Pluralism in Probabilistic Justification

David Atkinson and Jeanne Peijnenburg

1   Introduction

From Aristotle onwards, epistemic justification has been conceived as a form of inference. If a proposition E_n is epistemically justified by a proposition E_{n+1}, then according to the traditional view E_n is somehow inferred from E_{n+1}.

It took twenty-three centuries to modify this outlook. Today, many epistemologists construe epistemic justification in terms of probabilistic support rather than of inferential relations. In the modern view, E_n is epistemically justified by E_{n+1} if two requirements are fulfilled. First, E_{n+1} should probabilistically support E_n. By this we mean that the conditional probability of E_n, given E_{n+1}, exceeds the conditional probability of E_n, given not-E_{n+1}:

$$P(E_n|E_{n+1}) > P(E_n|\neg E_{n+1}) \, . \tag{1}$$

Second, the unconditional probability of $P(E_n)$ should not fall below some agreed threshold of acceptance.

This 'probabilistic turn' in epistemology opened the door to pluralism in epistemic justification. Imagine a sequence of propositions $E_0, E_1, E_2 \ldots$, such that E_0 is epistemically justified by E_1, which is epistemically justified by E_2, and so on. In 1956, when the probabilistic turn had not yet been fully made, Wilfrid Sellars still saw no more options than to construct this sequence as either a finite chain or a finite loop:

> "One seems forced to choose between the picture of an elephant which rests on a tortoise (What supports the tortoise?) and the picture of a great Hegelian serpent of knowledge with its tail in its mouth (Where does it begin?). Neither will do."[1]

Present-day epistemologists, however, are not confined to these two possibilities. Thanks to the interpretation of epistemic justification as probabilistic support, they have at their disposal many different ways of reconstructing justificationary processes. A target proposition, E_0, can be probabilistically justified by a chain or by a loop, and both the chain and the loop can be finite or infinite.[2] Moreover, in each of these four cases the conditional probabilities might be uniform, taking on the same values, or they might be nonuniform, differing throughout the chain or loop. This yields already eight different varieties of epistemic justification. In earlier papers we have discussed the four most intriguing ones — involving chains and loops of infinite

size.[3] We showed there that infinite chains and infinite loops can converge, yielding a unique and well-defined probability value for the target proposition E_0.

In the present paper we contribute to a pluralistic outlook by introducing even more possibilities for probabilistic justification. In contrast to the eight varieties above, all of which are one-dimensional, we will investigate probabilistic justification in more than one dimension. We shall concentrate again on structures of infinite size, and we show that many-dimensional networks can converge, too. Thus it makes sense to say that a target proposition, E_0, can be epistemically justified not only by an infinite one-dimensional chain or an infinite one-dimensional loop, but also by an infinite network of many dimensions.

We start, in Section 2, by recalling some facts about infinite, one-dimensional chains, where for convenience sake we restrict ourselves to chains that are uniform. In Section 3 we explain what happens when we go from a one-dimensional uniform chain to a two-dimensional uniform network. In Section 4, we contrast the properties of one-dimensional chains with those of two-dimensional networks. As we will see, there exists an intriguing difference between the two, which poses difficulties at an intuitive level. In Section 5 we indicate the relevance of this paper for disciplines outside epistemology and philosophy in general, by explaining an application of our analysis to genetics.

2 Infinite, uniform chains

Earlier we have shown that probabilistic epistemic chains of infinite length always converge. In the present section we summarize our findings, restricting ourselves for simplicity to uniform chains. However, the demonstration we give of convergence is different from earlier proofs, since we will now use fixed-point methods.[4]

The unconditional probabilities $P(E_n)$ and $P(E_{n+1})$ are related by the rule of total probability,

$$P(E_n) = P(E_n|E_{n+1})P(E_{n+1}) + P(E_n|\neg E_{n+1})P(\neg E_{n+1}). \tag{2}$$

As we have already indicated, we assume in this paper that the conditional probabilities are uniform, i.e. they are the same throughout the chain. However, it is important to keep in mind that the assumption of uniformity is not essential: the whole argument goes through without assuming uniformity, albeit in a somewhat more complicated form.

Under the assumption of uniformity, Eq.(2) may be rewritten in the form

$$P(E_n) = \beta + (\alpha - \beta) P(E_{n+1}), \tag{3}$$

where

$$\alpha = P(E_n|E_{n+1}) \quad \text{and} \quad \beta = P(E_n|\neg E_{n+1}). \tag{4}$$

Clearly $\alpha > \beta$ is equivalent to the condition of probabilistic support as expressed in (1).

Does the iteration (3) converge, giving a well-defined value for $P(E_0)$, $P(E_1)$, $P(E_2)$ and so on? If it does, then $P(E_n)$ and $P(E_{n+1})$ will have to be equal in the limit. Let us call this limiting value, if it exists, P_1^*. It is a fixed point of the iteration Eq.(3), i.e. it satisfies

$$P_1^* = \beta + (\alpha - \beta) P_1^*,$$

and this linear equation has the unique solution

$$P_1^* = \frac{\beta}{1 - \alpha + \beta},$$ (5)

where we exclude $\alpha = 1$ (the case in which E_{n+1} entails E_n).

To show that the iteration (3) does indeed converge, we write $P(E_{n-1}) = \beta + (\alpha - \beta) P(E_n)$, which is (3) with $n - 1$ in place of n. Subtracting this from Eq.(3), we obtain

$$P(E_n) - P(E_{n-1}) = (\alpha - \beta) \left[P(E_{n+1}) - P(E_n) \right].$$

Hence, by iteration,

$$
\begin{aligned}
P(E_n) - P(E_{n-1}) &= (\alpha - \beta)^2 \left[P(E_{n+2}) - P(E_{n+1}) \right] = \ldots \\
&= (\alpha - \beta)^s \left[P(E_{n+s}) - P(E_{n+s-1}) \right].
\end{aligned}
$$ (6)

We may take s to infinity on the right-hand side of (6), and since $(\alpha - \beta)^s$ tends to zero in this limit, it is clear that $P(E_n) - P(E_{n-1}) = 0$, i.e. $P(E_n) = P(E_{n-1})$ for all finite n. This shows that the iteration (3) converges. Indeed all the unconditional probabilities are equal to one another, and they are all equal to the fixed point, P_1^*.

3 Infinite, uniform networks

In this section we shall consider a two-dimensional probabilistic network, where a 'child' proposition is probabilistically justified by two 'parent' propositions, each of which is in turn probabilistically justified by two '(grand)parent' propositions, etc. So the network has a tree-like structure:
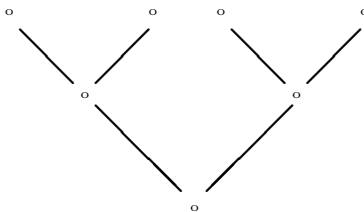

Figure 1: Justification in two dimensions

In the sequel, we shall often use 'child' and 'parent' for 'child proposition' and 'parent proposition'. We will talk about 'the probability that the child (parent) is true', or alternatively about 'the probability of the child (parent)'.

Much as in our treatment of the one-dimensional chain, we take it that a child proposition is justified by its parent propositions if two requirements are fulfilled. First, the parents must probabilistically support the child. By this we mean: the probability that the child is true,

given that both its parents are true, is larger than the probability that the child is true, given that both parent propositions are false.[5] In symbols:

$$P(E_n|E_{n+1}\&E'_{n+1}) > P(E_n|\neg E_{n+1}\&\neg E'_{n+1}), \tag{7}$$

where E_n stands for the child, E_{n+1} for the one parent, and E'_{n+1} for the other. We will refer to the conditional probability in which both parents are true by α and that in which both parents are false by β, so (7) becomes:

$$\alpha > \beta.$$

The second requirement for justification is that the unconditional probability of the child proposition may not lie below a certain threshold of acceptance. The threshold could be 0.5, or 0.9, or even higher, dependent on the context of the case.

For simplicity, we make three further assumptions. To begin with, we assume that the parents are independent of one another, so the probability that both parents are true equals the probability that one of them is true times the probability that the other one is also true:

$$P(E_{n+1}\&E'_{n+1}) = P(E_{n+1})P(E'_{n+1}). \tag{8}$$

Moreover, we explicitly assume gender symmetry, i.e. both parent propositions have the same probability of being true, so $P(E_{n+1}) = P(E'_{n+1})$. Finally, we suppose that the conditional probabilities are the same throughout the whole two-dimensional structure. In other words, just like the linear chain in the previous section, our quadratic two-parent network in the present section is uniform. However, as was the case for the one-dimensional chain, the uniformity assumption is not essential. Nor do we need the assumptions of independence or of gender symmetry. Our argument can be made without these three assumptions, although we shall not show that here.

The unconditional probability that the child is true, $P(E_n)$, can be written in terms of the triple joint probabilities associated with two parent propositions as follows:

$$P(E_n) = P(E_n\&E_{n+1}\&E'_{n+1}) + P(E_n\&E_{n+1}\&\neg E'_{n+1}) + P(E_n\&\neg E_{n+1}\&E'_{n+1}) +$$
$$P(E_n\&\neg E_{n+1}\&\neg E'_{n+1}). \tag{9}$$

The first term on the right-hand side is the joint probability that the child proposition and both parent propositions are all true. This term can be written as

$$P(E_n\&E_{n+1}\&E'_{n+1}) = \alpha P(E_{n+1}\&E'_{n+1}), \tag{10}$$

where α, as we said, is now the conditional probability that E_n is true, given that both parents are true. Since the parents are independent of one another, see (8), and since we supposed that each parent has the same probability of being true, Eq.(10) can be written as:

$$P(E_n\&E_{n+1}\&E'_{n+1}) = \alpha P^2(E_{n+1}). \tag{11}$$

The last term on the right-hand side of (9) is the joint probability that the child proposition is true, and both parent propositions are false. It can be written as

$$P(E_n\&\neg E_{n+1}\&\neg E'_{n+1}) = \beta P(\neg E_{n+1}\&\neg E'_{n+1})$$
$$= \beta P(\neg E_{n+1})P(\neg E'_{n+1}) = \beta\,[1 - P(E_{n+1})]^2, \tag{12}$$

where β is the conditional probability that E_n is true, given that both parents are false.

The second and third terms on the right-hand side of (9) are the joint probabilities that the child proposition is true, when one parent is true, and one is false. The terms can be written as

$$
\begin{aligned}
P(E_n \& E_{n+1} \& \neg E'_{n+1}) &= \gamma P(E_{n+1} \& \neg E'_{n+1}) \\
&= \gamma P(E_{n+1}) P(\neg E'_{n+1}) = \gamma P(E_{n+1})[1 - P(E_{n+1})], \qquad (13)
\end{aligned}
$$

$$
\begin{aligned}
P(E_n \& \neg E_{n+1} \& E'_{n+1}) &= \delta P(\neg E_{n+1} \& E'_{n+1}) \\
&= \delta P(\neg E_{n+1}) P(E'_{n+1}) = \delta P(E_{n+1})[1 - P(E_{n+1})], \qquad (14)
\end{aligned}
$$

where γ is the conditional probability that the child is true, given that only the first parent is true, and δ is the conditional probability that it is true, given that only the second parent is true.

On inserting the expressions (10),(12),(13), and (14) into (9), we find, after rearrangement,

$$
P(E_n) = \beta + 2(\varepsilon - \beta) P(E_{n+1}) + (\alpha + \beta - 2\varepsilon) P^2(E_{n+1}), \qquad (15)
$$

where ε is the average of the conditional probabilities that only one parent is true:

$$
\varepsilon = \tfrac{1}{2}(\gamma + \delta).
$$

In the special case that

$$
\alpha + \beta = 2\varepsilon, \qquad (16)
$$

this relation takes on the linear form

$$
P(E_n) = \beta + (\alpha - \beta) P(E_{n+1}),
$$

just like the one-dimensional chain (3). We know from Eq.(6) that this sequence converges if $\alpha > \beta$ (we exclude the case $\alpha = 1$).

When the special equality (16) does *not* hold, Eq.(15) has two fixed points, namely the two solutions of the quadratic equation

$$
P_2^* = \beta + 2(\varepsilon - \beta) P_2^* + (\alpha + \beta - 2\varepsilon) P_2^{*2}.
$$

We show in the appendix that only one of these fixed points is attracting, namely

$$
P_2^* = \frac{\beta + \tfrac{1}{2} - \varepsilon - \sqrt{\beta(1 - \alpha) + (\varepsilon - \tfrac{1}{2})^2}}{\alpha + \beta - 2\varepsilon} . \qquad (17)
$$

As in the one-dimensional case, all the unconditional probabilities are the same, being equal to the fixed point, P_2^*.

4 Contrasting Chains and Networks

It is interesting to compare the properties of the linear one-parent chain of Section 2 with those of the quadratic two-parent net of Section 3. In the present section we will discuss first a similarity between the two structures, and then an intriguing difference.

The similarity concerns sufficient conditions for convergence in the two cases. As we have seen, the requirement of probabilistic support for the chain is $P(E_n|E_{n+1}) > P(E_n|\neg E_{n+1})$, i.e. the child E_n is supported by its parent E_{n+1}. For the net, on the other hand, the condition is $P(E_n|E_{n+1}\&E'_{n+1}) > P(E_n|\neg E_{n+1}\&\neg E'_{n+1})$, i.e. the child E_n is supported by both of its parents E_{n+1} and E'_{n+1}. In both cases, the condition turns out to be sufficient for convergence: it ensures convergence of not only the one-dimensional infinite iteration, but also of the two-dimensional infinite net. Note that, for the net, the conditional probability that the child is true, given that only one parent is true, has no relevance to convergence. The iteration has an attracting fixed point if $\alpha > \beta$; in the two-dimensional case, it does not matter how large or small ε is.

There is, however, a crucial difference between the chain and the net. This difference pertains to the situation in which β is zero. For the chain, as we have seen, $\beta = P(E_n|\neg E_{n+1})$; and the infinite, uniform chain leads to the fixed point (5):

$$P_1^* = \frac{\beta}{1 - \alpha + \beta}\,,$$

which clearly vanishes if $\beta = 0$ (assuming $\alpha < 1$). Thus if the child of a parent proposition that is false is *never* true then, after an infinite number of generations, the child proposition will certainly be false. This should not come as a surprise. After all, here the probabilistic justification of the target proposition E_0 by an infinite chain E_1, $E_2 \ldots$, and so on, is such that only the conditional probability $\alpha = P(E_n|E_{n+1})$ is positive, the conditional probability $\beta = P(E_n|\neg E_{n+1})$ being zero. Consequently, Eq.(2) becomes $P(E_n) = P(E_n|E_{n+1})P(E_{n+1})$, and so each link of the infinite chain contributes to the monotonic diminution of the value of $P(E_0)$, resulting finally in zero.[6]

The situation in two dimensions is entirely different. Now $\beta = P(E_n|\neg E_{n+1}\&\neg E'_{n+1})$; and the infinite, uniform net leads to the fixed point (17):

$$P_2^* = \frac{\beta + \frac{1}{2} - \varepsilon - \sqrt{\beta(1 - \alpha) + (\varepsilon - \frac{1}{2})^2}}{\alpha + \beta - 2\varepsilon}\,.$$

In the case that β is zero, this formula reduces to

$$P_2^* = \frac{\frac{1}{2} - \varepsilon - |\frac{1}{2} - \varepsilon|}{\alpha - 2\varepsilon}\,. \tag{18}$$

Notice that this is zero *only in the case that $\varepsilon \leq \frac{1}{2}$*. When $\varepsilon > \frac{1}{2}$, the expression (18) becomes

$$P_2^* = \frac{2\varepsilon - 1}{2\varepsilon - \alpha}\,. \tag{19}$$

The interesting thing is that, if a child is false when both parents are false, then the unconditional probabilities $P(E_n)$ in the infinite net may, or may not be zero. It all depends on how probable it is that a child is true when *only one* of its parents is true. If this conditional probability is more than one half (that is, if the child is more likely to be true than false given that only one of the parents is true), then the unconditional probabilities $P(E_n)$ do not vanish. This is quite different from the one-dimensional situation, where $\beta = 0$ does imply that $P(E_n)$ vanishes. However, in order for the child to be justified, not only must ε be greater than one-half

when $\beta = 0$, but also α must be large enough to ensure that the unconditional probability that the child proposition is true does not lie below the threshold of acceptance. Precisely why $\varepsilon = \frac{1}{2}$ marks the boundary between a zero and a nonzero unconditional probability is intuitively still unclear to us.

5   Relevance and Applications

Do the above exercises have any applications? Are the formalisms that we have developed of any philosophical relevance or utility in the outside world?

As far as the relevance to philosophy is concerned, we can be brief. In the introduction we already alluded to the venerable tradition concerning justification in epistemology; in particular justification related to an infinite regress is a subject that has been much discussed.[7] Most philosophers in the tradition took the view that argumentation to infinite regress shows that a certain position is absurd. Philosophers ranging from Zeno, Plato and Aristotle to Aquinas, Descartes, Leibniz, Hume and Kant have all used a *regressus ad infinitum* as a *regressus ad absurdum*: in their view, any argument that leads to an infinite regress is thereby vitiated. In our paper we show that this view is mistaken if an infinite regress is probabilistic in nature. We have explained this first for a one-dimensional chain of propositions, and then for a two-dimensional network.

As to the applications in the outside world, they are numerous, especially in view of the fact that our simplifying assumptions (namely uniformity, and, in two dimensions, gender symmetry and independence) can be relaxed without affecting the essential findings. Here we will restrict ourselves to one application, taken from the genetics of a population in which background conditions remain stable over time.

Consider the inheritance of a gender-specific genetic disorder in a human population, such as the tendency to prostate cancer in the male, or to breast cancer in the female. The probability that a child will develop the condition at some time in its life is different if the parent of the same gender has the complaint, or if that parent does not. If the relevant external conditions remain the same over time, the two conditional probabilities, α and β, will be uniform, that is, the same from generation to generation. The one-dimensional formalism of Sect. 2 is then applicable, and we conclude that the probability of disease, which we can equate to the relative frequency of its incidence in a large population, after any transient effects have died out, is given by the fixed point (5):

$$P_1^* = \frac{\beta}{1 - \alpha + \beta}.$$

The values of α and β could be inferred from the statistics of two generations only, and one can then deduce the above relative frequency of incidence, which will be stable throughout the generations.

Our analysis of the two-dimensional case can be applied, for example, to the inheritance in a human population of albinism. The three conditional probabilities that a child will be normally coloured, namely α (if both parents are normally coloured), β (if neither parent is normally coloured, i.e. both are albinos), and ε (if one parent is normal and one is albino), can be estimated from the statistics of a large population. The relative frequency of normally

coloured individuals in a large population is then given by the fixed point (17):

$$P_2^* = \frac{\beta + \frac{1}{2} - \varepsilon - \sqrt{\beta(1-\alpha) + (\varepsilon - \frac{1}{2})^2}}{\alpha + \beta - 2\varepsilon}\,.$$

However, there is more to be said in this case.

To begin with, when two albinos mate, their children are nearly always albinos, i.e. they are almost never normally coloured, so $\beta = 0$ to a good approximation. Thus the situation obtains that we mentioned in Sect. 4.

Furthermore, our analysis of the two-dimensional case can be seen as a generalization of the famous Hardy-Weinberg rule in genetics.[8] The genetic fingerprint of every individual is given by his or her DNA, i.e. the double helix consisting of two strings of molecules called nucleotides. Sequences of nucleotides are grouped together to form genes. Many of these genes are such that they come in two possible forms or alleles, one of which stems from the father and one from the mother. Let us denote the one allele by the letter A, the other by a. So an individual's genetic make-up, as far as these particular genes are concerned, will be one of the following: AA, Aa, aA, aa. Now albinism arises from an allele, a, that is *recessive*: this indicates that only an individual with allele a in both strands of his or her DNA will be an albino. The allele A is called *dominant* because individuals carrying Aa, or aA are healthy, just like individuals carrying AA.

In a large population, suppose that the fraction of the recessive allele is q, and the fraction of the dominant allele is $p = 1 - q$. Then the albino fraction of the population, carrying aa, is q^2, while the healthy fraction that carries AA is p^2, and the healthy fraction that carries Aa or aA is $2pq$. Moreover, these fractions remain the same from generation to generation. This is the essence of the Hardy-Weinberg rule: it is based on the assumption of a theoretically infinite population and random mating.

In the Hardy-Weinberg model, it is possible to calculate the conditional probabilities in terms of q, the fraction of the recessive allele. We find

$$\alpha = 1 - \left(\frac{q}{1+q}\right)^2 \qquad \varepsilon = \frac{1}{1+q} \qquad \beta = 0\,, \tag{20}$$

where we suppress the details of the rather tedious calculation. The first thing to notice is that, if $0 < q < 1$, then ε is necessarily greater than one-half, and since $\beta = 0$ the formula (19) is applicable:

$$P_2^* = \frac{2\varepsilon - 1}{2\varepsilon - \alpha}\,.$$

With the values (20) we find $P_2^* = 1 - q^2$, which is clearly correct, since the albino fraction of the population, carrying two recessive alleles, aa, is q^2, and P_2^* is the complement of that.

When we take account of the fact that mutations from the recessive to the dominant allele are actually possible (very rarely), so that β is not quite zero, the general fixed-point formula (17) must be used instead of (19). This constitutes a modification of the Hardy-Weinberg model that can be readily handled by our methods.

Appendix

When the special equality (16) does *not* hold, the quadratic iteration (15) can be put into canonical form by means of the substitution

$$q_n = (\alpha + \beta - 2\varepsilon)\, P(E_n) - \beta + \varepsilon \,. \tag{21}$$

With this transformation, (15) becomes[9]

$$q_n = c + q_{n+1}^2 \,, \tag{22}$$

where

$$c = \varepsilon(1 - \varepsilon) - \beta(1 - \alpha) \,.$$

The conditional probabilities, α, β and ε are all real numbers in the unit interval, so it follows that $c \leq \varepsilon(1 - \varepsilon) = \frac{1}{4} - (\varepsilon - \frac{1}{2})^2 \leq \frac{1}{4}$. Further, since $\beta < \alpha$ (the condition of probabilistic support), $c \geq -\beta(1 - \alpha) > -\alpha(1 - \alpha) = -\frac{1}{4} + (\frac{1}{2} - \alpha)^2 \geq -\frac{1}{4}$. That is,

$$-\tfrac{1}{4} < c \leq \tfrac{1}{4} \,. \tag{23}$$

Consider the fixed point,

$$q^* = \tfrac{1}{2} - \sqrt{\tfrac{1}{4} - c} \,,$$

of the iteration (22). Via the inverse of the transformation Eq.(21), one can show that q^* corresponds to the fixed point P_2^* of Eq.(17).

To demonstrate that q^* is attracting, we change the variable from q_n to $s_n = q_n - q^*$, so that (22) becomes

$$s_n = s_{n+1} \left[1 - \sqrt{1 - 4c} + s_{n+1} \right] \,. \tag{24}$$

If

$$\left| 1 - \sqrt{1 - 4c} \right| < 1 \tag{25}$$

and s_{n+1} is very small, we conclude that c^* is attracting. Indeed, since

$$s_n - s_{n+1} = (s_{n+1} - s_{n+2}) \left[1 - \sqrt{1 - 4c} + s_{n+1} + s_{n+2} \right]$$

the mapping (24) is a contraction if $|s_n| \leq \rho$ and $\left| 1 - \sqrt{1 - 4c} + 2\rho \right| < 1$. Hence if $|s_N| \leq \rho$ for very large N, and ρ satisfies the above contraction constraint, the iteration backwards to s_0 will be attracted to zero, that is to say q_0 will be attracted to q^*. The domain of attraction of the fixed point, q^*, is $-\frac{3}{4} < c < \frac{1}{4}$. Attraction is trivially guaranteed also when $c = \frac{1}{4}$. So absolute convergence is assured if $|c| \leq \frac{1}{4}$; and we see from the inequalities (23) that this is consistent with the requirements $0 \leq \beta < \alpha < 1$ and $0 \leq \varepsilon \leq 1$.

Notes

[1]Wilfrid Sellars, "Empiricism and the Philosophy of Mind", in: Herbert Feigl/Michael Scriven (Eds.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis.* Minneapolis: University of Minnesota Press 1966, pp.253-329; p.300.

[2]The concept of an infinite loop might seem incoherent, but it is not. As we have shown elsewhere, an infinite loop differs in nontrivial ways from an infinite chain. See David Atkinson/Jeanne Peijnenburg, "Justification by Infinite Loops", in: *Notre Dame Journal of Formal Logic*, 51, 4, 2010, pp.407-416.

[3]Jeanne Peijnenburg/David Atkinson, "Probabilistic Justification and the Regress Problem", in: *Studia Logica*, 89, 3, 2008, pp.187-201; David Atkinson/Jeanne Peijnenburg, "Justification by an Infinity of Conditional Probabilities", in: *Notre Dame Journal of Formal Logic*, 50, 2, 2009, pp.183-193; David Atkinson/Jeanne Peijnenburg, "Justification by Infinite Loops", *loc. cit.*

[4]Of the earlier proofs, the most general one can be found in Appendix A of David Atkinson/Jeanne Peijnenburg, "The Solvability of Probabilistic Regresses: A Reply to Frederik Herzberg", in: *Studia Logica*, 94, 3, 2010, pp.347-353.

[5]Below we will consider the case in which one parent proposition is true while the other one is false. We will see that this situation has no bearing on the condition of probabilistic support.

[6]This result on the one-dimensional infinite chain may shed light on the position of C.I. Lewis and Bertrand Russell, who both claimed that an infinite chain of probabilistic relations must lead to probability zero for the target proposition. This claim is incorrect as a general statement, as Hans Reichenbach pointed out, precisely because Lewis and Russell had forgotten the term in the rule of total probability, which corresponds to the probability that the child proposition is true when its parent is false (see Eq.(2) in the main text). However, as we can see now, Lewis and Russell were right in a very restricted situation. If in one dimension the probability of a child, given the falsity of its parent proposition, is always zero, then the probability of the target proposition will progressively decrease from link to link, tending to zero in the limit of an infinite chain. See C.I. Lewis, "The Given Element in Empirical Knowledge", in: *The Philosophical Review*, 61, 2, 1952, pp.168-172; Hans Reichenbach, "Are Phenomenal Reports Absolutely Certain?", in: *The Philosophical Review*, 61, 2, 1952, pp.147-159; Bertrand Russell, *Human Knowledge: Its Scope and Limits.* London: George Allen and Unwin 1948; Jeanne Peijnenburg/David Atkinson, "Grounds and Limits: Reichenbach and Foundationalist Epistemology", forthcoming in *Synthese*, published online 14 July 2009, Open Access, http://www.springerlink.com/content/625275w685343723/fulltext.pdf; Jeanne Peijnenburg, "Ineffectual Foundations", in: *Mind*, 119, 2010, pp.1125-1133.

[7]Cf. Bonjour: "Considerations with respect to the regress argument [are] perhaps the most crucial in the entire theory of knowledge", Laurence Bonjour, *The Structure of Empirical Knowledge.* Cambridge (Mass.): Harvard University Press, 1985; p. 18.

[8]G.H. Hardy, "Mendelian Proportions in a Mixed Population", in: *Science,* 28, 1908, pp.49-50. W. Weinberg, "Über den Nachweis der Vererbung beim Menschen", in: *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg,* 64, 1908, pp.368-382.

[9]Elsewhere we propose to trace the connection between the two-dimensional net and the Mandelbrot fractal. (Benoît Mandelbrot, *The Fractal Geometry of Nature.* New York: W.H. Freeman and Co 1982 – second printing with update).