Approximating Bayesian inference through internal sampling

Joakim Sundh

Adam Sanborn

Jian-Qiao Zhu

Jake Spicer

Pablo León-Villagrá

Nick Chater

People must often make inferences about, and decisions concerning, a highly complex and unpredictable world, on the basis of sparse evidence. An "ideal" normative approach to such challenges is often modeled in terms of Bayesian probabilistic inference. But for real-world problems of perception, motor control, categorization, language comprehension, or common-sense reasoning, exact probabilistic calculations are computationally intractable. Instead, we suggest that the brain solves these hard probability problems approximately, by considering one, or a few, samples from the relevant distributions. By virtue of being an approximation, the sampling approach inevitably leads to systematic biases. Thus, if we assume that the brain carries over the same sampling approach to easy probability problems, where the "ideal" solution can readily be calculated, then a brain designed for probabilistic inference should be expected to display characteristic errors. We argue that many of the "heuristics and biases" found in human judgment and decision-making research can be reinterpreted as side effects of the sampling approach to probabilistic reasoning.

Keywords: Bayesian inference, Internal sampling, Probability, Cognitive biases, Approximation, Rationality

Subject index: Cognitive science, Behavioral science, Psychology, Behavioral economics, Computational modeling

Acknowledgement

Sundh, Sanborn, Zhu, Spicer, and León-Villagrá were supported by a European Research Council consolidator grant (817492-SAMPLING).

Introduction

Humans frequently make judgments and decisions regarding present or future states of a world of which they are at least partially ignorant; in psychology this is generally referred to as probability judgments and decision making under risk. Considering that there is little that is certain in life, we can reasonably assume that humanity would have to be equipped with some mechanism for making such inferences, ideally as quickly and accurately as possible, or we would have been hard-pressed to survive this long. Nevertheless, people regularly and systematically make simple mistakes (e.g., judgment biases) that, on closer inspection, seem rather obvious even without any formal schooling in either probability theory or psychology.

Many of these biases can be explained due to the fact that the information we receive from the world is in itself biased. The world we live in is large and complex, and we are often aware of only a limited part of it at a time, either because only a small portion of information is available to us or because we do not have the time or capacity to enumerate the totality of our surroundings. Instead, we typically experience only a few glimpses of the world around us, which we usually refer to as samples, and if those samples are biased in some way, then the judgments and decisions that we make based on that information are likely to be biased as well. It is important to note that this will be the case even if the observer has a complete model of how information is generated and processes it flawlessly (e.g., Le Mens & Denrell, 2011; Konovalova & Le Mens, 2020).

Yet, there remain certain biases that cannot be explained by the structure of the environment or biases in the information gathering process, primarily those that demonstrate incoherence in human cognition. Coherence implies adherence to the principles of logic and probability theory (e.g., Kolmogorov's axioms), which will necessarily be reflected in the information sampled from the world. For example, probability theory tells us that the probability of an event occurring plus the probability of an event not occurring will be equal to one (i.e., $P(A) + P(not\text{-}A) = 1$); conversely, if we count the number of rainy days during a week, then the number of rainy days plus the number of days without rain will obviously equal the total number of days of the week, regardless of whether our sample is representative of the average precipitation in our area or not. Therefore, since information sampled from the environment will necessarily be coherent according to the theorems of probability theory, the conclusions of a perfectly Bayesian mind should be coherent as well. Of course, perfect Bayesian accounts of actual cognitive processes are computationally intractable (e.g., van Rooij, 2008; van Rooij & Wareham, 2012), since the calculations involved in modeling every single possible state of any system relating to our physical environment would be overwhelmingly complex. Consequently, we must assume that the mind is limited to, at best, approximating the Bayesian solutions to inference problems.

Fortunately, we can extend the information-sampling paradigm to apply to the internal workings of the mind as well. In this case, we assume that, rather than explicitly calculating optimal Bayesian solutions, the brain approximates these solutions by considering a small number of samples from relevant distributions. While Bayesian calculation is difficult, sampling is often surprisingly easy; even when it is not possible to represent a distribution analytically it might be possible to sample from it, which is why sampling is often used as a tool in computational statistics and machine learning. Additionally, these sampling processes can explain many apparent idiosyncrasies in human probabilistic inference, including the aforementioned incoherence biases, because of a combination of internal sampling with compensatory strategies for small samples and computationally rational sampling algorithms. (For an in-depth introduction to the algorithms by which these sampling processes are realized, see Chapter X in the same volume.)

Approximating Bayesian inference by sampling

The internal sampling processes that we are discussing here differ from the process of sampling information from the environment (which, for clarity, we will refer to as ‘external sampling’) in that we are primarily concerned with sampling taking place within the mind itself. This is not to say that internal and external sampling processes do not each contribute to various quirks in human behavior but, as previously observed, the way internal sampling is realized can explain certain incoherence biases that external sampling cannot. It is clear that these two types of sampling interact, since much of the information one samples from the mind will have been influenced by what has been sampled from the environment (see the Conclusions section for a brief discussion on this point), but for simplicity we will limit ourselves to internal sampling in this chapter.

The basic principle of internal sampling is that, for each inference, a number of samples are drawn from an internal distribution, which is in turn based on information sampled from the environment (see Figure 1). As an illustration, let us consider the standard statistical example of estimating the proportion of blue and red balls in an urn. As per usual, one can sample balls from the urn and make a frequency judgment based on the proportion of blue and red balls in the sample. But we need not assume that the information associated with the sampled balls is discarded after being sampled, rather we can assume that it contributes to the basis of a second ‘internal’ urn. The obvious advantage of the internal urn is that whenever one is subsequently obliged to make a judgment regarding the proportion of blue and red balls in the external urn, one can use the internal urn as a proxy.

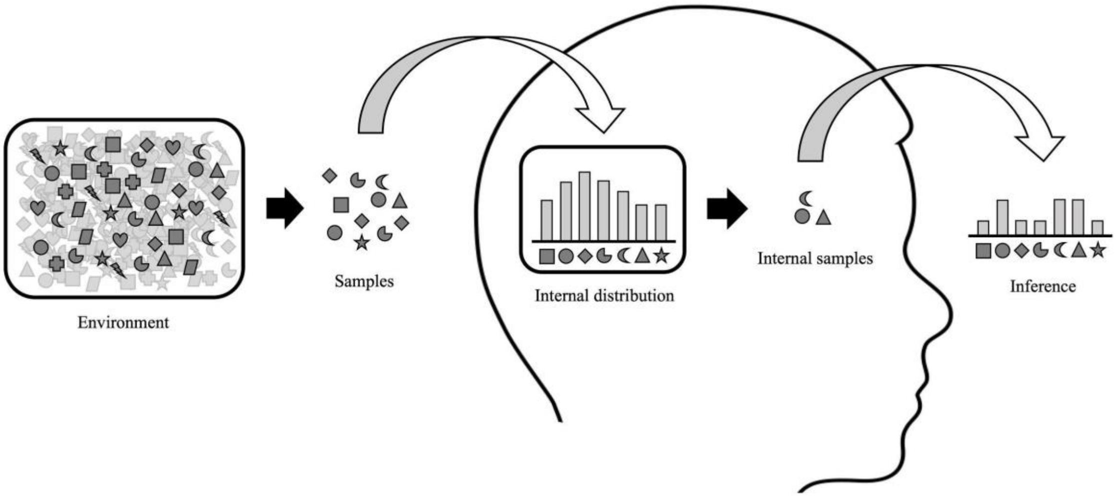

Figure 1. Schematic illustration of the internal sampling process: Selections of information are sampled from the environment, which contribute to an internal distribution. When an inference is made, a small number of samples are drawn from the internal distribution and an inference is created on the basis of those samples.

Sampling has the desirable quality that, in the limit (i.e., with an infinite number of samples), it will approach the "true" distribution. Of course, we cannot expect the brain to process an infinite number of samples, nor necessarily a large one. Human cognition is inevitably subject to computational constraints, implying an unstated upper bound on human rational inference, usually referred to as bounded rationality (Gigerenzer & Selten, 2002; Lieder & Griffiths, 2020; Simon, 1955). It is, therefore, more reasonable to assume that people only use enough samples to make "good enough" inferences about the world. Evidence further suggests that rational inferences might indeed only require very few samples, even as few as a single one in some situations (Vul, Goodman, Griffiths, & Tenenbaum, 2014).

The phenomena of probability matching illustrate this process nicely. Research has shown that, when choosing among alternatives with stochastic payoffs, people tend to choose alternatives in proportion to the probability of their payoffs rather than sticking with the alternative that would maximize their expected outcome, even when they have reliable knowledge of the outcome distributions (e.g., Koehler & James, 2009). If we assume that, for each decision, one draws a limited number of samples from an internal distribution (i.e., the internal urn) then alternatives will consequently be chosen according to their proportions in this distribution, which, assuming that the external sampling process is unbiased, will match the proportions of the environment (i.e., the external urn).

There are empirical as well as methodological hurdles for the internal sampling account to overcome. From the empirical perspective, there are a number of systematic biases, primarily those that imply incoherence in human probabilistic inference, that cannot be explained by sampling alone. From the theoretical perspective, a small number of samples might be enough to make rational decisions in many cases, but they often create obvious inaccuracies when making judgments. Additionally, although sampling is usually computationally easy, it might not be possible to represent the relevant distribution to sample from, meaning that independent and identically distributed sampling is not possible.

Fortunately, as we will demonstrate in the following sections, we can find solutions to the empirical hurdles by solving the methodological ones. Firstly, we will describe how Bayesian adjustment for a small number of samples, expressed by the *Bayesian sampler* model's prior on responses, can account for incoherence biases such as subadditivity and the conjunction fallacy, as well as repulsion and representativeness effects. Secondly, we will describe how sampling algorithms that do not require global knowledge of the sampling space can explain heavy-tailed changes and autocorrelations in human cognition, as well as supply an additional account for subadditivity, anchoring, and representativeness. We will see that both of these mechanisms generate departures from standard probability theory: the prior on responses through the small number of samples and how this is compensated for, and the alternative sampling methods through the introduction of correlations between samples and the consequent biases. These mechanisms can provide explanations of various classic heuristics and biases in the literature on judgements and decision making (see Table 1) and indeed, as we will see, the same phenomena can often be generated by *both* types of mechanisms.

Table 1.

Phenomena	**Mechanism**
Subadditivity	Explicit subadditivity: Each judgment is adjusted, so the sum of a number of judgments of the components of a category will be adjusted more than a single judgment of the category (**PR**) Implicit subadditivity/superadditivity: Combined judgments of common components of a category are judged as more likely than the category, while combined judgments of uncommon components of a category are judged as less likely than the category (**SA**)
Conjunction fallacy	Conjunctions between events are based on fewer samples than individual events, and are therefore judgments of conjunctions are adjusted more (**PR**) If sampling starts in a rich region for the individual events, more occurrences of the individual events are likely to be found and therefore they are likely to be judged as more likely (**SA**)
Anchoring	Gradually exploring the sample space means that responses will be biased by the starting point of the sampling processes (**SA**)
Repulsion	Optimal stopping implies that sampling is more likely to stop when evidence favors a particular decision, shifting judgments away from the boundary (**PR**)
Representativeness	Counting sufficiently similar “near misses” (i.e., approximate Bayesian computation) will bias judgments towards alternatives that are similar to available samples (**PR**)
Random number generation	Sampling algorithms moving through the distribution with momentum will struggle with generating truly random sequences (**SA**)
Heavy tailed changes in individual aggregate behavior (e.g., potentially in artificial and real financial markets)	Heavy tailed changes arise from the MC^3 sampling algorithm (Metropolis-coupled Markov chain Monte Carlo), in which several autocorrelated samplers run in parallel, at different levels of noise, with the possibility of switching the lowest noise level (the lowest temperature) to the current best guess (**SA**; see Chapter X for further details)
1/f noise	Changes in behavioral measures, from rhythmic tapping (Gilden, Thornton & Mallon, 1995), to reaction times in standard psychological experiments (van Orden, Holdon & Turvey, 2003) often show characteristic autocorrelations at all scales over time; MC^3 sampling generates this pattern, but most sampling algorithms, and indeed other types of cognitive model, do not (**SA**)

PR = Prior on Responses; **SA** = Sampling Algorithms

A prior on responses: Adjustment for limited samples

As previously described, internal sampling can explain the stochasticity of individual human probabilistic inference but cannot in itself explain biases that imply that people's judgments and decisions are incoherent on average. Firstly, people's probability judgments tend to be subadditive, meaning that when making separate probability judgments of the components of a category (e.g., "rain", "snow", or "any other type of precipitation"), the sum of these separate judgments is higher than the judgment of the category itself (e.g., "any type of precipitation"; e.g., Redelmeier, Koehler, Liberman, & Tversky, 1995). Secondly, people tend to make conjunction fallacies, by judging the probability of the conjunction of two (or more) events (e.g., rainy and windy weather) as higher than either of the individual events (e.g., rainy weather; e.g., Tentori, Crupi, & Russo, 2013; Tversky & Kahneman, 1983).

Internal sampling can potentially result in subadditivity and conjunction fallacies for individual judgments solely on account of the variability inherent in a small number of samples, but if the samples are representative of the internal distribution, then the average of repeated sampling will approach the average of the distribution, meaning that people's judgment should be unbiased on average. If people compensate for limited samples in a Bayesian manner, however, then biases such as these are unavoidable, and some compensation is often necessary in order to avoid obviously inaccurate inferences. For example, if one drew one blue ball from an urn with an unknown proportion of blue and red balls, then one would most likely not conclude that the urn contained only blue balls based on that evidence alone. Similarly, if one flipped a coin once and it came up heads, then one would certainly not conclude that the coin will always come up heads. Rather, assuming that any combination of red or blue balls in the urn is equally likely (i.e., a uniform prior), the Bayesian estimate is .67 blue balls. As for the coin, we usually have a strong prior that the probability of heads is .5, so we would reasonably require a rather long run of heads to change our opinion.

The Bayesian sampler model (Zhu, Sanborn, & Chater, 2020) formalizes this intuition with respect to human probability judgments. This model predicts that, for each probability judgment, a small number of instances are sampled from an internal representation, for example by drawing them from long-term memory or performing mental simulation, and the judgment is based on the proportion of outcomes in the sample after being adjusted according to a prior on responses. Given a symmetric prior expressed by the beta distribution Beta(β, β), the average judgment of the Bayesian sampler is

$$E\left(\hat{P}(A)\right) = \frac{N}{N+2\beta}P(A) + \frac{\beta}{N+2\beta}. \tag{1}$$

We can see that when the number of samples N increases, the first term will approach one and the second term will approach zero, while when the prior parameter β increases, the first term will approach zero and the second term will approach .5. Therefore, the Bayesian sampler predicts that, for a limited number of samples, human probability judgments will be adjusted towards the middle of the probability scale. This type of conservatism is indeed what we see in human behavior; as a rule, people's probability judgments tend to be less extreme than one would expect (Edwards, 1968; Erev, Wallsten, & Budescu, 1994; Hilbert, 2012; Peterson & Beach, 1967).

The Bayesian sampler can explain the subadditivity effect described above, by predicting that samples of lower (or higher) probabilities are adjusted more than probabilities close to .5. Because the components of a category are necessarily less probable than the category itself, the probabilities of each of them are likely to be more heavily adjusted than the probability of the category as a whole. The sum of the individual adjustments of the components is larger than the single adjustment of the category, which will in turn result in subadditivity. The conjunction fallacy, on the other hand, can be explained if we assume that fewer samples are used when judging the probabilities of conjunctions than when judging the probabilities of single events (a reasonable assumption, since conjunctions are more complex than individual events and therefore presumably more difficult to sample). In this case, conjunctions will be adjusted more than individual events due to the smaller number of samples, and conjunction fallacies will result, given that the probabilities of the conjunctions and the individual events are close enough to each other.

The Bayesian sampler thus differs from some other sampling-based models in that biases are not due to “naive” or “myopic” interpretation of small samples (e.g., Juslin, Winman, & Hansson, 2007) but rather due to a well-founded correction process that will, in the long run, improve accuracy. This does not necessarily imply that we can expect equivalent correction processes in all sampling-based inference; probability judgments in particular allow for a correction process that is both intuitively and mathematically accessible due to the application of a Beta prior, which might not always be the case. For example, naive use of small samples causes confidence intervals to be too narrow, but it is difficult to correct for this without knowing the form of the distribution. Nevertheless, the Bayesian sampler invites us to view traditional sampling models in a new light and ask ourselves what types of corrections (if any) people make in other situations. Ultimately, although beyond the scope of this chapter, applying a Bayesian perspective on extant models of internal as well as external sampling might allow for many new insights.

The use of a Beta prior also allows for the definition of an optimal stopping rule for sampling, which in turn provides an explanation for the repulsion bias sometimes observed in decision making, describing that comparisons of perceptual features against predefined boundaries can lead to subsequent estimates to be repulsed from that boundary. For example, judging whether the number of dots appearing in an array is above or below 25 causes an aversion to responses of 25 in later direct estimates (Zhu, Sanborn and Chater, 2019). In such cases, a decision-maker might take samples from a sensory representation to guide their answer, collecting evidence for either 'higher' or 'lower'. In combination with the prior, this evidence can then be used to determine the value of continued sampling: with each sample, the decision-maker can weigh the projected benefit of information from further samples against the time and effort required to retrieve those samples, stopping when the cost exceeds the benefit. Such a rule can then lead to biases in the aggregated samples since sampling is more likely to stop where collected evidence favors a particular decision, even if those samples do not accurately reflect the target distribution. Any subsequent judgement made based on those samples will then reflect that bias; in the case of the dots, a series of samples suggesting the response 'higher' could lead to early termination, but possible overestimation when attempting to determine the actual count, thus shifting responses away from the boundary (Zhu, Sanborn and Chater, 2019; see Chapter XX for treatment of related phenomena).

This viewpoint also provides a natural mechanism in terms of which we can understand the probabilistic biases that are often ascribed to the representativeness heuristic (Tversky & Kahneman, 1972). Suppose a person is asked to estimate the probability of, say, a fair coin falling heads five times in a row (HHHHH) versus a mix of heads and tails (e.g., HTTHH). To solve this directly by sampling sequences of five coin flips and comparing the relative frequencies of five heads versus the different mixes of heads and tails would require mentally simulating these coin flips many times; indeed, given that the true probability is just 1/32, a large sample of at least several hundred repetitions would be required to give a reasonably accurate estimate of either sequence. Clearly, for more complex events, the size of the sample required will grow very rapidly. But a sampling approach can still be applied, by drawing a small number of samples and seeing whether they are *similar* to the event of interest. For example, to work out the rough probability of crashing your car, it would be inefficient (and dangerous!) simply to wait for a sufficient number of crashes to obtain a reliable estimate. Observing a few near-misses however can inform us that this is, perhaps, more probable than we had expected, and consequently that we should drive more cautiously. Indeed, in computational statistics, this intuition is embodied in the method of Approximate Bayesian Computation (Beaumont, 2019), where a probability is estimated from the number of samples that are sufficiently similar to the target event. But this approximation may be misleading in some circumstances, depending on the psychology of similarity judgements. For example, if we mentally simulate or recall typical short sequences of coin flips, most will be an unpredictable mix of heads and tails, and therefore judged as similar to the equally patternless HTTHH, but few if any of a small sample of sequences will be judged as similar to HHHHH. Thus, counting "near misses" will lead to the erroneous conclusion that the irregular sequence is in fact more probable.

"A sense of location:" Sequential dependencies in internal sampling

So far, we've assumed that people can draw independent samples from their internal model of a probability distribution, but in reality, this is neither computationally feasible nor consistent with the operation of human memory. For example, if a person is asked to think of as many different animals as they can, starting with *lion* would probably prompt a sample of other animals native to Africa, such as *zebra*, *antelope*, *hippopotamus*, and so on, while starting with *whale* might lead to primarily sampling other animals that live in the ocean. The same principle can also be applied to problem-solving. If given a scrambled word such as CIBRPAMOLET most people will find it very difficult to discern what the unscrambled word is supposed to be, which is hardly surprising considering that there are 39,916,800 different possible combinations. From a sampling perspective, this is indeed a very large distribution to sample over and finding the correct solution by direct sampling would require quite a lot of time. On the other hand, if we are given a starting point closer to our goal, such as PROBELMATIC, then it is much easier to reach a correct conclusion (Sanborn, Zhu, Spicer, Sundh, León-Villagrá, & Chater, 2021).

Dependencies in internal sampling are perhaps most evident when people are asked to generate random sequences; human-generated sequences of random numbers show too many adjacent numbers, too few repetitions of digits, and continuing along the number line too often (e.g., 1-2-3) compared to truly random sequences (Wagenaar, 1972). These results have been taken as evidence that people are attempting to ensure that their sequences are locally representative of a random sequence (Kahneman & Tversky, 1972), or alternatively that they generate sequences using schemas (Baddeley, 1998).

Although fully random sampling is generally the most efficient way of ensuring that samples are representative, it is clearly psychologically and computationally unrealistic since it requires global knowledge of the target distribution. Instead, there are more psychologically reasonable alternatives that retain the guarantee of convergence to the target distribution in the limit of a large number of samples. In particular, Markov Chain Monte Carlo (MCMC) algorithms work by gradually exploring the sampling space, meaning that each sample will be dependent on the sample that came before, and therefore provides an alternative explanation for these deviations from truly random sequences. MCMC algorithms make local proposals, which account for the higher-than-expected proportion of adjacent numbers, and some of the most popular versions of this algorithm prefer to transition to new states when the probabilities of the states are equal, which can account for the fewer-than-expected repetitions. Continuing along the number line further than expected requires a more specialized mechanism, however, such as a sampling algorithm with some momentum in how it moves, which is inherent in some popular sampling algorithms like Hamiltonian Monte Carlo (Castillo, León-Villagrá, Chater, & Sanborn, 2021). Although algorithms such as

MCMC require fewer cognitive resources than direct sampling, it will create dependencies and autocorrelations, which, as noted, is something we also observe in human behavior (see Chapter X in this volume for an in-depth introduction to sampling algorithms and their implications).

While these explanations are difficult to distinguish in random number generation, different predictions can be made for newly trained representations. Castillo et al. (2021) performed such an experiment, in which participants learned a one-dimensional or two-dimensional grid of syllables and were then asked to generate random syllables. Participants' generated sequences were better matched by moving around the representation with momentum than by the transitions between syllables in either natural language or in the training they received on the representation.

The sampling algorithm's starting point. Applying algorithms with a sense of location to internal sampling means that where you start sampling can have a large influence on what is sampled, as in the relative difficulties of unscrambling CIBRPAMOLET and PROBELMATIC. The substantial influence of the starting point has thus been used to explain a number of *framing effects*: effects on how a question is asked upon the answer that is produced. For example, while in explicit subadditivity, the sum of separate judgments (e.g., of the probabilities of "rain", "snow", or "any other type of precipitation") is higher than the probability of the combined judgments (e.g., "any type of precipitation") there are implicit versions of the task that show different results. If you were asked for a combined judgement of the probability of "rain, snow, or any other type of precipitation", then this is composed of common components that might be judged as more likely than "any type of precipitation". However, if you were asked for a combined judgment of "diamond dust, virga, or any other type of precipitation", you may judge this to be less likely than "any type of precipitation". Implicit subadditivity effects such as these depend on the likelihood of the examples given, which can be explained by the starting point of internal sampling; starting with highly probable examples means that they are unmissable, while for the packed version of the question they could be missed. Conversely, starting with highly improbable examples can make it more difficult to bring the highly probable examples to mind (Dasgupta, Schulz, & Gershman, 2017; Sanborn & Chater, 2016; Sloman, Rottenstreich, Wisniewski, Hadjichristidis, & Fox, 2004).

The starting point of internal sampling can also be used to explain the *anchoring* effect, in which estimates are seemingly pulled towards values presented in preceding comparisons. For example, participants estimating the proportion of African countries in the U.N. guess higher numbers if first comparing this figure to 65% than without such a comparison (Tversky & Kahneman, 1974). These 'anchors' might act as starting points for MCMC chains that, with a limited number of iterations, the decision-maker is unable to sufficiently escape (Lieder, Griffiths, Huys, & Goodman, 2018).

Finally, the starting point of internal sampling can explain at least some forms of the conjunction fallacy. In Tversky and Kahneman's seminal 1983 paper, the first empirical evidence they provided resulted from asking participants to estimate across four pages of a novel either the number of words that had the form "_ _ _ _ i n g" or the number of words that had the form "_ _ _ _ _ n _". While the first question asked for a subset of the words asked for by the second, participants still estimated the number of words with the form "_ _ _ _ i n g" to be higher. Like with the implicit subadditivity example, we can think of this as a result of internal sampling that starts in a richer region of the internal representation when given the form "_ _ _ _ i n g", and which has more trouble finding this richer region from the starting point "_ _ _ _ _ n _" (Sanborn & Chater, 2016).

The sampling algorithm's movement. Aside from the starting point of internal sampling, the way in which sampling moves through a mental representation is also important for explaining aspects of human behavior. As noted in the animal-naming example, the contents of the mind that are relevant for answering a question might be divided into clusters (e.g., animals native to Africa, animals that live in the ocean, etc.), and so are internally sampled in a clustered fashion (Bousfield & Sedgewick, 1944; Hills, Jones, & Todd, 2012).

This problem of retrieving clustered items has been extensively studied in the animal foraging literature, as animals face the parallel challenge of obtaining food that is distributed into patches (e.g., berry bushes that cluster together with substantial distances between the clusters). The distances that animals travel have been found to correspond to power-law distributions in these kinds of environments. This means that while there is a high probability of travelling short distances, there is also a substantial probability of travelling large distances; the probability of each distance is proportional to that distance raised to a (negative) power. This implies that we should expect any type of concept generation or memory sampling to demonstrate similar patterns where information is similarly clustered in the mind. These distributions of movements appear to be adaptive, as the most effective blind search means moving according to a power-law distribution with an exponent of negative two (Viswanathan, Buldyrev, Havlin, da Luz, Raposo, & Stanley, 1999). Interestingly, the dynamics of internal sampling show a surprising correspondence to animal foraging. When retrieving animal names, there are long delays between retrieving names that come from different clusters (Hills, Jones, & Todd, 2012), and delays between retrievals are distributed according to a power-law distribution with an exponent that is close to negative two (Rhodes & Turvey, 2007). This suggests that internal sampling is able to move effectively in a clustered mental representation.

The way in which sampling moves through a mental representation does not just depend on the previous state, as it might if location were all that mattered. Instead, internal sampling shows long-range autocorrelations, with the next state of internal sampling depending on long-ago states as well. This has been demonstrated in a wide range of dependent measures, including repeated estimates of time intervals (e.g., one second), estimates of spatial intervals (e.g., one inch), and the response times of repeated trials of tasks such as lexical decision, mental rotation, visual search, and implicit association (Correll, 2008; Gilden, 1997; Gilden, Thornton, & Mallon, 1995). In particular, these dependencies have been characterized as 1/f noise, which is notable because such a pattern is not straightforward to produce (Gardner, 1978). This means, in practice, that the response you make does not just depend on your last response but also on responses you made many trials ago.

The combination of power-law distributions and 1/f noise in human cognition is interesting for two reasons, one theoretical and one practical. It is of interest to psychological theory because it is even more difficult to produce power-law distance distributions and 1/f noise in tandem; the most common models of each effect (i.e., Levy flights for distance distributions and fractal Brownian motion for 1/f noise) do not produce the other effect. It is also very useful for distinguishing between sampling algorithms. Of all the algorithms reviewed in Chapter X, only Metropolis-coupled MCMC seems capable of reliably producing both effects, which is of particular interest as this is also the best algorithm in this set for sampling from a clustered representation (Zhu, Sanborn, & Chater, 2018).

The combination of power-law distance distributions and 1/f noise is of practical interest because it is very similar to some of the stylized facts found in financial markets. Both individual asset prices and indices like the S&P 500 show heavy-tailed price increases and decreases: while there are many days in which the prices change a little bit, there are a few days in which it changes dramatically. In addition, markets have long-range autocorrelations in their volatility: the magnitude of a price change (though importantly not the direction of a price change) depends on the magnitude of previous price changes (Cont, 2001). The correspondence between these effects found in internal sampling and those found in financial markets allows for the tentative suggestion that there may be a relationship between these levels that past market-based explanations have overlooked (Chater, Sanborn, Zhu, & Spicer, 2020).

Putting the pieces together: The Bayesian sampler with a sense of location

In the above sections, we discussed two possible mechanisms for explaining various classic biases in the literature on judgements and decision making. These mechanisms do not necessarily have to act separately, but instead can work together within the same model. As an illustration, we return to the previously described numerosity task in which participants are asked to judge whether the number of dots presented briefly on a computer screen is greater than 25. We can integrate the two mechanisms by assuming that a sampling algorithm with a sense of location, such as MCMC, generates samples from a distribution of estimates of the number of dots observed, for example, 24-26-27-26. Next, samples are re-coded according to the decision boundary (i.e., greater than 25), so that they can be expressed as 0-1-1-1. The confidence in the hypothesis "greater than 25" is then a result of combining these re-coded samples with the Beta prior, and a (potentially varying) threshold in confidence is used to determine when to stop sampling and make a decision (Zhu, Sundh, Chater, & Sanborn, 2021).

This simple model reproduces an impressive number of stylized facts about the relationships between confidence, accuracy, and response time. It also gives a novel explanation of why erroneous response times are often slower than correct response times, especially when participants are instructed to focus on accuracy, being a result of the autocorrelation of the samples. Finally, beginning to sample for the next trial where sampling for the previous trial left off allows the model to produce the long-range autocorrelations in response times that can explain the general variability in this measure (Gilden, 1997; Zhu, et al., 2021).

Discussion

On the one hand, human judgment and decision-making are subject to a number of systematic biases. On the other hand, human behavior in much more complex problems than the simplified tasks commonly used in judgment and decision making studies have been shown to be close to ideal/normative Bayesian inference in areas such as perception (Kersten, Mamassian, & Yuille, 2004), categorization (Anderson, 1991; Lake, Salakhutdinov, & Tenenbaum, 2015; Sanborn, Griffiths, & Navarro, 2010), reasoning and argumentation (Hahn & Oaksford, 2007; Oaksford & Chater, 1994, 2020), and intuitive physics (Battaglia, Hamrick, & Tenenbaum, 2013; Sanborn, Mansinghka, & Griffiths, 2013). The common factor of these latter approaches, sometimes collectively labelled Bayesian cognitive science, is that probabilistic reasoning, rather than heuristics, is considered the core component of human cognition, which is often viewed as Bayesian in the sense that it is adapted to make rational inferences based on subjective uncertainty (Oaksford & Chater, 2009). Yet the aforementioned biases seemingly contradict this perspective, implying an apparent paradox when different areas of cognitive science are compared.

In this chapter, we have demonstrated how internal and external sampling can relate to each other, using the metaphor of an external and an internal urn illustrated in Figure 1, and that the concept of internal sampling can be used to extend the information sampling paradigm in order to create a more complete account of human probabilistic inference and associated biases. External sampling in itself can explain biases in the correspondence of judgments and decisions to the external environment, because biased information will persist even if one were a perfect Bayesian inference machine, yet it cannot account for biases that imply incoherence. To understand this better, we can first make the simplifying assumptions that the brain perfectly stores and retrieves unmodified and unbiased samples from the environment to make inferences (e.g., Shi, Griffiths, Feldman, & Sanborn, 2010). After storing a small number of external samples, draws from the internal urn will on average not be very close to the true environmental value: there will be a random deviation of the mean of the internal urn from the true environmental value. Only with a large number of external samples will draws from the internal urn be on average close to the true environmental value. This serves as a (highly simplified) model of expertise: more samples will mean that judgments will on average correspond more closely to the environment, given a suitable inference algorithm.

But the number of draws from the internal urn plays a separate role in the coherence of judgments. If we assume that the information that has previously been sampled from the environment (i.e., from the external urn) is in turn sampled with replacement from the mind (i.e., from the internal urn) for each individual judgment or decision, then using small samples that are adjusted appropriately can be used to explain many of the most well-known and persistent biases observed in human behavior. A person who draws a small number of samples will show greater inconsistency when asked the same question a second time, and assuming a prior on their responses, will show greater explicit subadditivity and conjunction fallacy effects. It is only with a very large number of samples that a person would be perfectly coherent and consistent, no matter the number of external samples in their urn. Thus, internal sampling processes, perhaps related to cognitive capacities such as working memory capacity (Lloyd, Sanborn, Leslie, & Lewandowsky, 2019), can unite seemingly contradictory results, by the assumption that people indeed make rational inferences, given computational restrictions and small samples.

In conjunction, the number of internal and external samples can explain the surprising finding that experts often demonstrate the same biases as amateurs, despite more experience and expertise (e.g., Redelmeier, Koehler, Liverman, & Tversky, 1994; Reyna & Lloyd, 2006; Tversky & Kahneman, 1983). Going back to the central metaphor, although the experts by virtue of their expensive experiences and greater opportunity for sampling presumably have a larger number of balls in their internal urn, they still draw the same number of samples when making a decision. Therefore, although experts might have an internal urn that better represents the nature of the world, incoherence biases will nevertheless occur.

Although most of the findings discussed in this chapter can also be explained by other (sometimes many other) theories, one of the main strengths of the internal sampling account is that it simultaneously explains so many of them. Additionally, because it will, in the limit, result in a perfectly calibrated inference, sampling is based on a rational foundation that other theories sometimes lack. As such, internal sampling stands out as one of the most complete accounts of incoherence in human probabilistic inference. Furthermore, there are some cognitive phenomena that sampling is uniquely equipped to tackle, such as probability matching (e.g., Koehler & James, 2009), since it is very difficult to account for such stochasticity without an underlying stochastic process such as sampling. An important next step in further validating the internal sampling account is to determine if human behavior exhibits more such characteristic patterns; for example, recent research has shown that the variance of human probability judgments is consistent with a binomial process, where the variance can be used to approximate the number of samples used (Sundh, Zhu, Chater, & Sanborn, under review).

Of course, there are many complications to this picture. Alongside the internal sampling mechanisms we discuss above, information samples from the environment are often biased, and participants at least partially correct for these biases as well as incorporate other information into their decisions (Hayes, Banner, Forrester, & Navarro, 2019). Instead of assuming a frequentist interpretation of these samples (e.g., Costello & Watts, 2014), we instead take a Bayesian perspective in which samples are drawn according to subjective degrees of belief that have already incorporated prior beliefs and whatever correction for environmental biases that people use. Furthermore, internal sampling differs from external sampling in the sense that we cannot observe the distributions or the samples directly. There are indications that sampling-based accounts of cognition are roughly compatible with those used in neural sampling models (Buesing, Bill, Nessler, & Maass, 2011; Fiser, Berkes, Orbán, & Lengyel, 2010; Hoyer & Hyvarinen, 2003; Moreno-Bote, Knill, & Pouget, 2011), but until this connection can be modelled on an implementation level (see Marr, 1982) internal sampling methods remain as-if models. Nevertheless, a clear strength of the internal sampling account is that we know that it could indeed be implemented by the structure and machinery of the brain.

Conclusions

In this chapter, we have considered how the human mind is able to deal with a complex and uncertain world. The “ideal” approach to dealing with uncertainty is often thought to be Bayesian probabilistic reasoning, but the calculations involved are hopelessly intractable for dealing with the challenges of the real world. Instead, we suggest the mind approximates Bayesian inference by sampling small numbers of items from the relevant probability distributions, where these samples are generated through mental simulation or by drawing on memory. Indeed, this is a widely adopted strategy in computational statistics and artificial intelligence, where various tractable schemes for sampling from complex distributions have been developed.

Approximate Bayesian inference using sampling will inevitably lead to systematic departures from precise Bayesian probabilistic calculation though. We have considered two systematic ways in which this is the case. First, if we draw samples, probability estimates from these samples have to be modified by prior knowledge (or many events will be assigned definitive probabilities of 0 or 1), but this adjustment will itself lead to systematic biases in probability judgments, as we saw arising from the Bayesian sampler model of probability judgement (inflating small probabilities will lead, for example, to subadditivity). Second, for problems of realistic complexity, samples cannot be drawn independently from the probability distribution of interest, which is, after all, complex and unknown. Instead, samples will be generated by local sampling methods such as Markov Chain Monte Carlo methods, and thus introduce "autocorrelated" samples, where each sample tends to be similar to prior samples. Such methods will, in the limit, sample accurately from the underlying distribution; but, of course, the cognitive system must make do instead with a small number of samples, so that the starting point of the sampling process, in particular, will have strong impacts on the sample drawn, and hence on the resulting probability judgements. We have seen that both sources of systematic bias can help explain well-known phenomena traditionally captured in the heuristics and biases framework pioneered by Kahneman and Tversky; thus, the conjunction fallacy, and biases associated with anchoring, representativeness, and others, seem naturally to arise from a sampling framework.

In contrast to many of the chapters of this book, we have focused primarily on internal sampling, through mental simulation or from memory, rather than potentially biased sampling arising from our interactions with the external world. But there are potentially interesting connections between sampling from the mind and sampling from the environment, which are likely to be interesting to explore in future research. For example, it may be that mental samples simply mirror samples we experience from the environment (Anderson, 1990; Stewart, Chater & Brown, 2006), so that biased sampling of the environment may be reflected in mental sampling. Or, in many social contexts, one person's mental sampling feeds into another person's environment (e.g., if one person's judgements, arising from mental sampling, are then communicated to others); so, biases in mental sampling may shape biases in environmental sampling. Other more complex interactions can, of course, also be imagined. We suggest that combining theories about mental sampling and sampling from the environment, together with the quirks and biases of both, is likely to be a useful direction for understanding both the remarkable human ability to cope with a complex and uncertain world, and our tendency to make systematic errors on even the most elementary reasoning problems.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review*, *98*(3), 409.

Baddeley, A. (1998). Random generation and the executive control of working memory. *The Quarterly Journal of Experimental Psychology: Section A*, *51*(4), 819-852.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327-18332.

Beaumont, M. A. (2019). Approximate Bayesian computation. *Annual Review of Statistics and its Application*, *6*, 379-403.

Bousfield, W. A., & Sedgewick, C. H. W. (1944). An analysis of sequences of restricted associative responses. *The Journal of General Psychology*, *30*(2), 149-165.

Buesing, L., Bill, J., Nessler, B., & Maass, W. (2011). Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons. *PLoS Computational Biology*, *7*(11), e1002211.

Castillo, L., León-Villagrá, P., Chater, N., & Sanborn, A. (2021). *Local sampling with momentum accounts for human random sequence generation*. Manuscript submitted for publication.

Chater, N., Sanborn, A. N., Zhu, J. Q., & Spicer, J. (2020) Macroeconomic Implications of the Sampling Brain. *Rebuilding Macroeconomics Working Paper Series*.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance, 1:2*, 223-236

Correll, J. (2008). 1/f noise and effort on implicit measures of bias. *Journal of personality and social psychology*, *94*(1), 48.

Costello, F., & Watts, P. (2014). Surprisingly rational: probability theory plus noise explains biases in judgment. *Psychological Review*, *121*(3), 463.

Dasgupta, I., Schulz, E., & Gershman, S. J. (2017). Where do hypotheses come from?. *Cognitive Psychology, 96,* 1-25.

Edwards, W. (1968). Conservatism in human information processing. *Formal Representation of Human Judgment.* Hoboken, NJ: John Wiley & Sons.

Erev, I., Wallsten, T. S., & Budescu, D. V. (1994). Simultaneous over-and underconfidence: The role of error in judgment processes. *Psychological Review, 101*(3), 519.

Fiser, J., Berkes, P., Orbán, G., & Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences*, *14*(3), 119-130.

Gardner, M. (1978). White and brown music, fractal curves and one-over-f fluctuations. *Scientific American*, *238*(4), 16-27.

Gigerenzer, G., & Selten, R. (Eds.). (2002). *Bounded rationality: The adaptive toolbox*. MIT press.

Gilden, D. L. (1997). Fluctuations in the time required for elementary decisions. *Psychological Science*, *8*(4), 296-301.

Gilden, D. L., Thornton, T., & Mallon, M. W. (1995). 1/f noise in human cognition. *Science*, *267*(5205), 1837-1839.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: a Bayesian approach to reasoning fallacies. *Psychological Review*, *114*(3), 704.

Hayes, B. K., Banner, S., Forrester, S., & Navarro, D. J. (2019). Selective sampling and inductive inference: Drawing inferences based on observed and missing evidence. *Cognitive psychology*, *113*, 101221.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological Bulletin*, *138*(2), 211.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological Review*, *119*(2), 431.

Hoyer, P. O., & Hyvärinen, A. (2003). Interpreting neural response variability as Monte Carlo sampling of the posterior. In *Advances in neural information processing systems* (pp. 293-300).

Juslin, P., Winman, A., & Hansson, P. (2007). The Naïve Intuitive Statistician: A Naïve Sampling Model of Intuitive Confidence Intervals. *Psychological Review*, *114*(3), 678–703.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430-454.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271-304.

Koehler, D. J., & James, G. (2009). Probability matching in choice under uncertainty: Intuition versus deliberation. *Cognition*, *113*(1), 123-127.

Konovalova, E., & Le Mens, G. (2020). An information sampling explanation for the in-group heterogeneity effect. *Psychological Review*, *127*(1), 47.

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332-1338.

Le Mens, G., & Denrell, J. (2011). Rational learning and information sampling: On the “naivety” assumption in sampling explanations of judgment biases. *Psychological Review*, *118*(2), 379.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*.

Lieder, F., Griffiths, T. L., Huys, Q. J., & Goodman, N. D. (2018). The anchoring bias reflects rational use of cognitive resources. *Psychonomic Bulletin & Review*, *25*(1), 322-349.

Lloyd, K., Sanborn, A., Leslie, D., & Lewandowsky, S. (2019). Why higher working memory capacity may help you learn: Sampling, search, and degrees of approximation. *Cognitive Science*, *43*(12), e12805.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. MIT press.

Moreno-Bote, R., Knill, D. C., & Pouget, A. (2011). Bayesian sampling in visual perception. *Proceedings of the National Academy of Sciences*, *108*(30), 12491-12496.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608.

Oaksford, M., & Chater, N. (2009). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, *71*, 305-330.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin*, *68*(1), 29.

Redelmeier, D. A., Koehler, D. J., Liberman, V., & Tversky, A. (1995). Probability judgment in medicine: Discounting unspecified possibilities. *Medical Decision Making*, *15*(3), 227-230.

Reyna, V. F., & Lloyd, F. J. (2006). Physician decision making and cardiac risk: effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied*, *12*(3), 179.

Rhodes, T., & Turvey, M. T. (2007). Human memory retrieval as Lévy foraging. *Physica A: Statistical Mechanics and its Applications*, *385*(1), 255-260.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, *20*(12), 883-893.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychological Review*, *117*(4), 1144.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review*, *120*(2), 411.

Sanborn, A. N., Zhu, J.-Q., Spicer, J., Sundh, J., León-Villagrá, P. & Chater, N. (in press). Sampling as the human approximation to probabilistic inference. In S. Muggleton & N. Chater (Eds). *Human-Like Machine Intelligence*. Oxford: Oxford University Press.

Shi, L., Griffiths, T. L., Feldman, N. H., & Sanborn, A. N. (2010). Exemplar models as a mechanism for performing Bayesian inference. *Psychonomic Bulletin & Review*, *17*(4), 443-464.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99-118.

Sloman, S., Rottenstreich, Y., Wisniewski, E., Hadjichristidis, C., & Fox, C. R. (2004). Typical versus atypical unpacking and superadditive probability judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 573.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology*, *53*(1), 1-26.

Sundh, J., Zhu, J. Q., Chater, N., & Sanborn, A. (2021). The mean-variance signature of Bayesian probability judgment. *Under review*.

Tentori, K., Crupi, V., & Russo, S. (2013). On the determinants of the conjunction fallacy: probability versus inductive confirmation. *Journal of Experimental Psychology: General*, *142*(1), 235.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*(4), 293.

Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, *32*(6), 939-984.

Van Rooij, I., & Wareham, T. (2012). Intractability and approximation of optimization theories of cognition. *Journal of Mathematical Psychology*, *56*(4), 232-247.

Viswanathan, G. M., Buldyrev, S. V., Havlin, S., Da Luz, M. G. E., Raposo, E. P., & Stanley, H. E. (1999). Optimizing the success of random searches. *Nature*, *401*(6756), 911-914.

Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science*, *38*(4), 599-637.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, *77*(1), 65.

Zhu, J. Q., Sanborn, A. N., & Chater, N. (2018). Mental sampling in multimodal representations. *Advances in Neural Information Processing Systems*.

Zhu, J. Q., Sanborn, A. N., & Chater, N. (2020). The Bayesian sampler: Generic Bayesian inference causes incoherence in human probability judgments. *Psychological Review*.

Zhu, J. Q., Sanborn, A., & Chater, N. (2019). Why Decisions Bias Perception: An Amortised Sequential Sampling Account. In *CogSci* (pp. 3220-3226).

Zhu, J. Q., Sundh, J., Chater, N., & Sanborn, A. N. (2021, February 4). The Autocorrelated Bayesian Sampler for Estimation, Choice, Confidence, and Response Times. https://doi.org/10.31234/osf.io/3qxf7