Philos Stud
DOI 10.1007/s11098-013-0267-7

Learning experiences and the value of knowledge

Simon M. Huttegger

Abstract Generalized probabilistic learning takes place in a black-box where present probabilities lead to future probabilities by way of a hidden learning process. The idea that generalized learning can be partially characterized by saying that it doesn't foreseeably lead to harmful decisions is explored. It is shown that a martingale principle follows for finite probability spaces.

Keywords Value of knowledge · Decision theory · Reflection principle · Probabilistic learning

1 Introduction

Bayesian conditioning is the standard model of probabilistic learning. Yet, as Jeffrey (1965, 1992) has convincingly argued, there are many modes of learning from experience that cannot be fit into the structure provided by conditioning. Jeffrey himself proposed probability kinematics as an alternative that applies to learning experiences where no proposition is learned for certain; the learning experience may only yield uncertain evidence, changing the probabilities of propositions to something less than one.

Probability kinematics still assumes that there is a set of propositions with respect to which learning takes place. But Jeffrey's undogmatic epistemology has been carried much further. Beyond probability kinematics one may consider learning

S. M. Huttegger (✉)
Department of Logic & Philosophy of Science, University of California, Irvine, 3151 Social Science Plaza A, Irvine, CA 92697, USA
e-mail: shuttegg@uci.edu

S. M. Huttegger
Munich Center for Mathematical Philosophy (MCMP), Geschwister-Scholl-Platz 1, 80539 München, Germany

Published online: 31 December 2013

situations where an agent just changes her probabilities in response to an experience without being able to express what happens in terms of a set of propositions. The learning process is hidden. We only know the inputs (initial probabilities) and the outputs (future probabilities). Skyrms (1990) appropriately calls this "black-box learning".

There is an intriguing principle, known as *reflection*, that has been put forward as a requirement for such belief changes (van Fraassen 1984). The principle says that your future probabilities after the learning experience should cohere with your current probabilities. It was suggested by Skyrms (1990) that this principle provides a plausible way to distinguish learning situations from situations where one expects probabilities to change for other reasons, such as getting drunk, having a brain lesion or having a dangerously low blood sugar level. Here I want to investigate one avenue for making this claim more precise. I develop an account in which the reflection principle is a necessary condition for a black-box probability update to count as a *genuine learning experience*. The link between the two is provided by the idea that genuine learning experiences should not lead you to expect to make harmful decisions (in certain decision situations) after changing your beliefs. The key to this idea is the *value of knowledge theorem* of decision theory, to which I turn first.

2 The value of knowledge

The value of new information manifests itself distinctly in the decision theory of Savage (1954) in which a precise formulation of the following is a theorem: *The expected utility of an uninformed decision cannot be greater than the prior expectation of an informed decision*. In epistemological contexts, any decision theory that violates this principle should, prima facie, be regarded with some suspicion; at least in certain idealized situations, expecting more knowledge should not lead you to expecting to make worse decisions.

The following treatment goes back to Good (1967).[1] In order to keep things simple, everything is assumed to be finite. This allows us to focus on the basic ideas and avoid thorny issues in the philosophical foundations of integrals and measures.[2]

The value of knowledge theorem is about a second order decision problem. The first order decision problem consists of n acts $A_1,\ldots,A_n$, m states of the world $S_1,\ldots,S_m$, and a utility function u for conjunctions of acts and states. The second order problem is that you can either make a choice in the first order problem now, or defer a choice until after the outcome of an experiment is revealed to you. The

[1] Savage (1954) also considers a version of this result. There is a note by F. P. Ramsey that anticipates the theorem (published as Ramsey 1990; see also Skyrms 1990 on this and on Good's argument).

[2] The arguments presented here do not hold generally in infinite probability spaces if the probability measure is not countably additive (Kadane et al. 1996, 2008). Many scholars—de Finetti (1974) first and foremost among them—think that countable additivity should not be thought of as a coherence requirement. See Seidenfeld (2001) for a recent overview of reasons to take finite additivity as basic. See also Zabell (2002) and references therein on finitely additive martingale theory, which would be relevant for an extension of our results to more general situations.

experiment consists of a finite partition $\{E_k\}$. One E_k will be announced to you as the true member of the partition.

Should you decide now or wait until you get more information? Suppose that you follow Savage's decision theory (1954) and choose the act A that maximizes your expected utility

$$\sum_i p(S_i)u(A\&S_i),$$

where $p(S_i)$ is your present probability for state S_i and $u(A\&S_i)$ is your utility for the outcome where the state of the world is S_i and you choose A.

If you decide now, the value of your decision is given by

$$\max_j \sum_i p(S_i)u\left(A_j\&S_i\right). \tag{1}$$

If you wait until after the experiment is performed, you may condition your probabilities on the new information. Suppose that E is the true member of $\{E_k\}$. For simplicity we assume that $p(E_k)$ is positive for every member of the partition. If you are a conditionalizer, you will update your degrees of belief to the future probabilities that are given by

$$p(S|E),$$

for any state S. After getting the information E, the posterior expected value for act A is equal to

$$\sum_i p(S_i|E)u(A\&S_i).$$

If you will choose the act that maximizes your future expected utility, then your prior expectation of your future expected utility is

$$\sum_k p(E_k) \max_j \sum_i p(S_i|E_k)u\left(A_j\&S_i\right). \tag{2}$$

This is the present value of making a decision after being informed about the outcome of the experiment.

It is a straightforward exercise to show that the value of choosing later (2) is greater than or equal to the value of choosing now (1), and is strictly greater unless the same act maximizes expected utility regardless of which of the E_k occurs (Good 1967). Let us call this the "value of knowledge relation".

The value of knowledge theorem tells you that more information won't result in making foreseeably harmful decisions; and if the experiment is at all relevant to the choice of an act, more information is actually expected to be helpful. This should be understood in terms of your present epistemic state. Information may be harmful in specific situations, but not with respect to your prior expectations. Hence, if you follow Savage's decision theory you will usually choose to wait until you have more information.

A little observation that turns out to be of some importance later should be noted at this point. We may take the outcomes of the experiment to be part of the state of the world. In this case the utilities may depend on acts, states and outcomes of experiments.[3] The sum (1) corresponds to

$$\max_j \sum_{i,k} p(S_i \& E_k) u\left(A_j \& S_i \& E_k\right),$$

and (2) corresponds to

$$\sum_k p(E_k) \max_j \sum_{i,k} p(S_i \& E_k | E_k) u\left(A_j \& S_i \& E_k\right),$$

which is equal to

$$\sum_k p(E_k) \max_j \sum_{i,k} p(S_i | E_k) u\left(A_j \& S_i \& E_k\right),$$

since $p(S_i|E_k) = p(S_i \ \& \ E_k|E_k)$. The value of knowledge theorem clearly also holds for these expressions.

Skyrms (1990) investigates an extension of the value of knowledge theorem for black-box learning.[4] Suppose that you are in a learning situation where you expect your probabilities to change from p to one of the future probabilities in the set $\{p_f\}$ after making an observation or undergoing some other kind of experience. For simplicity we assume that there are only finitely many p_f that you consider to be possible. p and p_f are probability functions[5] on some finite Boolean algebra that consists of all Boolean combinations of the states of the world. For each state S your probability $p(S)$ changes to one of the values $p_f(S)$. There is not assumed to be an evidential proposition that describes the information that was obtained. What is being learned is only given by the effect on posterior probabilities. We assume that prior to the learning experience you have second order probabilities $p(p_f)$ over the possible posterior probabilities p_f.

Is there any rule that governs belief dynamics in this extremely general setting? Probabilists have sometimes put forward the following principle:

$$p\left(A|p_f\right) = p_f(A). \tag{M}$$

Principle (M) is also known as reflection. It says that your future probability for A should be equal to your prior conditional probability for A given your future probability. The relation given in (M) has deep roots in probability theory, where it expresses the martingale condition.[6] Martingales are enormously important because they provide the foundation for many laws of large numbers. I shall refer to (M) as the 'martingale principle' because 'reflection' is also used for other, related requirements (see van Fraassen 1995).

[3] If the utilities don't depend on the outcomes of the experiment, the terms here reduce to the earlier ones.

[4] Graves (1989) presents a generalization for probability kinematics.

[5] That each posterior p_f is a probability function means that you expect to be statically coherent.

[6] That's why it's called (M) here and in Skyrms (1990).

As Skyrms (1990) shows, the martingale principle also leads to the value of knowledge relation. Suppose that (M) holds. Then (assuming that each $p(p_f)$ is positive) your value for choosing an act now is

$$\begin{aligned}\max_j \sum_i p(S_i)u(A_j\&S_i) &= \max_j \sum_i \sum_f p(S_i|p_f)p(p_f)u(A_j\&S_i)\\ &= \max_j \sum_f \sum_i p_f(S_i)p(p_f)u(A_j\&S_i).\end{aligned}$$

The value of choosing after the learning experience is

$$\sum_f p(p_f) \max_j \sum_i p_f(S_i)u(A_j\&S_i).$$

The latter term cannot be less than the former term on general mathematical grounds. The value of knowledge relation continues to hold.[7]

The value of knowledge theorem is not unconditionally true (Kadane et al. 2008). Let me highlight some of the underlying assumptions. In the first place, the experiment is assumed to be essentially costless; otherwise you may not always wish to wait for more information even if it would be useful. Moreover, you know that you are an expected utility maximizer and that you will be one after learning the true member of the partition. Third, the states, acts and utilities are the same before and after the learning experience. Also, having the learning experience does not by itself alter your probabilities for states of the world (although the outcomes of the experience usually do); the learning experience and the states of the world are probabilistically independent. Furthermore, in the classical theorem you know that you will update by conditioning; in Skyrms' extension, you know that you will honor the martingale principle. Finally, by working within Savage's decision theory we import several further assumptions into the value of knowledge theorem, notably the probabilistic independence of states and acts (choosing an act does not give any information about the state).

These are idealized assumptions, many of which serve to minimize the differences between the current decision problem and the future decision problem. The only aspect that is supposed to change is that you have gained information. The idealization thus seems to be legitimate if one's goal is to isolate the influence of learning experiences on decision making. This requires that only the learning experience may alter the probabilities of states, and this is the only change that takes place. The value of knowledge in real life situations is of course intermingled with other aspects that are not captured in the theorem.

3 Genuine learning experiences

The generalized value of knowledge theorem proceeds from the premise that the martingale principle (M) captures the dynamics of beliefs in learning situations.

[7] See Myrvold (2012) for a similar treatment with epistemic utility functions.

Only then do we again obtain the result that learning usually leads one to expect to make better decisions. Thus the martingale principle is supposed to hold for genuine learning experiences (this claim or similar ones can be found in, e.g., Jeffrey 1988; Skyrms 1990, 1997; Myrvold 2012; Huttegger 2013); it is, after all, intended to be a generalization of Bayesian conditioning.

This means, in particular, that the martingale principle should not be applied to belief changes in epistemologically defective situations. In situations of memory loss, of being brainwashed or being under the influence of drugs, (M) should obviously not hold. If you believe that in an hour you will think you can fly because you're about to consume some funny looking pills, then you should not already now have that belief.

So, the martingale principle is claimed to apply if you learn something in the black-box, but not if you learn nothing or other things happen besides learning. The problem with this explanation of the martingale principle is that the underlying learning situation is a black-box; we only know the input (current degrees of belief) and the output (posterior degrees of belief) but not how the agent gets from here to there. Why should we think that what happens inside the black-box really is a legitimate way to obtain information? We cannot point to anything concrete that's going on in the black-box since then it would not be a black-box anymore. What we need is a way to say more about what's happening inside the black-box without opening it. This means that any successful approach has to be indirect.

Let me start by mentioning two approaches to justify the martingale principle for belief dynamics. One is in terms of diachronic Dutch book arguments (Goldstein 1983; van Fraassen 1984) and the other one in terms of minimizing inaccuracy (Huttegger 2013). The first one shows that violations of (M) lead to inconsistencies between present and future degrees of belief. Such violations lead to inaccurate future beliefs on the second approach. Both results suggest that violations of the martingale principle lead to future beliefs that are in some sense non-ideal from your current perspective. But they don't establish an explicit link between the martingale principle and what genuine learning experiences are.

I submit that one way to establish such a link is to exploit the value of knowledge relation. The idea is to make precise how genuine learning should affect what you think about particular decision situations. If you anticipate to change your beliefs to p_f and you presently think that the belief change is based on a genuine learning experience, then you should expect the new beliefs to help you in making decisions in *all relevant decision problems*.

The rationale for this implication is given by the paradigm case of Bayesian learning. We have seen that in Savage style decision problems conditioning implies the value of knowledge relation. Since we think of the black-box process as an extension of conditioning, generalized learning should also entail the value of knowledge relation. The latter is of course only required to hold for those decision situations that meet the assumptions mentioned in the preceding section; that is, the underlying decision problems are in the style of Savage, the utilities are the same before and after the learning experience, you are certain to act as a Bayesian maximizer now and in the future, and so on. We consider only these decision problems because in others the conclusion of the value of knowledge theorem need not always hold for the paradigm case of Bayesian conditioning. These other kinds of decision situations may not always be suitable for detecting whether a shift from p to p_f constitutes a genuine learning experience, while the uncontroversial decision problems used in the preceding section are candidates for detecting learning experiences that go beyond mere conditioning.

Let me now try to make this idea more precise. We again consider a situation where you change your probabilities for states $p(S)$ to $p_f(S)$, p_f being one of a finite number of posterior probabilities in $\{p_f\}$, and where you have prior second order probabilities over posterior probabilities $p(p_f)$. We also assume that the $p(p_f)$ are all positive. In order to have a rich enough set of decision problems, we suppose that for every finite combination of utility values, each of which lies in an interval around zero, there are acts $A_1,\ldots,A_n$ such that $u(A_i\&S_j)$ assume these values.[8] A genuine learning situation is partially characterized in the following way:

Postulate. If a belief change from p to $\{p_f\}$ constitutes a *genuine learning situation*, then

$$\sum_f p(p_f) \max_j \sum_i p_f(S_i)u(A_j\&S_i) \geq \max_j \sum_i p(S_i)u(A_j\&S_i), \tag{3}$$

for all utility values $u(A_i\&S_j)$ with strict inequality unless the same act maximizes expected utility irrespective of which of the p_f occurs.

The left side of (3) is the prior expectation of the value of your posterior choice after the learning experience, which is described by the shift from p to p_f. The right side is the value of choosing an act now.

We take the states S_i to describe the world precisely enough so that S_i includes information about which posterior p_f obtains. As was noted above, the analogous situation for the case of conditioning (states specify the outcomes of experiments) does not endanger the value of knowledge theorem. Since the classical value of knowledge theorem is the primary source for our postulate, requiring a fine enough description of states in the general framework appears to be legitimate. That the description of states is sufficiently precise implies that utilities are allowed to depend on posteriors.

The requirement that we should be able to vary the utilities as we please (at least within certain bounds) captures the idea that a genuine learning experience should not be harmful, and perhaps help, for making decisions whatever the decision problem at hand is. If a belief change leads you to foreseeably make worse choices than you could already make now in some decision situations, then it cannot be a pure learning experience. Perhaps you are bolder after having taken those funny looking pills, for example. From your current perspective, this might help you in some decision problems, but it will be harmful in others.

Notice that the postulate only formulates a necessary condition for genuine learning. No claim is made as to the sufficiency of this criterion. There may well be other features of genuine learning that are not captured here.

[8] More formally, suppose that there are $n \cdot m$ intervals around zero where m is the number of states. For $i = 1,\ldots, n$, $j = 1,\ldots, m$ let u_{ij} be an arbitrary number in the ijth interval. We assume that for all u_{ij} there are acts $A_1,\ldots, A_n$ such that $u(A_i\&S_j) = u_{ij}$.

Finally, notice also that the plausibility of the postulate does not depend on whether or not you think that the value of knowledge theorem is a requirement for all rational decision theories. As I mentioned earlier, the theorem fails to hold in many alternatives to Savage's decision theory. The reason why these alternatives don't play a role here is because for them the value of knowledge relation already fails for conditioning and is hence not a trustworthy indicator that something is learned.

4 From learning to the martingale principle

It is now not difficult to derive the martingale condition (M) from (3). Since we assume that all $p(p_f)$ are positive (3) implies that

$$\sum_f p(p_f) \max_j \sum_i p_f(S_i) u(A_j \& S_i) \geq \max_j \sum_f \sum_i p(S_i | p_f) p(p_f) u(A_j \& S_i).$$

Suppose that all utilities for conjunctions of acts and states are zero except for $u(A\&S)$, which is assumed to be positive. Such an assignment of utilities exists because of our assumption that the class of possible decision situations is sufficiently rich. This particular assignment is easy to generate. Just consider a payoff scheme where only the outcome $A\&S$ results in a positive payoff, say, \$1; every other outcome has payoff zero.

Now (3) reduces to

$$\sum_f p(p_f) p_f(S) u(A\&S) \geq \sum_f p(S|p_f) p(p_f) u(A\&S).$$

We now assume that $p_f(S) = 0$ if p_f is not the posterior in state S. This requires that you will always be certain about the posterior you arrived at after the learning experience; otherwise you would think it were possible that you get utility $u(A\&S)$ even though you are not in state S and really get a utility of zero. The corresponding relation holds automatically for the conditional probabilities $p(S|p_f)$. If p_f is not the posterior given in S, then $p(S|p_f) = 0$.[9]

With this assumption it follows that

$$p(p_f) p_f(S) u(A\&S) \geq p(S|p_f) p(p_f) u(A\&S),$$

where p_f is the posterior in state S. This of course implies

$$p_f(S) \geq p(S|p_f). \tag{4}$$

The argument can clearly be repeated for all states S. If p_f is not the posterior in state S, then (4) holds because in this case $p_f(S) = 0 = p(S|p_f)$. Hence (4) holds for all S and all posteriors.

[9] It follows that this issue does not arise in Skyrms' value of knowledge theorem. If the martingale principle is assumed to hold, then $p_f(S) = p(S|p_f) = 0$ if p_f is not the posterior in state S.

Now suppose that for some state S

$$p_f(S) > p\left(S|p_f\right).$$

Then there is a state S' such that $p(S'|p_f) > p_f(S')$, for otherwise we would have

$$\sum_i p_f(S_i) > \sum_i p\left(S_i|p_f\right) = 1.$$

(The equality holds because the states constitute a partition.) But this contradicts that (4) holds for all states. It follows that $p_f(S) = p(S|p_f)$ for all S. $\square$

The proof shows that *if updating your probabilities constitutes a genuine learning situation*, *then principle* (M) *holds*. Connecting learning to the value of knowledge allows one to argue that the martingale principle is a necessary condition for genuine learning.

Recall that if

$$p\left(S|p_f\right) = p_f(S),$$

for all states S, then (3) holds for all decision problems (this follows from Skyrms' argument mentioned earlier). Hence the martingale principle and the value of knowledge relation are in a certain sense equivalent. This was claimed in Skyrms (1997), who refers to a kind of conditional bet in order to support the claim of necessity. The bet is conditional on the posterior, and so is related to the decision situation used in the foregoing argument. However, the example does not make clear under what conditions necessity of the martingale principle holds; nor does Skyrms clearly link the value of knowledge relation to genuine learning situations as we do in the postulate.

5 Conclusion

The martingale principle should in general only be applied to situations where new beliefs are expected to be solely arrived at by learning. The principle holds whenever you expect that a belief change does not result in making harmful decisions in certain decision problems. These are the decision problems for which the classical value of knowledge theorem holds. So they are also the ones that should hold for generalized learning experiences that are not influenced by factors other than obtaining more information. This forges the link between the martingale principle and learning.

For ideal learning situations the martingale principle thus constitutes a necessary condition. This is a limiting case; the principle will fail in non-ideal cases where learning is one factor among others that influence a belief change. Even in those cases the martingale principle should hold approximately as long as the situation approximates the ideal case. Saying something more precise about this approximation is a difficult problem, but not one that is specific for the martingale principle. The philosophy of scientific modeling teaches us that one of the goals of modeling is to isolate certain factors of interest and to understand their influence apart from what happens in processes where it acts together with many other factors. It is, however, usually very hard to say something about how the different factors work together. Thus, notwithstanding the fact that one can always do more, having an understanding of the role of one factor provides us with valuable information. The same is true for the martingale principle.

References

de Finetti, B. (1974). *Theory of probability* (Vol. 1). London: Wiley.

Goldstein, M. (1983). The prevision of a prevision. *Journal of the American Statistical Association, 78*, 817–819.

Good, I. J. (1967). On the principle of total evidence. *British Journal for the Philosophy of Science, 17*, 319–321.

Graves, P. (1989). The total evidence principle for probability kinematics. *Philosophy of Science, 56*, 317–324.

Huttegger, S. M. (2013). In defense of reflection. *Philosophy of Science, 80*, 413–433.

Jeffrey, R. C. (1965). *The logic of decision*. New York: McGraw-Hill. 3rd Revised edition, 1983. Chicago: University of Chicago Press.

Jeffrey, R. C. (1988). Conditioning, kinematics, and exchangeability. In B. Skyrms & W. L. Harper (Eds.) *Causation*, *chance*, *and credence* (Vol. 1, pp. 221–255). Dordrecht: Kluwer.

Jeffrey, R. C. (1992). *Probability and the art of judgement*. Cambridge, MA: Cambridge University Press.

Kadane, J. B., Schervish M. J., & Seidenfeld, T. (1996). Reasoning to a foregone conclusion. *Journal of the American Statistical Association, 91*, 1228–1236.

Kadane, J. B., Schervish, M., & Seidenfeld, T. (2008). Is ignorance bliss? *The Journal of Philosophy, 105*, 5–36.

Myrvold, W. C. (2012). Epistemic values and the value of learning. *Synthese, 187*, 547–568.

Ramsey, F. P. (1990). Weight or the value of knowledge. *British Journal for the Philosophy of Science, 41*, 1–4.

Savage, L. J. (1954). *The foundations of statistics*. New York: Dover Publications.

Seidenfeld, T. (2001). Remarks on the theory of conditional probability: Some issues of finite versus countable additivity. In V. F. Hendricks (Ed.), *Probability theory* (pp. 167–178). Amsterdam: Kluwer.

Skyrms, B. (1990). *The dynamics of rational deliberation*. Cambridge, MA: Harvard University Press.

Skyrms, B. (1997). The structure of radical probabilism. *Erkenntnis, 45*, 285–297.

van Fraassen, B. C. (1984). Belief and the will. *Journal of Philosophy, 81*, 235–256.

van Fraassen, B. C. (1995). Belief and the problem of Ulysses and the Sirens. *Philosophical Studies, 77*, 7–37.

Zabell, S. L. (2002). It all adds up: The dynamic coherence of radical probabilism. *Philosophy of Science, 69*, 98–103.