy.

5.1. **The expanding pool of actually proposed alternatives.**

A meta-inductivist can be optimal for a finite pool of strategies, like the finite pool of actually proposed strategies, but, crucially, we need to assume that this pool is *fixed*. Yet it is only plausible that the pool of actually proposed strategies will *expand* in the course of time: informed by the actual history of events, brand *new* strategies may be proposed.

The meta-inductive strategies we have been considering cannot guarantee optimality with respect to new strategies—simply because they do not allow for dynamically incorporating new strategies in their pool. Imagine that we fix the pool of strategies that have been proposed by this time in history, and design and follow a meta-inductive strategy that is optimal relative to this pool. But in the future a new strategy might be proposed, and this strategy might continue to be forever much more successful than all the original strategies—and hence than our meta-inductive strategy. This means that our meta-inductive strategy is no longer optimal in the sense of being as good as we can possibly be: surely we could have followed the new and much more successful strategy instead.

5.2. **Truly analytical optimality.** Is this really a problem for Schurz's argument, though? Was the goal, specified above, not to justify following object-induction among the alternative strategies we have *now*?

Yes, this is still the goal, and the relevant empirical fact (E) is still that object-induction has been most successful among the alternatives we have *now*. However, I now claim, the *analytical* step (A), to be truly analytical, must involve a notion of optimality that is robust against all possible empirical circumstances: against all possible histories of events, but also against *all possible evolutions of the pool of strategies*.

Again, the crucial component of analytical optimality is that it covers every possible history: it should not and does not depend on the contingent fact of the actual history of events we have seen occur. But likewise, it should not depend on the contingent fact of the actual alternative strategies that we have seen proposed. This does *not* mean that we must demand optimality relative to all possible finite pools of alternative strategies *at the same time* (this would be Arnold's infeasible demand of optimality relative to all possible strategies); but it does mean that we must demand optimality relative to all possible *expanding pools*, or histories of finite pools.

Otherwise, the meta-inductive method is simply not optimal in the sense of analytically the best we could do. The meta-inductive method that we fix at this point of time, relative to the current pool of alternative strategies, was not guaranteed to be optimal: it might not have been if other, better, strategies had been proposed in the past. And it might still fail to be, if other, better, strategies are proposed in the future. As such, it is not a strategy one is justified to follow without any empirical assumptions, and it cannot fulfill the analytical role required for Schurz's argument.

5.3. **Dynamic optimality.** What are the prospects for the design of a 'dynamic' meta-inductive strategy that *is* optimal in the above sense?

Such a strategy must allow for dynamically adding new strategies to its pool as they appear, while somehow preserving optimality guarantees with respect to all the available strategies in each round. There are some choices to be made here, starting with a suitable standard of optimality.

It appears too strict, for instance, to demand that the meta-inductivist keep its loss rate low with respect to new strategies on *past data*: it cannot, of course, guard itself against new strategies that simply fit their past predictions to the past data and thereby can claim to have a perfect score. This demand indeed goes beyond a notion of optimality as the best we could do, since we could only have followed a strategy from the moment it is actually available. On the other hand, it also appears infeasible to first measure the success of a new strategy from the moment it comes in. As an extreme scenario: in each round a new strategy appears that makes an initial perfect prediction and then stops predicting well; now after each round the best strategy again has a perfect score while the meta-inductivist has not necessarily been doing very well.

So perhaps we need to argue for a middle way, where new strategies are assigned some 'virtual' loss for the rounds where they were not yet participating. This is indeed an approach taken in the literature that comes closest to our problem, the framework of 'specialists,' experts that are in each round allowed to 'sleep' and refrain from making predictions (Freund et al., 1997). The 'abstention trick' due to Chernov and Vovk (2009) advocates the assignment to asleep strategies of the same predictions as the meta-inductivist; using this trick Mourtada and Maillard (2017) derive bounds for the specific case of growing expert pools. However, it remains to be argued that these results are truly applicable to the current context: that they can still support both the analytical and the empirical step of the argument. I will leave this here at this briefest of sketches, and suggest further investigation as a challenge for Schurz's research programme.

6. Conclusion

I identified as a challenge for Schurz's proposed meta-inductive justification of induction the need for a notion of optimality that is robust against newly proposed prediction strategies. Notably, this challenge finds a parallel in the problem of new theory in the traditional Bayesian framework (Earman, 1992, 195ff; Gillies, 2001). This suggests that the aims of reliability and of optimality are confronted with much the same structural difficulties, and that, unless this challenge can indeed be met, a shift of focus to optimality might not be such an effective means of avoiding foundational problems as Schurz advocates.

References

E. Arnold. Can the best-alternative justification solve Hume's problem? On the limits of a promising approach. *Philosophy of Science*, 77(4):584–593, 2010.

N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning and Games*. Cambridge University Press, Cambridge, 2006.

A. Chernov and V. G. Vovk. Prediction with expert evaluators' advice. In R. Gavaldà, G. Lugosi, T. Zeugmann, and S. Zilles, editors, *Proceedings of the 20th International Conference on Algorithmic Learning Theory*, volume 5809 of *Lecture Notes in Computer Science*, pages 8–22. Springer, 2009.

J. Earman. *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. A Bradford Book. MIT Press, Cambridge, MA, 1992.

Y. Freund, R. E. Schapire, Y. Singer, and M. K. Warmuth. Using and combining predictors that specialize. In *Proceedings of the 29th Annual ACM Symposion on the Theory of Computing*, pages 334–343, New York, 1997. ACM Press.

D. A. Gillies. Bayesianism and the fixity of the theoretical framework. In D. Corfield and J. Williamson, editors, *Foundations of Bayesianism*, volume 24 of *Applied Logic Series*, pages 363–379. Springer, 2001.

C. Howson. *Hume's Problem: Induction and the Justification of Belief*. Oxford University Press, New York, 2000.

P. Lipton. *Inference to the Best Explanation*. Routledge, London, second edition, 2004.

J. Mourtada and O.-A. Maillard. Efficient tracking of a growing number of experts. In S. Hanneke and L. Reyzin, editors, *Proceedings of the 28th International Conference on Algorithmic Learning Theory*, volume 76 of *Proceedings of Machine Learning Research*, pages 517–539. PMLR, 2017.

H. Putnam. 'Degree of confirmation' and inductive logic. In P. A. Schilpp, editor, *The Philosophy of Rudolf Carnap*, volume XI of *The Library of Living Philosophers*, pages 761–783. Open Court, LaSalle, IL, 1963.

W. C. Salmon. *The Foundations of Scientific Inference*. University of Pittsburgh Press, Pittsburgh, PA, 1967.

G. Schurz. The meta-inductivist's winning strategy in the prediction game: A new approach to Hume's problem. *Philosophy of Science*, 75(3):278–305, 2008.

G. Schurz. Meta-induction and social epistemology: computer simulations of prediction games. *Episteme*, 6(2):200–220, 2009.

G. Schurz. Optimality justifications: new foundations for foundation-oriented epistemology. *Synthese*, 195(9):3877–3897, 2018.

G. Schurz. *Hume's Problem Solved: The Optimality of Meta-Induction*. 20xx. Manuscript in preparation.

T. F. Sterkenburg. *Universal Prediction: A Philosophical Investigation*. PhD Dissertation, University of Groningen, 2018.

T. F. Sterkenburg. The meta-inductive justification of induction. To appear in *Episteme*, 20xx.

V. G. Vovk. Competitive on-line statistics. *International Statistical Review*, 69(2):213–248, 2001.

Munich Center for Mathematical Philosohy, LMU Munich
E-mail address: tom.sterkenburg@lmu.de