This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

7

The View from Social Epistemology

We ended the previous chapter by noting that there are distinctively social aspects to learning. Not only do we learn *from* others, as when we learn from our parents or our teachers, but we often can acquire new knowledge only if we join forces *with* others. This is true not only for everyday knowledge but also and perhaps in particular for scientific knowledge. Indeed, scientists who might want to work in relative isolation—in the manner of Newton or, more extremely, of the exceedingly reclusive Henry Cavendish—would today stand little chance of making any appreciable progress; modern science has become a community effort (Gribbin, 2002, p. 359). While this has been long recognized especially in the works of sociologists of science (e.g., Kuhn, 1962; Merton, 1973), the more general lesson that the communal aspects of knowledge acquisition call for a rethinking of *epistemology* sank in only after the publication of Goldman's seminal 1999 book on social epistemology.

Once it is appreciated to what extent that we, qua epistemic subjects, are dependent on our functioning as members of groups of agents bent on the same epistemic goal, it becomes natural to ask how we might best organize such groups, supposing that we want to optimize their performance. This is a broad question, of course. Even if we consider a well-delineated case, such as a group of academic researchers employed by the same laboratory, the question could concern topics ranging from strategies for organizing the workplace to after-work socializing or decisions about how to allocate grant money and other resources within a group (Kitcher, 1993). This chapter focuses instead on some of the most fundamental *epistemic* mechanisms that groups might decide to put into place or might be encouraged to adopt, for how their members ought to interact. In particular, it focuses on mechanisms for how group members should update on new evidence while also being

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

informed about the changing belief states of others in the group, whom they know to acquire evidence flowing from a common source (all group members study, for example, the same economic phenomenon, or the same virus, or the same genetic mutation). To nobody's surprise, we are especially interested in comparing Bayesian and abductive versions of social learning.

The notion of *optimizing* group performance will also be understood in an epistemic sense, as relating to the question of how to get at the *truth* of whatever issue the group is working on. Whereas, as stated in chapter4, truth is generally regarded as the overarching epistemic goal, we are often forced to make a trade-off between speed and accuracy: we do want to get at the truth, but we also want to get there reasonably fast, which may mean that we should be content if we can quickly become *highly confident* in the truth or quickly get *close* to the truth.

Thus, the question to be investigated is how members of a group of agents should update on the receipt of new evidence in a setting in which they also have access to relevant beliefs of other group members, supposing that the group wants to strike the best balance between speed (getting at the truth fast) and accuracy (minimizing error rates). The main methodological tool to be used is again that of computational agent-based modeling, but now we also want to explicitly take into account the socio-epistemic interactions. To that end, we build on a model developed by Rainer Hegselmann and Ulrich Krause for studying opinion dynamics in groups of interacting agents focused on a common research question (Hegselmann & Krause, 2002, 2005, 2006, 2009; for related models, see Deffuant et al., 2000; Dittmer, 2001; Weisbuch et al., 2002; Pluchino, Latora, & Rapisarda, 2006; Semeshenko, Gordon, & Nadal, 2008; De Langhe & Greiff, 2010). In this Hegselmann–Krause (HK) model, as it is now commonly known, each agent holds at any point in time only a single opinion on an issue that is the same throughout the process. Later work has generalized the model to those in which agents could at each point in time have opinions on multiple issues (e.g., Lorenz, 2003, 2008; Riegler & Douven, 2009; Wenmackers, Vanpoucke, & Douven, 2012, 2014). In recent joint work with Sylvia Wenmackers (Douven & Wenmackers, 2017), I have presented an extension of the HK model in which each agent at each moment assigns *probabilities* to all members of a set of relevant possibilities.

Douven and Wenmackers's (2017) aim was to compare different update rules within a social setting. Specifically, they compared Bayes's rule with various instances of EXPL. One way in which I want to go beyond Douven and

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

Wenmackers's work is by considering, next to EXPL, the other formalizations of explanatory reasoning that we examined in the previous chapter. Another generalization concerns the way that an agent determines which of her group members she takes to be her *epistemic peers*, who in Douven and Wenmackers (2017) are those members whose probabilities she takes into account in computing her new probabilities.[1] Whereas Douven and Wenmackers used one particular metric for defining epistemic peerhood, in this chapter I want to consider others as well. Furthermore, I compare different ways in which group members may accommodate the information about their peers' probabilities, thereby going beyond previous research, Douven and Wenmackers having considered only the linear pooling of probabilities (Stone, 1961; for details, see the following).

Whereas the main conclusion from Douven and Wenmackers (2017) was that in a social setting explanatory reasoning can be superior to Bayesian updating both in terms of speed and in terms of accuracy, the present chapter aims at showing how we may be able to find an optimal combination of update rule, rule for determining peerhood, and rule for taking into account one's peers' probabilities, where optimality is still a matter of balancing speed and accuracy. I am again using evolutionary computing by letting combinations of the aforementioned kind compete for survival and reproduction, where in a first batch of simulations their success depends directly on how fast and accurate they are, and in a second batch, more aligned with the simulations reported in the previous chapter, on survival probabilities and so more indirectly on speed and accuracy.

While my main motivation continues to be a comparison of Bayesian and abductive updating, this chapter has a secondary motivation as well. In recent years, much has been made of the so-called wisdom of crowds effect, the often-reproduced finding that averaged group estimates of some parameter value tend to be closer to the truth not only than most of the individual estimates but even than expert estimates (Surowiecki, 2004; Page, 2007). At the same time, Jan Lorenz et al. (2011) report evidence showing that if group members

1. The intuitive idea behind distinguishing peers from non-peers is that the former are and the latter are not considered competent enough to make it worth taking their opinions into account. For background on the notion of epistemic peer, see, for instance, Christensen (2007, 2011, 2016), Douven (2010), Goldman (2010), Ballantyne and Coffman (2011, 2012), De Cruz and De Smedt (2013), De Langhe (2013), Vallinder and Olsson (2013), and De Cruz (2017).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

are made aware of their fellow members' estimates, that will not only lead members to adjust their own estimates, but the various resulting adjustments also will tend to have a negative effect on the accuracy of the averaged group estimate. This might seem to give a reason for trying to *prevent* the kind of interactions among group members that were experimentally explored by Lorenz and coauthors. In a similar vein, Kevin Zollman (2007) found that while increasing the connectivity in an epistemic network tends to increase the speed with which the network converges on an opinion, it at the same time tends to *decrease* the chances that the network will arrive at the truth (see also Kummerfeld & Zollman, 2016, and Hahn, Hansen, & Olsson, 2020; but cf. Mason & Watts, 2012, and Rosenstock, Bruner, & O'Connor, 2017).

However, while these authors found a potential negative effect of epistemic interactions among the members of research groups, it is to be noticed that such interactions are not *bound* to be detrimental. To the contrary, as Douven and Wenmackers's (2017) results made manifest, whether the influence will have a positive or a negative effect on the accuracy of the group estimate may all depend on *how* the individual group members' beliefs are influenced by those of their peers. There are update procedures that take into account both new information about the parameter value at issue *and* information regarding the beliefs about that value held by others in the group, and which greatly increase not only the accuracy of the group estimate but also the speed with which that estimate converges to the true value, supposing that group members are repeatedly offered new evidence as well as new information about the beliefs of their peers. Groups might decide to adopt those update procedures to counteract the kind of effects of social influence discovered by Lorenz et al. (2011) and others. Naturally, that presupposes that such procedures have been identified. In the following, I describe in some detail how this could be accomplished.

7.1 The Hegselmann–Krause Model and Beyond

We look at groups of agents, each of whom repeatedly updates her belief state upon the receipt of new information. The new information is of two kinds: on one hand, agents receive new information directly about the subject matter of their research, while on the other hand, they receive new information about the belief states of others in their group after *they* received new "direct" information about that subject matter. All agents always first receive a piece of

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

direct information—which may be different for different agents, even though the information comes from the same source—and then after each of them has updated on that, they receive information about the new belief states of their colleagues, which may occasion a further adjustment of an agent's own belief state. For the first part of this process, I resort to the same update rules that figured in the simulations reported in the previous chapter: Bayes's rule and instances of EXPL, Good's rule, and Popper's rule. The second part involves a variant of the previously mentioned HK model.

Computational agent-based modeling has become a mainstay in the social sciences, used to study a great variety of phenomena and processes—the emergence of social norms, racial segregation, voting behavior, and so on—that are too complex to be investigated using strictly analytical tools. The technique is also rapidly gaining popularity in the study of social aspects of knowledge and belief, fueled by the observation hinted at previously, to wit, that our being socially embedded in one or more communities of truth seekers is crucial to our epistemic success.[2]

Arguably, the most widely used computational agent-based model in this field is the HK model. Hegselmann and Krause themselves used the model mainly to address *descriptive* issues, such as the conditions under which initially diverging opinions in a community of interacting agents tend to converge or when they tend to polarize. More recently, the model has also been applied to a variety of *normative* issues, such as questions concerning reliabilism and truthful assertion (Olsson, 2008), peer disagreement (De Langhe, 2013), and efficient truth approximation in science (Douven & Kelp, 2011).

In the version of the HK model that is most of interest to us (though not in the original, most basic version of the model), communities of agents seek to determine the value τ of some unspecified parameter, where it is given that $\tau \in (0, 1)$. The agents update their estimates of τ repeatedly at discrete time steps, where each agent updates on the basis of (1) information that she receives about τ, and (2) the estimates of τ of those agents who are within her bounded confidence interval (BCI), meaning that their estimate is within some fixed distance ε of the agent's own estimate. More formally, the estimate

2. For some more background on computational agent-based models, see Douven (2019c).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

of agent x_i after the $(u + 1)$-st update is given by

$$x_i(u + 1) \;=\; \alpha \frac{1}{|X_i(u)|} \sum_{j \in X_i(u)} x_j(u) \;+\; (1 - \alpha)\tau,$$

with $x_j(u)$ the estimate of agent x_j after the uth update,

$$X_i(u) \;:=\; \{j : |x_i(u) - x_j(u)| \leqslant \varepsilon\}$$

the set of agents within agent x_i's BCI after the uth update, and $\alpha \in [0, 1]$ a parameter determining the weight that the agent gives to the "social" part of the updating relative to the "evidential" part.[3]

One reason why the HK model is so popular is its great flexibility. In general, it takes little effort to extend the model or to adapt it to a researcher's specific needs. For example, researchers have used simple extensions of the model to study situations in which agents receive evidence beset by some degree of measurement error (e.g., Douven, 2010; Douven & Riegler, 2010; De Langhe, 2013). Matthew Crosscombe and Jonathan Lawry (2016) present another extension to study agents with (partially) vague beliefs. The model has also been extended to situations in which agents are equipped with richer belief states in that they hold opinions on various related or unrelated matters (see, e.g., Lorenz, 2003, 2008; Riegler & Douven, 2009; Wenmackers, Vanpoucke, & Douven, 2012, 2014) and to ones admitting different types of agents, some of which may exhibit epistemically irresponsible behavior (Hegselmann & Krause, 2015; Douven & Hegselmann, 2021; Glass & Glass, 2021).

We are interested mostly in an extension of the HK model proposed in Douven and Wenmackers (2017). In this extension, agents' belief states at a given time are not single real numbers (as in the original model) but probability functions, defined on a set of self-consistent, mutually exclusive, and jointly exhaustive hypotheses $\Psi = \{\psi_i\}_{1 \leqslant i \leqslant n}$. It is still the case, however, that the agents take into account both evidence directly pertaining to the members of Ψ and the belief states (i.e., probability functions) of those in their BCI.

3. In this model, and also in Douven and Wenmackers's extension of it, subsequently discussed, all agents in the BCI are treated on a par. If information about the credibility or expertise of these agents were available, one might plausibly wish to weigh them differently (see, e.g., Tamargo et al., 2014, and Gottifredi et al., 2018). This can easily be accomplished in the framework used in this chapter (see footnote 4 in this chapter).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

Naturally, in this model the BCI has a definition different from the previously stated one. Here, two agents are in each other's BCI precisely if the sum of the absolute differences in the probabilities that they assign to the hypotheses is below a given threshold value, or formally, agent i with belief state Pr_i is within the BCI of agent j with belief state Pr_j precisely if $\Delta(\mathrm{Pr}_i, \mathrm{Pr}_j) \leqslant \varepsilon$, where

$$\Delta(\mathrm{Pr}_i, \mathrm{Pr}_j) \ := \ \sum_{k=1}^{n} \left|\mathrm{Pr}_i(\psi_k) - \mathrm{Pr}_j(\psi_k)\right|.$$

Note that, given this definition, an agent is always in her own BCI (as is also the case in the original HK model and in the extensions mentioned previously).

The actual updating in Douven and Wenmackers's model then proceeds as follows: In a first, "evidential," step, agent j updates her probabilities on a new piece of evidence she receives, using either Bayes's rule or EXPL; that is, where Pr_j^u is her probability function after the uth update, she first forms $\mathrm{Pr}_j^{u_{\mathrm{evid}}}$ via either Bayes's rule or EXPL. Then, in a second, "averaging," step, the agent averages over the probability functions of the agents in her resulting BCI, $X\left(\mathrm{Pr}_j^{u_{\mathrm{evid}}}\right)$, which yields $\mathrm{Pr}_j^{u_{\mathrm{av}}}$. Finally, she sets $\mathrm{Pr}_j^{u+1} = \mathrm{Pr}_j^{u_{\mathrm{av}}}$.

The averaging takes place by a method now known as "linear pooling" (Stone, 1961), discussion of which goes back at least to Pierre-Simon Laplace's *Essai Philosophique sur la Probabilité* from 1814 (Bacharach, 1979). In this method, agents average by taking the *straight arithmetic average* of the probabilities of the agents within their BCI. Formally, for all $\psi \in \Psi$ and all $\mathrm{Pr}_k^{u_{\mathrm{evid}}} \in X\left(\mathrm{Pr}_j^{u_{\mathrm{evid}}}\right)$, it holds that

$$\mathrm{Pr}_j^{u_{\mathrm{av}}}(\psi) \quad = \quad \frac{\sum_k \mathrm{Pr}_k^{u_{\mathrm{evid}}}(\psi)}{\left|X\left(\mathrm{Pr}_j^{u_{\mathrm{evid}}}\right)\right|}.$$

The resulting function again satisfies the axioms of probability theory (Stone, 1961).[4]

Douven and Wenmackers's (2017) explicit aim was to compare Bayes's rule with EXPL in a social setting. Their starting point was the finding from Douven (2013), discussed in chapter4, that in an individual setting (without

4. There exist weighted versions of this and the other pooling methods subsequently discussed (see, e.g., Genest & Zidek, 1986), which could be used in the kind of situation in which some agents are known to be more credible or more expert than others (see the previous footnote).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

any averaging of probability functions taking place) EXPL tends to outperform Bayes's rule in terms of speed, but Bayes's rule tends to outperform EXPL in terms of accuracy. One could conclude from this that it is a matter of context which of these rules we ought to use (supposing that the choice is between only those two rules), given that typically we care about both speed and accuracy but sometimes more about one and sometimes more about the other.

But Douven and Wenmackers wondered whether a speed–accuracy trade-off would be needed if agents were to change their probabilities not only on the basis of direct evidence but also as a result of social influence. To address this question, they ran computer simulations using the extension of the HK model just described, with groups of simultaneously and repeatedly updating agents who received evidence that although from the same source (i.e., the same probability distribution) for all agents, was randomly generated per agent. Douven and Wenmackers systematically studied different communities of agents, which differed along two dimensions, to wit, by the value for c that they substituted in EXPL (with agents setting c equal to 0 *de facto* amounting to Bayesians), and by the value of ε, that is, by their being more or less liberal in considering others as their epistemic peers.

The agents in Douven and Wenmackers's simulations all receive binomial data (i.e., independent and identically distributed binary data), the data points coming in one at a time, and their goal is to estimate on that basis the probability of "success." For the sake of concreteness, one may imagine the data to consist of the outcomes of repeated tosses of a coin with unknown bias, where this bias (for example, the probability of heads) can take eleven possible values, ranging from 0 (certain to land tails) to 1 (certain to land heads) in increments of .1. Initially, all these values are equally likely to indicate the true bias of the coin. Thus, at the start of the updating process, it is reasonable for the agents to invest equal confidence in all given bias possibilities. Importantly, while each agent in the community receives data from a different coin, it is known to every agent that all coins have the *same* bias. The community of agents is tasked to estimate this bias.

As said, the relevant dimensions of evaluation were speed and accuracy. The former was operationalized in terms of the percentage of the community that assigns a high probability to the truth (where "high probability" was understood as a probability above .9), and the latter in terms of the Brier score, which we already encountered a number of times.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

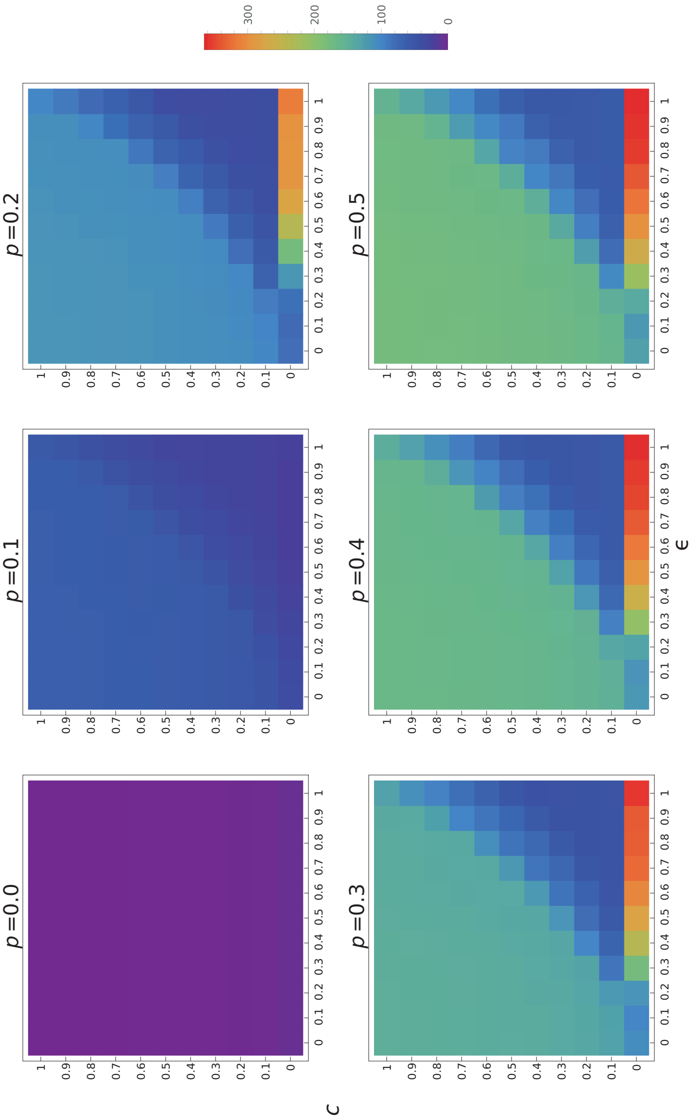

Figure 7.1: Average total Brier scores for six bias possibilities; $p = x$ indicates that the chance for each of the coins to come up heads equals x. (*Source*: Douven & Wenmackers, 2017.)

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

Douven and Wenmackers's simulations can be thought of as effectively comparing 121 (= 11 × 11) update procedures with each other, one for each combination of c and ε, with both in $\{0, 0.1, 0.2, \ldots, 1.0\}$. The communities that these authors considered all consisted of fifty agents, each of whom received per simulation the outcomes from 500 tosses. Given that as previous research had shown, results might depend on the specific bias of the coin, simulations were run for biases $p \in \{0, .1, .2, .3, .4, .5\}$.[5] For each of these bias possibilities, fifty simulations were run.

In all simulations, average Brier scores per agent were calculated after each update. Per simulation, these averages were then summed over all 500 tosses, and finally the average over the 50 simulations was calculated. The summarized results are shown in figure 7.1, which is reproduced from Douven and Wenmackers (2017). The bottom row of each of the six plots represents the communities of Bayesian updaters, which differ only by their willingness to count others as their epistemic peers; the square in the bottom-left corner of each plot represents communities of "pure" Bayesians, who attach no value to explanatory strength and are unwilling to be influenced by anyone else in their group. As Douven and Wenmackers observe, among the Bayesian communities the pure ones do best on the count of Brier scores: as clearly visible, Bayesians who *are* willing to be influenced by others incur higher Brier penalties on average. More significantly, *non*-Bayesian updaters who are willing to be influenced by others do better still than pure Bayesians.

With regard to speed, it had already been found that on average, and in the statistical model considered in the simulations, EXPL leads agents to assign a high probability to the truth faster than does Bayes's rule. This result was reproduced in the kind of social setting assumed by Douven and Wenmackers. Figure 7.2 shows graphically for different combinations of c and ε the average proportions of agents (averaged over the simulations) that assign a high probability to the truth, for the relevant bias possibilities, and after each toss (compare figure 6 in Douven & Wenmackers, 2017).

Douven and Wenmackers (2017) were interested in comparing Bayes's rule and EXPL in a social setting along the dimensions of speed and accuracy. They found that although in an individual setting the two rules appeared to have different strengths and weaknesses—Bayes's rule was more accurate

5. For symmetry reasons (see ch. 4, footnote 11), it would have been superfluous to run simulations for the other bias possibilities.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

than EXPL, but EXPL was faster than Bayes's rule—this picture changed in a social setting, where EXPL appeared to be both faster and more accurate than Bayes's rule.[6]

It was known that in the kind of individual setting studied in Douven (2013), EXPL users typically had a lower Brier penalty than Bayesians but still tended to have a slightly higher *average* Brier score, because *when* they incurred a higher Brier score than Bayesians it tended to be *much* higher. The cause was that EXPL users tend to pick up the lead in the evidence faster than Bayesians do, which ultimately has the downside that they are more prone to follow *misleading* evidence. As Douven and Wenmackers (2017) showed, in a social setting the risk of being pushed off track by misleading evidence is much diminished due to the averaging mechanism that is called at each step. Specifically, if an agent receives misleading evidence, then the averaging part still brings about a kind of regression to the mean, thereby preventing the agent from straying too far from the true bias probability.[7]

6. Trpin and Pellert (2019) show that these results generalize to cases in which the evidence is uncertain. They effectively compare Jeffrey conditionalization (see section 3.2.1) with an explanation-based counterpart of that rule.

7. Bayesians might object that instead of averaging over the probability functions of their peers, agents should reason back from how those functions are updated to the evidence that the others received and then update their own probability function on the pooled evidence (this would amount to a supra Bayesian approach, in the sense of Keeney & Raiffa, 1976). As pointed out in Douven and Wenmackers (2017, section 5), however, this objection misses the purpose of the statistical toy model used in their paper and also here, which is meant to set up a simulation and not as the ultimate case of interest. In particular, we have no special interest in cases that are analytically tractable, in which all the evidence is at least in principle available to all the agents, and in which the only epistemic goal is to minimize expected Brier penalty. We are interested in learning rules that apply also to situations in which the model is known to be at best an idealization, in which, for instance, information regarding other agents' evidence is or may be incomplete and mediated by probability assignments, in which we have to estimate on how much independent data the assignment of each agent is based, and so on. More practically speaking, and also entirely in line with the view on rationality taken in the previous chapter, it is worth stressing that even in a context that fits the toy model perfectly, processing constraints and in particular working memory constraints may make the "reasoning back" strategy infeasible. This is most evidently true when we assume the updating to take place under time pressure, as is for instance eminently plausible in the kind of situation considered in section 7.4. Bayesians might respond that in such a situation, averaging over others' probability functions is equally problematic. That is not so, however. It has been long known that people can accurately estimate averages without carrying out the actual computations (Spencer, 1961, 1963; Beach & Swenson, 1966; Bulger, Hiles, & Lowe, 1969; Lovie & Lovie, 1976; Malmi & Samson, 1983; Lindskog, Winman, & Juslin, 2013a, 2013b).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

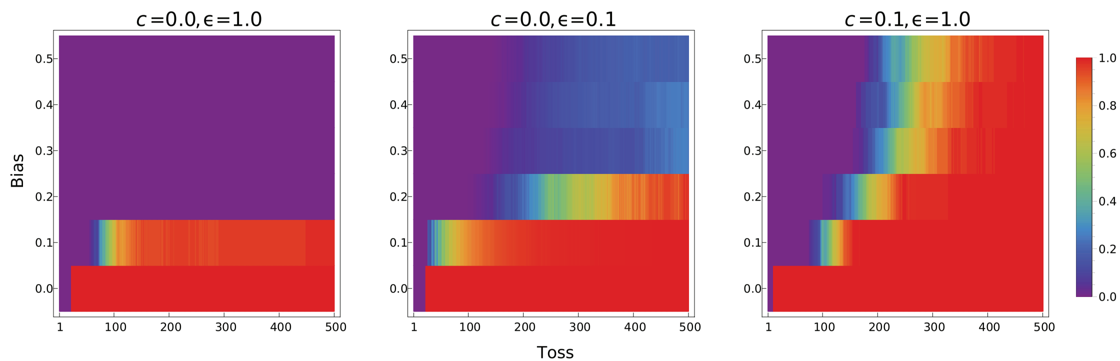

Figure 7.2: Proportions per toss and per bias of agents who assign a high probability to the truth, for specific combinations of c and ε; results are averaged over all simulations.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

7.2 Further Generalizing Group Learning

We want to compare update rules in the context of determining how to epistemically organize a research group in order for that group to strike the best balance between being fast and being accurate. Part of the recommendation to a group might be to adopt Bayes's rule, or EXPL, as their members' way of accommodating direct evidence. First, as discussed in section 6.3, there exist other formalizations of abductive reasoning than EXPL, and these should certainly also be considered in a comparison. Second, Douven and Wenmackers (2017) let the agents in their simulations average via linear pooling, but one can think of other ways of merging probability functions, some of which might offer advantages in striking the best balance between speed and accuracy. Finally, we measured distances between agents' probability functions by adding up proposition-wise absolute differences between probabilities; this is the so-called Manhattan distance. Here, too, there are variants that could be tried and that if adopted by a group of agents to determine epistemic peerhood, might help that group to do better epistemically.

While the linear pooling of probability functions assumed by Douven and Wenmackers (2017) has been advocated by various authors (e.g., Aczél & Wagner, 1980; Lehrer & Wagner, 1981), others have expressed a preference for *geometric pooling* (e.g., Genest, 1984), in which one takes the geometric instead of the arithmetic average of probabilities (and then renormalizes, which becomes necessary).[8] Formally, the geometric average of numbers $x_1, x_2, \ldots, x_n$ is $\sqrt[n]{x_1 x_2 \cdots x_n}$. So, instead of setting for any update u and with $\Pr_k^{u_{\text{evid}}} \in X\left(\Pr_j^{u_{\text{evid}}}\right)$ and $\Psi = \left\{\psi_i\right\}_{1 \leqslant i \leqslant n}$ a set of jointly exhaustive and mutually exclusive self-consistent propositions,

$$\Pr_j^{u_{\text{av}}}(\psi_i) = \frac{\sum_k \Pr_k^{u_{\text{evid}}}(\psi_i)}{\left|X\left(\Pr_j^{u_{\text{evid}}}\right)\right|}, \qquad \text{for all } i: 1 \leqslant i \leqslant n,$$

8. In fairness, while there is strong empirical support for the thought that people can accurately estimate *arithmetic* averages (see footnote 7 in this chapter), I am not aware of any evidence indicating that they are good at estimating geometric or harmonic averages (the latter to be introduced shortly). So, what follows may be more of interest to, for example, roboticists studying opinion pooling (e.g., Lawry & Lee, 2020; Lee, 2020) than to cognitive scientists and social epistemologists (at least to those social epistemologists who share my conviction that we should occupy ourselves with developing theories of rationality for real people rather than elegant formal models that describe the behavior of agents bearing an at best vague semblance to ordinary mortals).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

the proposal is to set

$$\Pr_j^{u_{\mathrm{av}}}(\psi_i) \;=\; a\,\sqrt[m]{\prod\nolimits_k \Pr_k^{u_{\mathrm{evid}}}(\psi_i)}, \qquad \text{for all } i\colon 1 \leqslant i \leqslant n,$$

with $m = \left|X\left(\Pr_j^{u_{\mathrm{evid}}}\right)\right|$ and with $a = \left(\sum_{i=1}^{n} \sqrt[m]{\prod_k \Pr_k^{u_{\mathrm{evid}}}(\psi_i)}\right)^{-1}$, to ensure that the probabilities assigned to the members of Ψ sum to 1 again.

For numbers $x_1, x_2, \ldots, x_n$, their harmonic average equals

$$\left(\frac{x_1^{-1} + x_2^{-1} + \cdots + x_n^{-1}}{n}\right)^{-1}.$$

I am not aware of any recommendations in the literature to aggregate probability functions via the harmonic average—"harmonic pooling," one might call this—nor can I think of good a priori reasons for using this procedure. However, neither can I rule out a priori that for the purpose of making group learning fast and accurate, using the harmonic average has an edge over using either the arithmetic or the geometric average. Thus, I also consider the following definition of an average probability function:

$$\Pr_j^{u_{\mathrm{av}}}(\psi_i) \;=\; a\left(\frac{\sum_k \Pr_k^{u_{\mathrm{evid}}}(\psi_i)^{-1}}{\left|X\left(\Pr_j^{u_{\mathrm{evid}}}\right)\right|}\right)^{-1}, \qquad \text{for all } i\colon 1 \leqslant i \leqslant n,$$

with a ensuring that the probabilities sum to 1.

With regard to measuring distances between probability functions—in order to determine who is an epistemic peer and who is not—the Manhattan distance used by Douven and Wenmackers (2017) is but one possibility. There are many alternatives. Here, we focus on just two, but readers are invited to use the computer code in the Supplementary Materials to experiment with other distances (see the following for details). The two considered are the Euclidean metric, according to which

$$\Delta(\Pr_i, \Pr_j) \;:=\; \sum_{k=1}^{n} \left(\Pr_i(\psi_k) - \Pr_j(\psi_k)\right)^2,$$

and the Kullback–Leibler (KL) divergence, according to which

$$\Delta(\Pr_i, \Pr_j) \;:=\; \sum_{k=1}^{n} \Pr_i(\psi_k)\ln\left(\frac{\Pr_i(\psi_k)}{\Pr_j(\psi_k)}\right).$$

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

Note that, supposing the KL divergence, because this is not symmetric it will not generally hold that $\Delta(\Pr_i, \Pr_j)$ is equal to $\Delta(\Pr_j, \Pr_i)$. However, in the present context, the asymmetry may actually be an advantage, given that in reality, respecting someone as one's epistemic peer is not always reciprocated.

We can treat the extension of the HK model from Douven and Wenmackers (2017) as a template, into which one rule from each of three groups can be plugged in, viz., one of four evidential update rules: Bayes's rule, EXPL, Good's rule, and Popper's rule; one of three averaging rules: linear pooling, geometric pooling, and harmonic pooling; and one of three rules for determining epistemic peerhood: the Manhattan metric, the Euclidean metric, and the KL divergence. Each of the resulting combinations of rules yields a schema, with slots for c and ε still to be filled in. Does it matter which combination of rules we choose and what values we set for c and ε?

It will probably come as no surprise that the answer is positive. This is easy to verify by rerunning the simulations reported in Douven and Wenmackers (2017) for other instances of the template. By way of illustration, consider figure 7.3, in which the plots summarize results from simulations in which linear pooling was replaced by geometric pooling, but which were otherwise the same as the simulations underlying figure 7.1. We see that this seemingly small difference leads to a *completely* different picture. Now, Bayesians perform very well on the count of accuracy, supposing that they are open to social influence, although on the condition of a greater openness to such influence (even if only slightly greater), EXPL users who assign *some* (not-too-large) bonus for explanatory goodness do well, too. By way of further illustration, figure 7.4 shows that we obtain a somewhat different picture if harmonic pooling is assumed and we replace the Manhattan metric by the Euclidean one.[9]

But accuracy, here measured by the Brier score, is only one dimension of comparison, speed being the other one, and the goal ultimately is to strike the best balance between them. More importantly, we are considering only an illustration here, which by no means provides a systematic answer to our question of how best to organize a group epistemically (which update rule to recommend to the group and which principles to recommend for regulating social influence on belief states).

9. The color scaling was set separately for the plots in figures 7.1, 7.3, and 7.4. Using the same scale in the three figures would have rendered the differences in figures 7.3 and 7.4 difficult to discern.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

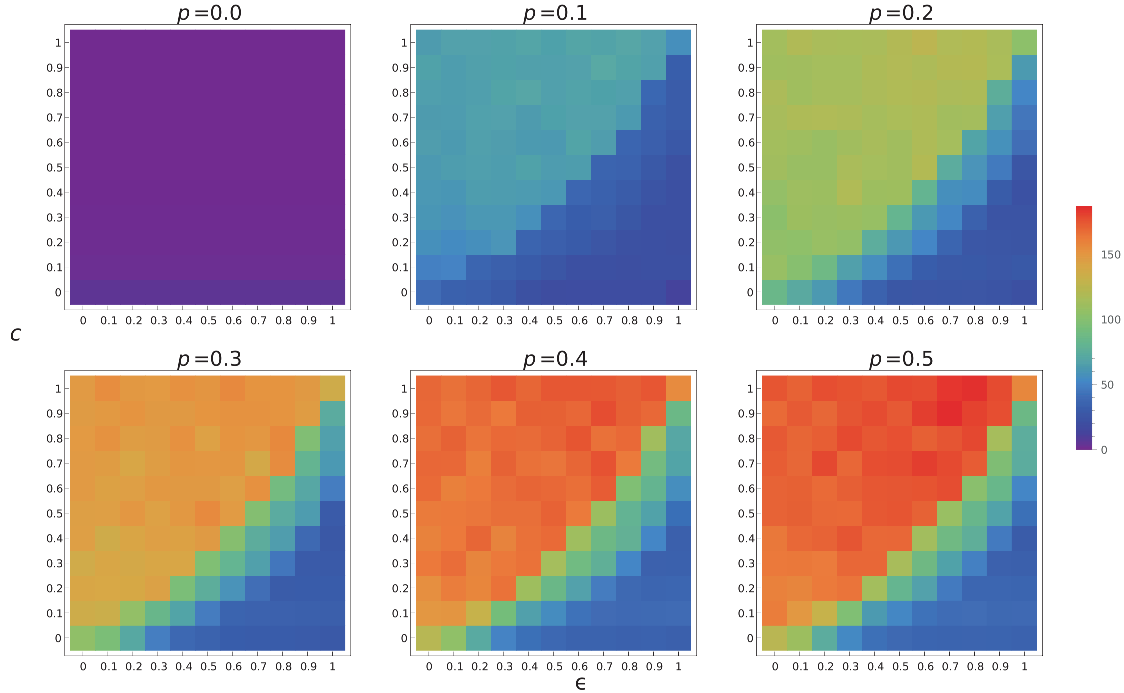

Figure 7.3: Average total Brier scores for six bias possibilities for the combination of EXPL, the Manhattan metric, and geometric pooling; $p = x$ indicates that the chance for each of the coins to come up heads equals x.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

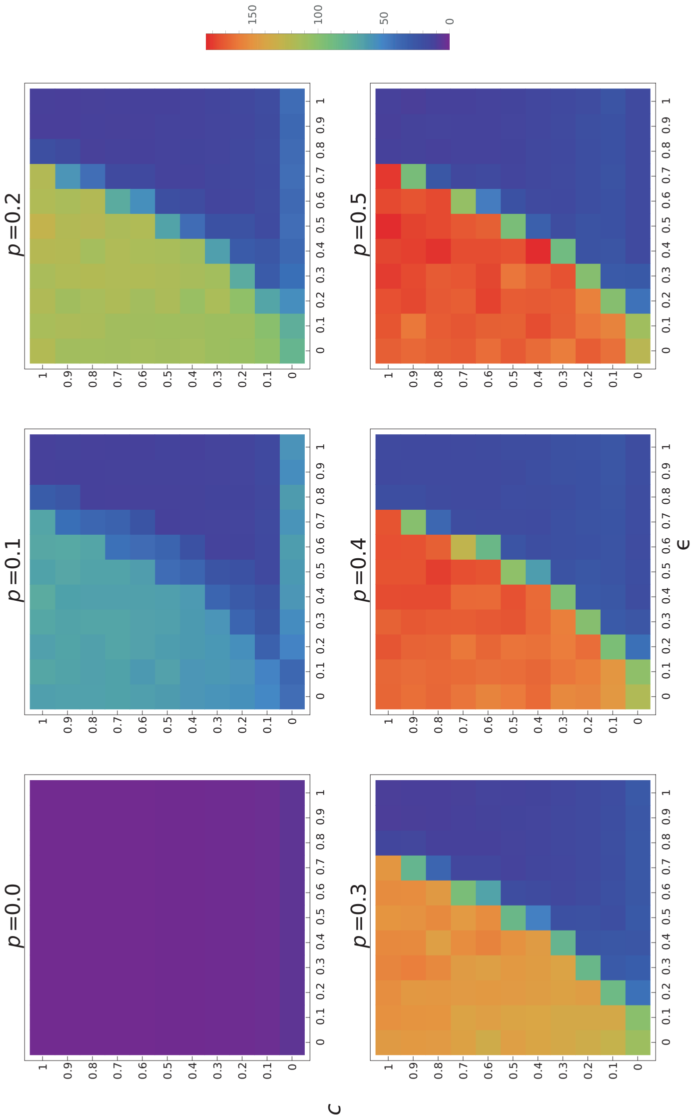

Figure 7.4: Average total Brier scores for six bias possibilities for the combination of EXPL, the Euclidean metric, and harmonic pooling; $p = x$ indicates that the chance for each of the coins to come up heads equals x.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

To forestall misunderstanding, it is to be noted that, in answering this question, we do not have to solve a multiobjective optimization problem, with speed and accuracy as objectives (Deb, 2001; Qian, Yu, & Zhou, 2013; Rakshit & Konar, 2015). As mentioned repeatedly, we are thinking of striking the best *balance* between speed and accuracy as our objective. In doing so, we assume that what counts as "best" depends to an *equal* degree on speed and accuracy. This will not always be the case, of course, but for now it serves as a useful simplifying assumption; we come back to this in the final section of this chapter. Although by thinking of the task before us as a single-objective optimization problem we are avoiding some complications, it remains the case that the search space is mixed, with both discrete (update rule, distance, and pooling method) and continuous variables (values for c and ε). This on its own makes analytical methods difficult to apply. A posteriori methods, specifically computational methods, have been applied with greater success to such problems. In the next section, we again make use of the method of evolutionary computing that was introduced in the previous chapter.

7.3 Evolutionary Computing and Optimal Group Learning

The optimization technique used in the following is again a form of agent-based optimization. Recall from chapter 6 that in this technique, different types of agents are thought of as representing different solutions to a given problem and that these solutions are scored with respect to the relevant objective or objectives. From those scores we derive the agents' fitness levels, on the basis of which, in a next step, we select agents either deterministically or stochastically in order to retain them or allow them to reproduce (or both) to form a new population, which then enters into a new competition for survival and reproduction. As previously mentioned, this process of competition, selection, and reproduction can be repeated several times until some predetermined criterion is met.

The evolutionary computing used is slightly more complicated than that used in chapter 6. Rather than individual agents, we consider *groups* of agents, where these groups differ by having adopted different epistemic principles of the kind examined in the previous sections. One could consider groups whose members use (partly) different epistemic principles in order to investigate whether such epistemic in-group diversity might carry evolutionary advan-

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

tages, but the groups that I looked at consisted of agents who all assumed the same set of epistemic principles.

Sometimes in evolutionary computing, the predetermined stopping criterion is a minimally acceptable fitness level to be attained by some agent (or more generally, by some solution to the target problem). For our present concerns, there does not appear to be such a minimal fitness level that could be set a priori. We would *like* groups to get to the truth immediately and unexceptionally, but that is an unrealistic goal to attempt. What *is* a realistic goal is not clear from the outset but rather something that one would hope to learn from the simulations. So I decided to use the most straightforward stopping criterion to be found in the evolutionary computing literature, to wit, stopping after a fixed number of generations. However, the number I chose—fifty—was established on the basis of some pilot runs that suggested that this number of generations would suffice to achieve satisfactory convergence, in that solutions in the fiftieth generation would deviate only slightly from each other. (This was confirmed by the actual runs carried out, as subsequently described.)

Although all evolutionary algorithms build on the same ideas of variation and selection, they can differ greatly in their details, most notably concerning the extent to which they rely on the basic evolutionary operations of crossover, mutation, and selection as well as the extent to which such operations are applied stochastically or deterministically. In the previous chapter there was no mutation involved, and there was no crossover involved either, in the sense that there was no "mixing" of "genetic" material. In place of crossover, there was a simple copying operation: the 50 percent best doctors were selected in each round and allowed to move on to the next generation, together with an identical copy of them. The algorithm I used for the simulations reported in this section involved selection and crossover, detailed subsequently, and to a lesser extent, mutation. In this algorithm, selection proceeded deterministically but crossover and mutation proceeded stochastically. Selection was done on the basis of speed and accuracy, the latter operationalized in terms of the average Brier score incurred by a group and the former in terms of the first time a majority of the group assigned a high probability (understood as a

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

probability above .9) to the true hypothesis.[10] Both criteria factored equally in determining fitness values.

The simulations that I ran were similar to those reported in Douven and Wenmackers in that the agents in any one group received binomial data from a common source, where the data one agent received were independent from those received by any other agent. In this situation each agent repeatedly flips her own coin even though all coins have the same (initially unknown) bias, and the shared task of the group is to find out what that bias is. For concreteness, I describe the data in the following as the outcomes of coin flips.

More exactly, the algorithm used consists of the following five steps:

Step 1. Create a starting population of groups each consisting of fifty agents, such that all combinations of update rules (Bayes's rule, EXPL, Good's rule, and Popper's rule), distances (Euclidean, Manhattan, and KL divergence), and pooling methods (linear pooling, geometric pooling, and harmonic pooling) are represented by exactly one group, yielding a total of thirty-six groups; for the groups consisting of non-Bayesian updaters (users of EXPL, Good's rule, or Popper's rule), choose randomly per group an explanatory bonus value from the standard uniform distribution; and for each group, choose randomly a value for ε (i.e., the "peer radius") also from the standard uniform distribution.

Step 2. For every agent in every group, generate a sequence of 500 coin flips, where all coins have the same randomly chosen bias (the choices limited in the same exact way as in Douven & Wenmackers, 2017, so that all possible biases are multiples of 0.1). Let agents update their probabilities after each toss according to whichever combination of rules their group has adopted, using the generalized HK model described previously, and after each update, calculate Brier scores for all agents.[11] Also, register after each update the percentage of agents

10. I also experimented with different scoring rules and different threshold values. These experiments did not lead to any noteworthy qualitative differences. Comparing forms of abductive reasoning with Bayesian reasoning using different stopping rules (see footnote 10 in ch. 4) is left as an avenue for future research.

11. To forestall misunderstanding, it is to be noted that the agents in the model all belong to the same group, so that the algorithm enables only intragroup communication and not intergroup communication.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

who assign a probability greater than .9 to whichever the true bias hypothesis is in this run. For each group, repeat this scoring process twenty-five times; so choose anew a bias for the coins, let the agents update on 500 coin flips, and score each agent with respect to accuracy (Brier score) and speed (probability greater than .9 assigned to the truth), and keep doing this until all agents have gone through twenty-five such rounds of scoring. Then calculate the group's average Brier score (i.e., the total Brier score incurred by the group's members averaged over the twenty-five runs) and also the average toss number after which a majority of agents in the group assign a probability greater than .9 to the truth (the average being over the twenty-five runs).

Step 3. Rank the thirty-six groups in the population twice, once according to their average Brier score and once according to their average speed (expressed in terms of steps needed to achieve truthcloseness), with lower ranks being better in both cases. Then add up the ranks and select the eighteen groups with the lowest resulting sums (following, e.g., Baker, 1985). Draw (with replacement) from these "fittest" groups eighteen pairs, and then select randomly and independently from each pair an update rule, a distance, and a pooling method; draw a value for the explanation bonus c from a normal distribution truncated between 0 and 1 inclusive with as mean the mean of the c values of the members of the pair and as standard deviation the standard deviation of the c values of the whole previous generation; and determine the value for ε in exactly the same way. Put the results of these selections together to obtain eighteen new "offspring" groups.

Step 4. Form the next generation of groups by conjoining the eighteen best groups from the parent generation and their offspring (another eighteen groups) formed as explained in step 3.

Step 5. Iterate steps 2–4.

Figure 7.5 shows a schematic representation of the given steps.

I have two comments on this algorithm. First, as mentioned, selection occurs deterministically while variation occurs stochastically. Given the way the values of c and ε are picked for the offspring, an element of random

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

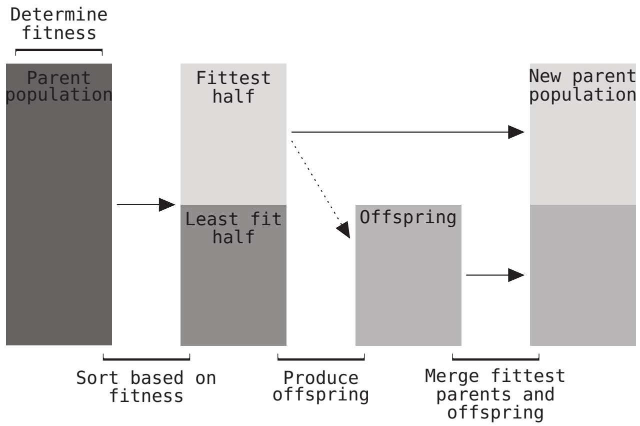

Figure 7.5: Schematic representation of how the evolutionary algorithm creates a new generation of groups. (Solid arrows suggest that groups move from one stage to the next, and the dotted arrow suggests that although the groups themselves do not move to a next stage, their genetic material does.)

mutation is built into the variation part. In the piloting stage of the present research, I experimented with stochastic selection as well, in particular with an algorithm in which a group's rank on the basis of the scoring determined its chances of being selected for producing offspring. Because this did not lead to results qualitatively different from those obtained with deterministic selection but did slow down the computations quite a bit, I went for deterministic selection in the full simulations.

Second, these simulations, too, were conducted using the Julia language. Appendix E explains how to obtain and use the code for the simulations. Readers are encouraged to go through the code in the notebook and to experiment with it; doing so should further enhance understanding of the simulations. The code will also allow readers to try out variant update rules, distances, and aggregation functions. For example, the distances that I used are from the `Distances.jl` package for Julia, which currently supports a total of twenty-seven distances, a number of which are certainly plausible options for determining peerhood (e.g., the Jensen–Shannon divergence and the Hellinger distance).

As said, one run consisted of an evolutionary sequence of fifty generations in which except for the starting generation, each generation was the result

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

of the process described in steps 2–4. The plan was to start by conducting fifteen runs and to continue with additional runs depending on how much variance was found in the first fifteen.

Turning now to an evaluation of the results that were obtained, one can get a first impression of them from Figures 7.6 and 7.7. These figures show the outcomes from a single, randomly chosen run. Figure 7.6 consists of three series of fifty stacked histograms, each histogram showing the percentage of tokens of each update rule type (top plot), distance type (middle plot), and pooling type (bottom plot) present in the corresponding generation. It is clear from this figure that after nineteen generations, all groups agree on a best combination of update rule, distance, and pooling method, to wit, EXPL, the Euclidean metric, and harmonic pooling. Figure 7.7 shows that there is great agreement among the groups about the best values for c and ε, which are both in the higher range. Figure 7.8 shows that how the population of groups evolves has a noticeable impact on how accurate they are on average, as well as on how fast they are on average.

Whereas the aforementioned figures show the results of a randomly selected run, very similar results were obtained in the other runs. Notably, there appeared to be a strong convergence not only in that all runs showed all groups to endorse, after twenty to thirty generations, a combination of EXPL, the Euclidean metric, and the harmonic pooling method, but also in that in each run all groups came to hold values for both c and ε typically in the vicinity of .9. More specifically, pooling the results from all runs, I found, averaging over all last-generation groups, a c value of .94 (± 0.04), and an ε value of .91 (± 0.05). These values hardly differ from the average values of the top groups in the last generations from the fifteen runs: .95 (± 0.03) for c, and .89 (± 0.07) for ε. In view of these results, I refrained from conducting runs in addition to the fifteen with which I started.

To see the effect that the evolutionary process had on the criteria of interest, note that the average total Brier score incurred by the groups in the first generation was 6,468.19 (± 730.07) whereas that incurred by the groups in the last generation was 616.08 (± 114.59). This amounts to an average Brier score *per agent* of 129.36 for first-generation groups, and 12.32 for last-generation groups. Similarly, the first toss number at which a majority of agents in a group assigned a high probability to the truth was, on average, 153.60 (± 6.20) for groups in the first generation, and 31.47 (± 1.16) for groups in the last generation.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

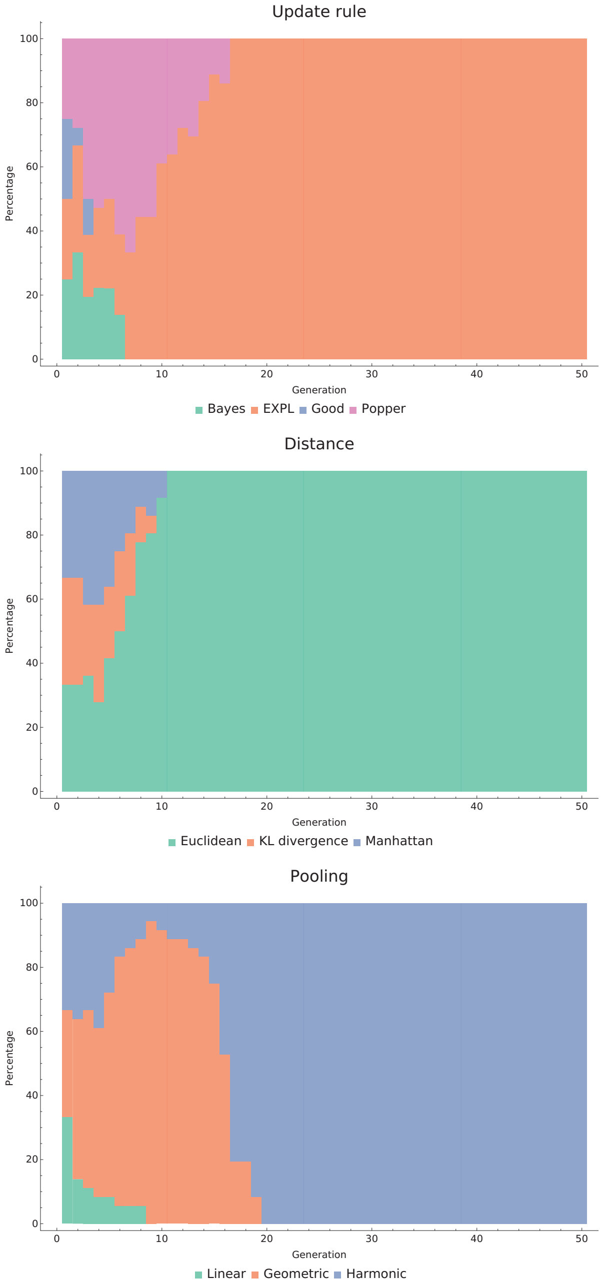

Figure 7.6: For a randomly chosen run, per-generation percentages of rule types (*top*), distance types (*middle*), and pooling types (*bottom*).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

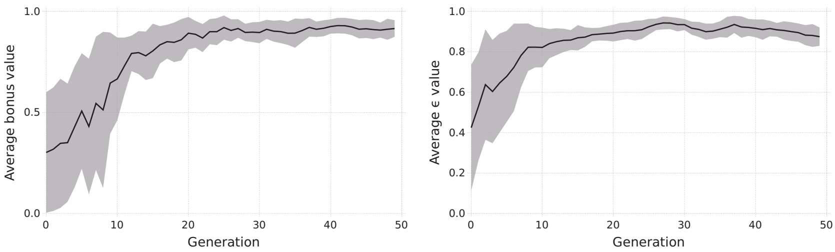

Figure 7.7: Average bonus and ε values for the fifty generations in the same run that generated the data shown in figure 7.6; shaded bands indicate one standard deviation from the mean.

To bring the significance of the overall outcome into further relief, one can run the fitness-level determining procedure used in the evolutionary computations on a group of *noninteracting* Bayesians (agents updating via Bayes's rule but not letting themselves be influenced by the opinions of their peers) and, for comparison, on a group of agents that has adopted the setting that came out best in the foregoing, so a combination of EXPL as an update rule, with $c = 0.95$; the Euclidean metric for determining peerhood, with $\varepsilon = .89$ as determining the criterion for peerhood; and the harmonic pooling function for aggregating probability functions. For either group, (1) draw a random

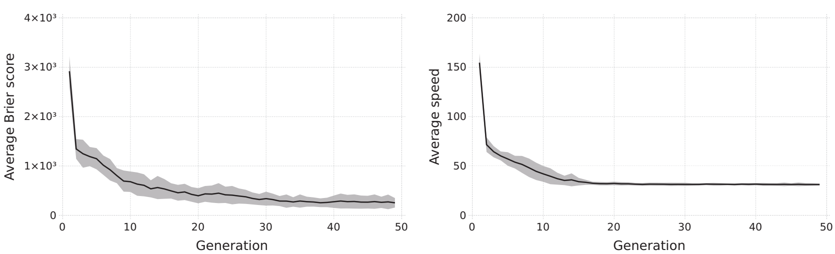

Figure 7.8: Average total Brier scores and average first toss after which the majority assigns a high probability to the truth, for the fifty generations in the same run that generated the data shown in figures 7.6 and 7.7; shaded bands indicate one standard deviation from the mean.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

bias hypothesis; (2) generate for each agent her own sequence of 500 tosses with the drawn bias; (3) let each agent consecutively update on those tosses and, for those in the second group, also take into account her peers' opinions in the way prescribed by the said epistemic principles; (4) register the Brier scores for each agent after each update; (5) register the first toss after which a majority of the group assigns a probability above .9 to whichever bias hypothesis was drawn in step (1); and (6) repeat the foregoing five steps twenty-five times just as in the evolutionary computations. I ran these simulations ten times, in which the average total Brier score was 3,798.82 (± 347.43) for the Bayesians and 524.91 (± 59.57) for the group using the "best" settings and in which the average toss at which a majority assigned a probability above .9 to the truth was 118.58 (± 10.00) for the Bayesians and 29.90 (± 4.12) for the other group. In other words, by using the settings that the evolutionary computations showed to be optimal, a group can, in the type of situation considered here, on average achieve a more than sevenfold increase in accuracy and at the same time an almost fourfold increase in speed relative to a group of strict Bayesian updaters.

7.4 Optimal Group Learning in Action

There are a great many variations one could consider in the framework developed in the foregoing. In this section, I want to look at a variation that puts the simulations reported in the previous chapter in a social perspective. Instead of *single* doctors, consider *teams* of doctors, each team being, for example, fifty strong. All doctors in a team are independently gathering evidence about one and the same patient whose life they are jointly trying to save. Suppose that a team decides to intervene precisely if a majority of its members assign a probability above .9 to the same hypothesis. Teams are evaluated in the same way individual doctors were evaluated in the simulations presented in chapter 6. In particular, they are evaluated on the basis of treating one hundred patients, where the relevant score is the average probability that a patient they treat will live. Here, I look only at the kind of environment in which that probability comes from a Weibull distribution. Interested readers may wish to use the code in the Jupyter notebook belonging to this section to explore a Gamma environment (or other environments, by modeling probability of death via still other probability distributions).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

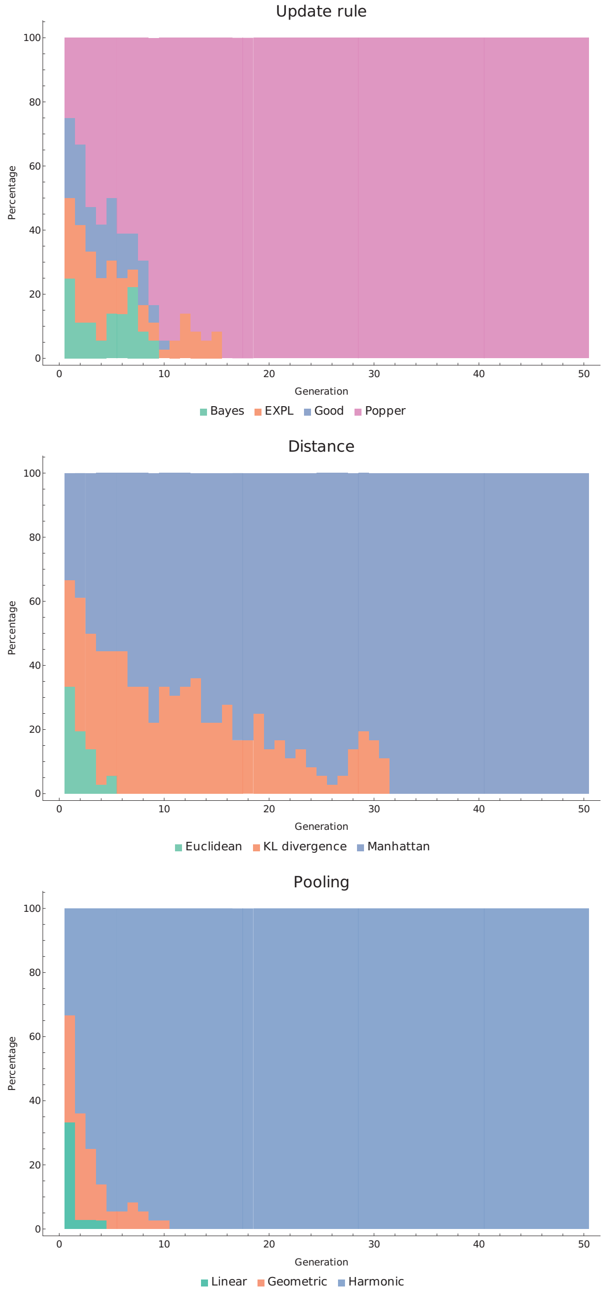

Figure 7.9: For a randomly chosen run, per-generation percentages of rule types (*top*), distance types (*middle*), and pooling types (*bottom*).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

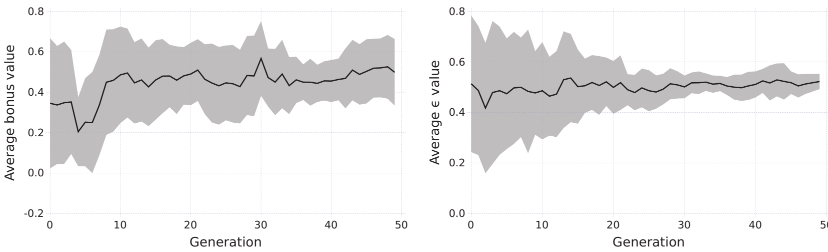

Figure 7.10: Average bonus and ε values for the fifty generations in the same run that generated the data shown in figure 7.9; shaded bands indicate one standard deviation from the mean.

To see whether we can make epistemic recommendations to such teams, we rerun the evolutionary procedure described in the previous section. We keep everything the same, except, as said, how the teams are scored (average survival probabilities in place of speed and accuracy). Figures 7.9 and 7.10 show the outcomes of a randomly selected simulation. This simulation is representative of the lot at least insofar as Popper's rule is again the overall winner: Of the 36 × 15 = 540 groups in the last generations, 276 consisted of Popperians, 193 of EXPL users, and 71 of Bayesians, groups using Good's rule never making it until the end. This is in line with what we found in chapter 6 for individual doctors in a Weibull environment. And identical to what we found in the previous section, all groups in all simulations came to use the harmonic pooling method. As for distances, the outcome is different now: whereas in the previous simulations all last-generation groups used the Euclidean metric, in the new simulations this is so for only 180 of the last-generation groups, the other groups using either the KL divergence (252) or the Manhattan metric (108). Taken over all last-generation groups in all simulations, the average bonus value assigned by last-generation groups is .49 (± 0.16) and the average ε value is .52 (± 0.03).

To emphasize once again the importance of *social* aspects to learning, I mention that averaged over all simulations, the probability that a patient would live went up from .78 (± 0.01) in the first generation to .95 (± 0.01) in the last. Compare this with the fact that in the simulations presented in the previous chapter, in which patients were treated by individual doctors, the average probability for the last generation that a patient would live was

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

only .87 (± 0.02). In real life, there is a reasonable expectation that a patient's chances of survival will go up when they are looked at by several doctors instead of just by one. But that is at least partly because we would expect each of the doctors to bring her or his own expertise, complementing that doctor's colleagues' expertise. Nothing of this kind is involved in our simulations. That we can nevertheless expect a better outcome from a group of doctors can be explained only by the fact that they learn in part by taking the opinions of their colleagues into account.

Further details can be obtained from the Supplementary Materials. For now, I note that the new results buttress the main message of this book that abduction is not a rule that is to be followed mechanically but rather designates a mode of reasoning that needs to be filled in depending on the situation, where it may require skill to find the best fill-in for any given case. We see here in particular that it may matter to how we ought to reason abductively whether we are acting completely on our own or are part of a group and revise our opinions not only on the basis of evidence that we receive directly from the world but also in response to the opinions of other group members.

7.5 The Upshot

This chapter continued the comparison among update rules begun in chapter 6, specifically pitting Bayes's rule against various probabilistic versions of abduction. It did so as part of a broader research question, to wit, whether we can recommend best epistemic practices to groups of agents who independently receive evidence from a common source. To address this question, we again used the method of agent-based optimization, and more specifically an evolutionary algorithm, although now the units of selection were groups of agents, where the agents within any one group could communicate with each other and thus impact each other's belief states. Assuming that a group is interested in striking the best balance between speed and accuracy, and making certain assumptions about the evidence-generating process, it was found that one can make a clear recommendation in the form of a unique combination of update rule, distance measure, and pooling method, together with an indication of the optimal range for the c and ε values, the former indicating the explanatory bonus and the latter willingness to count others as peers.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

This result was reached by assuming that speed and accuracy mattered equally. As said, however, this assumption is not guaranteed to hold. In some situations, we may be willing to sacrifice quite a bit of accuracy for the benefit of greater speed, whereas in other situations the opposite may hold. The code that was used for the simulations and that is in the Supplementary Materials allows interested readers to verify that giving different weights to speed and accuracy can lead to vastly different outcomes. For instance, rerunning the simulations with all the weight given to accuracy consistently led to a combination of Popper's rule, the Euclidean metric, and geometric pooling as the optimal combination, with values for c and ε in the mid-to-high range. Apart from the relative importance of speed and accuracy, the evidence in the simulations also had properties that do not hold generally—for instance, if the data really concern coin tosses, one might be more likely to encounter biases close to .5 than ones more to either extreme—and varying the evidence-generating process can affect the outcomes as well. All this is underpinned by the simulations reported in the previous section, in which as in the simulations presented in chapter 6, we focused not on speed and accuracy directly but rather on survival probabilities, and we let those determine, indirectly, which trade-off between speed and accuracy to choose. This led to results concerning best combinations of epistemic principles that were different from those obtained in section 7.3.

As discussed in the previous chapter, cognitive psychologists working on theories of ecological rationality have generally rejected as an illusion the idea, cherished by philosophers, that rationality can be captured by a number of universally valid principles and have argued that we can understand the notion of rationality only if we take into account (1) the various ways—biological and cognitive—in which humans are limited; (2) the environments in which humans are to operate and to interact; and (3) the goals that humans have in those environments. Biological, cognitive, and environmental constraints can all vary greatly from one individual to another and even from one moment to another. As explained earlier, this led Elqayam, Gigerenzer, and others to conclude that whether someone reasons or acts rationally depends on whether the reasoning or behavior helps the person to achieve whatever goal she pursues in whichever environment she is in, given the cognitive resources available to her in that environment. These researchers were concerned with the rationality of individuals, not of groups, but the results from this chapter suggest that their proposal makes sense at the group level as well.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

The upshot is that we *can* make recommendations for how to organize research groups epistemically, but these recommendations will depend on what type of phenomenon or process the group is studying and on what the aim of a group is. In view of the aforementioned work on ecological rationality, a stronger conclusion could not have been expected. And as for the overarching topic of this book, we have seen further evidence that it can make a lot of sense to urge the members of groups to reason abductively (in some form or other), and more sense than to recommend Bayes's rule. Again, this is not to discredit Bayes's rule, but rather to push back against the claim that violations of that rule necessarily indicate epistemically irresponsible behavior.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a portion of the eBook at doi:10.7551/mitpress/14179.001.0001

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

This is a section of doi:10.7551/mitpress/14179.001.0001

The Art of Abduction

By: Igor Douven

Citation:

The Art of Abduction
By: Igor Douven
DOI: 10.7551/mitpress/14179.001.0001
ISBN (electronic): 9780262369923
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024

MIT Press Direct

© 2022 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC BY-NC-ND license. Subject to such license, all rights are reserved.

CC BY-NC-ND

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in EB Garamond by the author in LuaLaTeX.

Library of Congress Cataloging-in-Publication Data

Names: Douven, Igor, author.
Title: The art of abduction / Igor Douven.
Description: [Cambridge, Massachusetts] : Massachusetts Institute of Technology, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2021031324 | ISBN 9780262046701 (paperback)
Subjects: LCSH: Abduction (Logic) | Reasoning. | Practical reason. | Bayesian statistical decision theory.
Classification: LCC BC199.A26 D68 2022 | DDC 160—dc23
LC record available at https://lccn.loc.gov/2021031324

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2235448/c003700_9780262369923.pdf by guest on 30 May 2024