RUDOLF CARNAP
# Probability and Content Measure

## 1. On Knowledge and Error

Before I begin the discussion of questions about probability and content measure, in which my views differ from those of Karl Popper, I should like to emphasize that I agree to a large extent with Popper's views on general questions of the theory of knowledge and the methodology of science, as represented in his *Logic of Scientific Discovery* [8] and especially in two more recent papers.[1] The picture he draws of a sharp contrast between his conception and that of the "empiricists" (or "positivists," as he frequently says) is not correct as far as my conception is concerned.

First, I agree completely with Popper's criticism of classical rationalism and classical empiricism inasmuch as both regard the source of knowledge in question (reason or senses, respectively) as "authoritative," i.e, as yielding definitive certainty. There was indeed a tendency in this direction in my book *Der logische Aufbau der Welt* (1928, but written in 1924-1925, before I came to Vienna) and later in Schlick's paper "Das Fundament der Erkenntnis." But in the Vienna Circle, I always emphasized, together with Neurath and the majority of the other members, that every factual statement is not certain (in the strict sense of this

[1] [9], reprinted as Introduction in [6]; and [10], Sects. I-X (pp. 215-233), and [6], Addenda, Sect. 2, pp. 388-391. Compare also [7], Sect. VI, and my reply in [3], Sect. 31, pp. 995-998.

NOTE: This paper is a revised version of a discussion memorandum which I contributed in December 1962 to the discussions on probability at the Minnesota Center for the Philosophy of Science. The paper was written in a period of work supported by a grant from the National Science Foundation.

word) but may always be subject to re-examination and modification. I further agree with Popper that all knowledge is basically guessing, and that the scientist has the task, not of going beyond guessing in order to come to certainties, but rather of improving his guesses.

In this paper I shall leave aside the question of the nature of the theory which I call "inductive logic." This theory is not, as Popper thinks, in contradiction to my conception in matters of general methodology as indicated above, but, on the contrary, is based on this general conception. In my view, the purpose of inductive logic is precisely to improve our guesses and, what is of even more fundamental importance, to improve bur general methods for making guesses, and especially for assigning numbers to our guesses according to certain rules. And these rules are likewise regarded as tentative; that is to say, as liable to be replaced later by other rules which then appear preferable to us. We can never claim that our method is perfect. I say all this only in order to make quite clear that inductive logic is compatible with the basic attitude of scientists; namely, the attitude of looking for continuous improvement while rejecting any absolutism. However, in the present context I cannot go into the details of these questions. I believe that my paper [1] gives an exposition of my view on the nature of inductive logic which is clearer and from my present point of view more adequate than that which I gave in my book [2]. This exposition emphasizes especially the application of inductive logic to the choice of a rational, practical decision, e.g., with respect to a bet, a business investment, or any other action whose outcome is uncertain.

## 2. Probability and Content Measure

In this paper I intend to explain my views on a special problem which Popper has discussed, namely, the question how a good scientist or a reasonable businessman or engineer chooses his hypotheses, i.e, the suppositions on the basis of which he makes his practical decisions; and especially the question how this choice is influenced by taking into consideration the probability or the content measure of a hypothesis. Here I understand the term "probability," like Popper, in the sense of logical probability, either as absolute (Popper's $P(x)$, my $m$-function) or relative to a given evidence (Popper's $P(x, y)$, my $c$-function). I use for $P(x)$ sometimes also the term "initial probability of $x$" (instead of the ambiguous term "a priori probability" in the traditional language). If $e$ is the total evidence which the observer possesses at the present moment, then let us call $P(x, e)$ "the present probability of $x$."

For the purpose of this paper, it is not essential what we take as fl exact definition of "content measure." I shall here, for the sake of simplicity, follow Popper in defining the content measure (for which he uses the term "content") as the reciprocal of the initial probability:[2]

(1) $Ct(x) =_{df} 1/P(x)$.

I wish to emphasize that I use in this definition the initial probability $P(x)$. This seems natural because the amount of information convey by $x$ is not changed by the evidence $e$, while the probability is changed. (It is also possible to introduce the relative measure of content of $x$ with respect to $e$, as I have done in [4]. But this is not a measure of the content of $x$ itself, but rather of that part of the content of $x$ which is not contained in the content of $e$.)

Now Popper asserts the following:

(2) The larger the probability of a statement, the smaller its content measure.

With this I certainly agree, provided that we mean here initial probability. Obviously the definition (1) yields the following:

(3) If $P(x) > P(z)$, then $Ct(x) < Ct(z)$.

But when a scientist considers a hypothesis $x$, be it a physical or biologic law or a prediction, and in this context speaks of its probability, then he is thinking not of $P(x)$, but of the present probability $P(x, e)$. Therefore he might interpret the thesis (2) as follows:

(2′) If $P(x,e) > P(z, e)$, then $Ct(x) < Ct(z)$.

However, (2′) is not generally valid. It is indeed valid in the special case that logical implication holds in one direction between the two hypotheses $x$ and $z$ under comparison:

(4) If $z$ L-implies $x$, but $x$ does not L-imply $z$, then
(a) $Ct(x) < Ct(z)$,
(b) $P(x) > P(z)$,
(c) For any evidence $y$ (not L-false), $P(x,y) \geq P(z, y)$.

[2] In my terminology I distinguish between "content" and "content measure" (or "measure of information"). A statement specifying the content of $x$ is an answer to the question of *what* is said by $x$ (this may be explicated, for example by the class of the non-L-true sentences which arc L-implied $x$). A statement giving g the content measure tells us *how* much is said by $x$. In [4] I proposed two different explicata for

An implication relation of the kind indicated holds in many of the cases occurring in the discussions on content measure and probability; we shall ten later give examples of this kind. Therefore, in these cases (2′) holds. But it is important to be aware that this is in general not the case. (On the other hand, the assertion (2′) would be valid for the r*elative* measure of content.)

Now we come to the question which property of a hypothesis makes it enable if there is a choice among several hypotheses. For the sake of simplicity let us consider a situation where the choice is between just two proposed hypotheses. There are under discussion the following two rules:

(5a) Choose the hypothesis with the higher content measure.
(5b) Choose the hypothesis with the higher probability.

On the basis of (2) Popper thinks that these two rules are incompatible. Sometimes he gives the impression of defending rule (5a). At any rate he believes that Keynes, I, and others plead for (5b). Actually I would reject both rules if formulated in this simple form. Elsewhere I have discussed a rule like (5b)[3] which advises choosing among incompatible hypotheses the one with the highest probability. There I show in detail, with the help of examples, why a rule of this kind should be rejected. I think that most of the authors who at the present time use a concept subjective or personal probability would likewise reject (5b) for similar reasons. I am in agreement with Popper's emphasis on the importance of the content measure for judging hypotheses and for choosing a hypothesis. And it is true that there are special cases in which the simple rule (5a) is valid, e.g., those in which the content measures of the two hypotheses differ but the other relevant factors are equal and among them, in particular, the probabilities. Thus I would agree with the following rule as a modifcation of (5a);

(6a) If two hypotheses have different content measures, while their probabilities (and other circumstances) are equal, then choose the hypothesis with the higher content measure.

this concept. I called the first one "content measure" and defined it by $P(\sim x)$ or $1 - P(x)$; the second, under the term "measure of information," was defined as $\log_2 (1/P(x))$. This distinction is not essential for the present discussion. In assimilation to Popper, I shall me here the first term, with a simplified version of the second definition. (The two definitions just given are also mentioned by Popper as alternatives.)

[3] See the discussion of rule $R_2$ in [2], Sect. 50C.

But then I should like to point out that the counterpart to this rule is likewise valid:

(6b) If two hypotheses have different probabilities but their content measures (and other circumstances) are equal, then choose the hypothesis with the higher probability.

If (2) were unrestrictedly valid, then obviously the conditions in rule (6a) would be impossible, and likewise those in (6b). Thus each of the two rules would be inapplicable. The same holds if *relative* content measure is used. However, these rules are here to be understood in my sense of the terms as explained above; that is to say, the content measure is understood as constant (not changing with growing evidence), and probability is understood in the sense of present probability. Thus the rules are to be interpreted as follows (for the sake of simplicity I omit reference to other circumstances):

(7a) If $Ct\ (x) > Ct(z)$, and $P(x,e) = P\ (z,e)$, then choose $x$.
(7b) If $P\ (x,e) > P\ (z,e)$, and $Ct\ (x) = Ct\ (z)$, then choose $x$.

Here in each of the two rules the two conditions are compatible, as we shall see in an example.

## 3. An Example

I shall now show with the help of a simplified example that higher content measure (in my sense) is not always associated with lower probability, but may also be associated with equal, and even with higher, probability.

When I consider the application of the concept of probability in science then I usually have in mind in the first place the probability of predictions and only secondarily the probability of laws or theories. Once we see clearly which features of predictions are desirable, then we may; say that a given theory is preferable to another one if the predictions yielded by the first theory possess on the average more of the desirable features than the predictions yielded by the other theory.

Suppose that a practically acting man *X*, say an engineer, asks a theoretician (in this case, a meteorologist) for a prediction *h* of the temperature at this place tomorrow at noon, a prediction to be based on a comprehensive body of evidence *e* available today. *X* intends to use the diction in order to make an adjustment of a certain apparatus which h will leave at this place so that the apparatus will automatically carry out something tomorrow at noon, say, take a photo or register the results of certain measurements. *X* is aware of the fact that it is impossible for him to obtain information today about the exact temperature for tomorrow noon; what he asks for is merely an interval prediction of the form "the temperature will be between *n* and *m*," where *n* and *m* are integers interpreted as degrees of the Fahrenheit scale. Let us assume that for the given place and time only temperatures between 0 and 100 are to be expected. Let us further assume that the hundred unit intervals of this total interval have equal initial probability; hence for each hypothesis *h* for a unit interval $(n, n + 1)$, $P(h) = .01$. For any hypothesis *h*, with an interval $(m,n)$ of the length $l = n - m$, the following holds. First, $P(h) = l/100$; thus the initial probability is proportional to *l*. Then we have by (1): $Ct(h) = 100/l$; that is, the content measure is inversely proportional to *l*. Intervals of equal length have equal initial probabilities and equal content measures. We assume that *X* knows these values. The evidence

Table 1. Values for Four Hypotheses

| Hypothesis ($h$) | Interval | Length ($l$) | $P(h)$ | $Ct(h)$ | $P(h, e)$ |
|---|---|---|---|---|---|
| $h_1$.................. | (70,71) | 1 | .01 | 100 | .04 |
| $h_2$................ | (69,72) | 3 | .03 | 33 | .12 |
| $h_3$................ | (86,87) | 1 | .01 | 100 | .01 |
| $h_4$................ | (84,89) | 5 | .05 | 20 | .05 |

*e* available today to the theoretician may include many results of previous meteorological observations which are relevant for the hypotheses to be predicted, including the weather situation of today. We assume that the theoretician has rules for the computation of logical probability. With the help of these rules he calculates, on the basis of the given evidence *e*, a probability density function for the temperature tomorrow at noon. Let us suppose that this is a normal function with the mean 70 and the standard deviation 10. Then the P-density at 70 is 0.04, and at 86.5, 0.01. If *h* is a hypothesis with an interval of a small length *l*, with the center point 70 (or 70.5), then $P(h,e)$ is approximately 0.04 *l*; and similarly, if center point is 86.5, then $P(h,e) \approx 0.01\, l$. Hence we obtain the values Table 1 as examples.

Now I shall show—which really should be obvious—that, if *X* applies either rule (5a) or rule (5b), and thus bases his choice either on content measure alone or on probability alone, he will sometimes be led to choices which are clearly wrong. Since the maximum of the density is at 70, an interval which includes this point is clearly preferable to an interval far away from this point. But each of the rules (5a) and (5b) is in many cases in conflict with this common-sense consideration. Suppose *X* always follows rule (5a). Thus, when given the choice between the two hypotheses $h_2$ and $h_3$. he will prefer $h_3$ because this is a more precise prediction and therefore has a greater content measure than $h_2$. (The relative content measure $Ct\ (h,e) = 1/P\ (h,e)$ has the value 8 for $h_2$ and 100 for $h_3$; thus it would lead even more strongly to the wrong preference.) On the other hand, let us suppose that *X* always heeds rule (5b). Then, if he has to choose between $h_1$ and $h_4$, he will have to choose $h_4$ because the present probability of $h_4$ is higher than that of $h_1$.

We can also easily see now that increase in content measure (in my sense) is not incompatible with increase in present probability. For example, in comparison to $h_4$, the hypothesis $h_2$ has both higher content measure and higher present probability.

The most important point is the following. Any prediction with an interval whose center point is at or close to 70 is preferable to any prediction with an interval around 87 or any other point removed from 70. Thus we recognize that the feature which is primarily important is the location of the interval; its length is also of interest, but only secondarily. In other words, what is primarily important is the probability density itself (more exactly speaking, the mean density in the interval in question), while the probability and the content measure are only of secondary importance.

## 4. The Optimum Interval Hypotheses

In view of our preceding considerations, we may now disregard the intervals which do not include the point 70. Then furthermore, among the intervals which include this point, we shall restrict our attention to a still narrower class, namely, that of the *optimum intervals*. We call an interval hypothesis *h* optimum if there is no other hypothesis *h′* which is preferable to *h* either according to rule (7a) or according to (7b). Thus, for example, the hypothesis with the interval (65,75) is optimum. For if *h′* is any other hypothesis with an interval of the same length 10, but with a different (present) probability, e.g., the interval (68,78), then the probability of $h'$ is lower and therefore $h$ is preferable according to rule (7b). And if $h''$ is any other hypothesis which has the same probability as $h$ (or a nearly equal but somewhat lower probability) but a different content measure (i.e., different length), then the content measure of $h''$ smaller than that of $h$ (the interval is longer than that of $h$) and therefore according to (7a) $h$ is preferable.

In Table 2 I shall give for a few selected values of $P(h,e)$ the length $l'$ of those symmetrical intervals around the mean (here 70) for which the present probability, in other words, the integral of the probability density over the interval, has exactly the value of $P(h,e)$ given in the column. (These values of $l'$ can be taken from any table for the probability integral, i.e., the integral of the normal density.) These values $l'$ refer to a continuous scale. But in our example we have restricted ourselves to a discrete scale with only integers as values. Therefore we have ' take here an interval with integer boundaries and with an integer value $l$ close to the calculated value $l'$, as stated in the table. (Since our scale is discrete, the optimum interval is often not uniquely determined.

Table 2. The Optimum Interval Hypotheses for Some Values of the Present Probability

| $P(h, e)$ | $l'$ | $l$ | Optimum Interval |
|---|---|---|---|
| .5................ | 13.4 | 13 | (64,77) |
| .6................ | 16.8 | 17 | (62,79) |
| .7 ................ | 20.8 | 21 | (60,81) |
| .8................ | 25.6 | 26 | (57,83) |
| .9................ | 32.8 | 33 | (54,87) |
| .95.............. | 39.2 | 39 | (51,90) |
| .98............... | 46.6 | 47 | (47,94) |
| .99............... | 51.6 | 52 | (44,96) |

For any odd length $l$ there are two optimum intervals: for example, for $l = 3$, both the interval (69,72) and the interval (68,71); both interval hypotheses have the same content measure and, because of the symmetry of our curve, equal probability. In order to have a unique optimum interval, we make the arbitrary convention that for the length 2n + 1 we shall take the interval (70 - n,70 + n + 1).)

Thus we restrict our consideration now to the class of optimum interval hypotheses. In this class (made unique by convention as indicated above) the following holds. Of any two distinct optimum intervals one is included in the other. Therefore of any two distinct optimum interval hypotheses, one of them L-implies the other. Hence, theorem (4) is now applicable. And therefore, within the restricted class, Popper's thesis (2′) is valid, since all cases in which (7a) or (7b) would be applicable are now excluded through the restriction to optimum intervals. In this special situation we can thus increase at most one of the two magnitudes, either probability or content measure, but not both. Are now Popper's views here correct, first the view that we have to make a choice between two alternative methods, namely, either aiming at high probability or aiming at high content measure, and second the view that a good scientist would always choose the latter method?

It seems to me that in this situation it would be wrong simply to aim at high content measure; but it would also be wrong to aim at high probability. Both high content measure and high probability are valuable features of a hypothesis. But it is unreasonable to increase as much as possible either the one or the other of the two magnitudes. The reasonable procedure consists in making a compromise solution. To some extent it is good for *X* to aim at high content measure provided that he does so within reasonable bounds. If *X* were to disregard bounds, then he would choose the interval (70,70) of length zero and hence with infinite content measure. But clearly this would be unreasonable; the probability would be zero, and that means that it would be almost certain that the chosen prediction is false and that the actual temperature lies outside of the chosen interval and therefore his adjustment would not fulfill its purpose. On the other hand, *X* might well aim at a high probability, but here again he must be careful to remain within reasonable bounds. Otherwise he would take the total interval, with probability, one. But thereby he would disregard the experience collected so far, and that would be clearly unreasonable. In order to act reasonably he must take into consideration the known experiences, and he must make his arrangements in such a way that they take care of the situation as well as possible with respect to the possible temperatures. In the last described he would adjust also to very high and to very low temperatures, although both of them are practically excluded, and thereby he would make his arrangements inefficient.

## 5. Utility

The conflict between striving for high content measure and striving for high probability is merely a special case of the following general kind of situation. Frequently in our life we have to make a choice between different things or actions of such a kind that we are motivated by two conflicting tendencies. We find two different features of which each is desirable; but the situation within the given class of cases among which we have to choose is such that, generally speaking, the higher the degree of the first feature, the lower the degree of the second. In situations of this kind, rules which advise us to choose either the maximum degree of the first feature or the maximum degree of the second feature, as rules (5a) and (5b) do, are inappropriate. The reasonable solution consists in trying to find a suitable compromise; that is to say, in choosing a thing which has a fairly high degree of both features. But how shall we determine the exact point at which the two conflicting tendencies are, so to speak, in balance? In those situations where, for any possible action, we know which of the possible outcomes that concern us directly (e.g, gain or loss of money, things, time, health, and the like) will result from the action, the answer can easily be given: we should choose that action whose outcome has the maximum utility. In general, however, the situation is such that for each of the possible actions different outcomes are possible, and we do not know which of these outcomes will actually occur. In a case of this kind we should choose that action for which the expectation value of the utility of the outcome is a maximum. The expectation value is the weighted mean of the utilities of the different outcomes, each utility value of an outcome being multiplied with a factor which is to be taken higher the more reason we have to expect just this outcome to occur. As Ramsey and de Finetti have shown, if we wish to act in a reasonable way, these factors must satisfy the axioms of the calculus of probability (compare my [1], statement 6, pages 307-308). This use of probability values may be regarded as the basic role of the concept of probability (the term "probability" understood, as generally in this paper, in the sense of either logical probability or personal probability).

We must now recognize that the discussions in the preceding sections of this paper concerning the choice of a hypothesis have only a limited validity, because they left utility out of consideration. Suppose that some-one asks himself the question: "Shall I accept the hypothesis *h* that my house will bum down at some time during the next year, or shall I rather accept its negation *h′*?" If he disregards utility considerations, then he would presumably accept the hypothesis *h′*, because it seems to be a more natural choice than the assumption *h* that his house will burn down. However, the acceptance of h′ would lead him to the decision not to take out fire insurance; and this would clearly be unreasonable.

In our previous example, *X* must take into account the gains or losses he would have in each of the hundred possible cases with respect to the temperature tomorrow if he were to adjust his apparatus in accordance with one or the other of the interval hypotheses among which he has to choose. For each of these hypotheses *X* will calculate the expectation value of his gain (counting a loss as a negative gain). Then *X* should accept as a basis for his adjustment of the apparatus that hypothesis for which the expectation value of the gain is a maximum. (This is my rule $R_4$ in [2], Section 50E; in Section 51A of the same work I explain in detail a further refinement, rule $R_5$, which advises maximizing the expectation value not of the gain but of the corresponding utility. For our present discussion we may leave out of consideration the distinction between gain and utility.)

## 6. Hypotheses of Purely Theoretical Interest

So far I have discussed the problem of the choice of a hypothesis with; respect to a practical decision to be taken by the man *X*, where the decision together with unknown circumstances in the environment of *X* will lead to a tangible gain or loss. But what about the choice of a hypothesis in the case not of an engineer or another man concerned with practical actions but of a theoretical scientist? For example, *X* may be an astronomer who considers an astronomical prediction in a situation where he does not intend to base any practical decision on the hypothesis, and therefore there is no tangible gain or loss connected with the acceptance of the hypothesis. In my view it would be advisable not to talk about the *acceptance* of this or that hypothesis in a situation of this kind, at least not if we wish to make a more exact analysis of the situation. It is true that it is customary to speak in terms of acceptance. But in my view this is an oversimplification which, though in general harmless, should still be avoided if we wish to see the situation more clearly. Instead of saying:

"Among the given hypotheses $h_1$, $h_2$, etc, I choose the hypothesis $h_1$," it would be better to say; "I assign to the hypothesis $h_1$ a high probability" (and, if possible, in more exact terms: "the probability with the numerical value $q_1$"), "to the hypothesis $h_2$ a smaller probability" (and again, if possible: "the numerical value $q_2$"), and so on. For *X* to pick out one of the hypotheses and to declare that he accepts it would give only a crude indication of the knowledge that *X* possesses with respect to the matter in question. A more exact description of his knowledge is given rather by the specification of the probability values, say by a specification of a probability distribution over the values of a certain magnitude (e.g, the probability density of temperature in our previous example). It is frequently said that actual scientists never make such ascriptions of probability values to their hypotheses. But I think that this is not quite correct. It is true that in the period of classical physics this was not frequently done. Today, however, a physicist would very frequently give a probability button for a magnitude. (I am speaking here of a probability distribution in the logical sense, characterizing his knowledge with respect to given evidence; I am not speaking of the likewise frequently occurring but quite different case where a physicist gives a probability distribution a certain magnitude in the statistical, empirical sense, namely, as indicating frequencies to be expected.) He may give the probability distribution either for a particular empirical magnitude, e.g, the length *A* of the great axis of the orbit of the earth around the sun, or for the value of general parameter which occurs in one or several basic laws, e.g, the velocity of light *c*. In cases of this latter kind we have to do with a constant, e.g., *A* or *c*. Therefore it makes no sense to speak about frequencies. There is only one value; but the exact value is not known, and for this reason the physicist gives a probability distribution. He does this frequently by a statement of the following form: "$k = 0.356 \pm 0.0012$." This statement is not meant as a declaration of the acceptance of the interval (0.3548, 0.3572) but rather as a specification of a probability distribution (which may for instance be interpreted in the following way: "The probability density for *k* on the basis of the available evidence is the normal function with mean 0.356 and standard deviation 0.0012"). But even if it were true that at the present time no scientist would ever give a specification of probability values (in the sense indicated) for hypotheses of theoretical science, then I would still say that it is advisable to do so, because this is a better method of formulating the knowledge situation than the formulation in terns of acceptance.[4]

## REFERENCES

1. Carnap, Rudolf. “The Aim of Inductive Logic,” in *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*, Ernest Nagel, Patrick Suppes, and Alfred Tarski, eds., pp. 303-318. Stanford, Calif.: Stanford University Press, 1962.

2. Carnap, Rudolf. *Logical Foundations of Probability*. Chicago: University of Chicago Press, 1950; second ed.,1962.

3. Carnap, Rudolf. “Replies and Systematic Expositions,” in [11], pp. 859-1013.

4. Carnap, Rudolf, and Yehoshua Bar-Hillel. *An Outline of a Theory of Semantic Information.* Cambridge, Mass.: Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952. Reprinted in Bar-Hillel, Yehoshua, *Language and Information: Selected Essays*. Reading, Mass.: Addison-Wesley, 1964.

5. Jeffrey, R. C. “Valuation and Acceptance of Scientific Hypotheses,” *Philosophy of Science,* 23:237-246 (1956).

6. Popper, Karl R. *Conjectures and Refutations*. New York: Basic Books, 1962.

7. Popper, Karl R. ‘“The Demarcation between Science and Metaphysics;’ in [11], pp. 183-226. Also in [6], pp. 253-292.

8. Popper, Karl R. *The Logic of Scientific Discovery*. New York: Basic Books, 1959.

9. Popper Karl R. “On the Sources of Knowledge and of Ignorance,” in [6], pp. 3-30.

10. Popper, Karl R. ‘“Truth, Rationality, and the Growth of Scientific Knowledge,” in [6] pp. 215-250.

11. Schilpp, Paul Arthur, ed. *The Philosophy of Rudolf Carnap*. La Salle, Ill.: Open Court, 1964.

[4] Elsewhere I have stated my reasons against the view that the end result of an act of inductive reasoning is the acceptance of a hypothesis (see the last paragraph of Sect. 25 in [3]; cf. Jeffrey [5]). And I have tried to show that the changed point of view makes it possible to give an answer to Hume’s skeptical objection against induction (see the last two paragraphs of my paper [1]).