Chapter 3
The Probabilistic Regress

Abstract
During more than twenty years Clarence Irving Lewis and Hans Reichenbach pursued an unresolved debate that is relevant to the question of whether infinite epistemic chains make sense. Lewis, the nay-sayer, held that any probability statement presupposes a certainty, but Reichenbach profoundly disagreed. We present an example of a benign probabilistic regress, thus showing that Reichenbach was right. While in general one lacks a criterion for distinguishing a benign from a vicious regress, in the case of probabilistic regresses the watershed can be precisely delineated. The vast majority ('the usual class') is benign, while its complement ('the exceptional class') is vicious.

3.1 A New Twist

The previous chapter indicated how intricate the debate about epistemic justification has become. A mixed bag of knotty details and drawbacks complicates the subject, giving rise to a variety of different positions. But although nobody knows what exactly epistemic justification is, the idea that it involves probabilistic support is widespread among epistemologists of all sorts and conditions. Internalists, externalists, foundationalists, anti-foundationalists, evidentialists and reliabilists: most of them assume that 'A_j justifies A_i' implies that A_i somehow receives probabilistic support from A_j.

In this chapter and the ones to follow we want to make clear how significant this turn towards probability actually is, and what surprising consequences it has. The debate about epistemic regresses acquires a completely

D. Atkinson, J. Peijnenburg, *Fading Foundations*, Synthese Library 383,
DOI 10.1007/978-3-319-58295-5_3

new twist when Kolmogorovian probability is brought into the picture; for as we will see a probabilistic regress turns out to be immune to many of the objections that have routinely been raised against the traditional regress of entailments. The situation is to a certain extent reminiscent of the two causal regresses that we encountered in Chapter 1. Whereas a causal series *per se* only makes sense if it has a first member, this is not so for a causal series *per accidens*. Similarly, as we will argue, a traditional regress of entailments needs a first member, but a regress of probabilistic support may not.

In the present chapter we will describe the concept of a probabilistic regress, that is a regress in which (1.1) of Chapter 1,

$$q \longleftarrow A_1 \longleftarrow A_2 \longleftarrow A_3 \longleftarrow A_4 \ldots$$

is reinterpreted as: q is probabilistically supported by A_1, which is probabilistically supported by A_2, and so on, *ad infinitum*.[1] It is assumed that every link in this chain satisfies the condition of probabilistic support (2.1). As we have seen, this condition is quite weak, falling considerably short of the title 'justification'. But for our purposes this minimal requirement is enough.

Our exposition of a probabilistic regress takes as its starting point a historical debate between Hans Reichenbach (1891-1953) and Clarence Irving Lewis (1883-1964). Lewis and Reichenbach are both early defenders of the view that epistemic justification is probabilistic in character, holding that A_j might justify A_i even if the former does not logically entail the latter but only provides probabilistic support. They disagree, however, as to the implications of this claim. Lewis insists that probabilistic justification must spring from a ground that is certain, whereas Reichenbach maintains that probabilistic justification remains coherent, even if it is not rooted in firm ground. The disagreement between Lewis and Reichenbach extended over more than two decades, from 1930 until 1952, and it is well documented in letters and in journal contributions.

In Sections 3.2 and 3.3 we will give an overview of the dispute. We first describe Lewis's main claim, viz. that any proposition of the form 'q is probable' or 'q is made probable by A_1' must presuppose a proposition that is certain. Lewis's argument for this claim is that without such a presupposition we will end up with a probabilistic regress that has the absurd consequence of always yielding probability value zero for q. Next we describe Reichenbach's objection to this argument. We then explain that Lewis is not convinced by it and challenges Reichenbach to produce a counterexample,

[1] The term 'probabilistic regress' was coined by Frederik Herzberg (Herzberg 2010).

i.e. a probabilistic regress that yields a number other than zero for the target proposition q.

Reichenbach never took up Lewis's challenge, but we will meet it in Section 3.4. By presenting a probabilistic regress that converges to a non-zero limit, we demonstrate that a target can have a definite and computable value, even if it is probabilistically justified by a series that continues *ad infinitum*. In this manner we show that Reichenbach rather than Lewis was correct, and also that a probabilistic regress can make sense.

The counterexample to Lewis in Section 3.4 has a simple, uniform structure. In Section 3.5 we offer a nonuniform and thus more general counterexample. Both counterexamples belong to what we call 'the usual class', i.e. the class of probabilistic regresses that yield a well-defined probability for the target proposition. We distinguish it from 'the exceptional class', which contains the probabilistic regresses that are not well-defined. In Section 3.6 we will spell out the conditions for membership of the usual and the exceptional classes. As it turns out, exceptional probabilistic regresses are characterized by the fact that here probabilistic support comes very close to entailment. Not surprisingly, therefore, probabilistic regresses in the exceptional class need a ground in order to bestow a value on the target, and in that sense count as vicious.

The uniform and the nonuniform counterexamples in 3.4 and 3.5 are rather abstract in nature; but in Section 3.7 we offer two real-life probabilistic regresses, based on the development of bacteria.

3.2 The Lewis-Reichenbach Dispute

In 1929 Lewis published his first major work, *Mind and the world order. An outline of a theory of knowledge*.[2] Here he starts from the traditional view that our knowledge is partly mathematical and partly empirical. The mathematical part deals with knowledge that is *a priori* and analytic; the empirical part concerns our knowledge of nature. This knowledge of nature, says Lewis, is always only probable:

> ...all empirical knowledge is probable only ...our knowledge of nature is a knowledge of probabilities.[3]

[2] The present section is based on Peijnenburg and Atkinson 2011.
[3] Lewis 1929, 309-310.

Since the crucial issue for any theory of knowledge is the character of empirical knowledge, it follows that

> ...the problem of our knowledge ...is that of the validity of our probability judgements.[4]

What about the validity of probability statements? In *Mind and the world order,* Lewis stresses time and time again that probability judgements only make sense if they are based on something that is certain:

> The validity of probability judgements rests upon ...truths which must be certain.[5]

> ...the immediate premises are, very likely, themselves only probable, and perhaps in turn based upon premises only probable. Unless this backward-leading chain comes to rest finally in certainty, no probability-judgment can be valid at all.[6]

Lewis is not the only philosopher who has argued that probability judgements presuppose certainties. The idea can already be found in David Hume's *Treatise of human nature* and it has also been defended by, among others, Keith Lehrer, Richard Fumerton, and Nicholas Rescher.[7] Lewis is however one of the few who discusses the claim in more detail. His explanation can be summarized as follows.

A statement of the form 'q is probable' or 'the probability of q is x' is in fact elliptical for 'q is probable, given A_1', or 'the probability of q given A_1 is x', where x is a number between one and zero. In symbols: the unconditional $P(q) = x$ is elliptical for the conditional $P(q|A_1) = x$. In many cases, A_1 is itself only probable, so we obtain 'A_1 is probable', which is shorthand for 'A_1 is probable, given A_2'. Again, if A_2 is only probable, we need A_3, et cetera. A probabilistic regress threatens. Lewis's claim is that in the end we must encounter a statement, p, that is certain (or has probability 1 — we will not distinguish here between these two cases):

$$q \longleftarrow A_1 \longleftarrow A_2 \longleftarrow A_3 \longleftarrow A_4 \longleftarrow \ldots \longleftarrow p.$$

Denying that this is so, and claiming that such a certain p is not needed, says Lewis, amounts to making nonsense of the original statement ('q is

[4] Ibid., 308.
[5] Ibid., 311.
[6] Ibid., 328-329.
[7] Hume 1738/1961, 178; Lehrer 1974, 143; Fumerton 2004, 162; Fumerton and Hasan 2010; Rescher 2010, 36-37.

probable') itself. Thus we can only give a probability value to a target, q, if we suppose that there is a ground or foundation, p, that is certain.[8].

Reichenbach read *Mind and the world order* soon after it came out. Although he concurred with many of Lewis's reasonings, he profoundly disagreed with the claim that probability statements only make sense if they are based on certainties. On July 29, 1930, he sent Lewis a letter, enclosing some of his own manuscripts. Unfortunately this letter is now lost. We only know of its existence from a reply that Lewis wrote to Reichenbach, dated August 26, 1930.[9] We are unable to infer from this reply what exactly Reichenbach had written, since Lewis mainly writes about the manuscripts that Reichenbach had sent him.[10]

Between 1930 and 1940 a correspondence developed, which was partly about practical matters (Reichenbach had fled Berlin in 1933 and went to Istanbul, from where he tried to find an academic position in the U.S.A.), and partly about Lewis's claim that probability judgements presuppose certainties. As far as the latter is concerned, it is clear that Reichenbach's arguments did not convince Lewis, for sixteen years later, in his book *An analysis of knowledge and valuation*, Lewis stresses the same point again:

> If anything is to be probable, then something must be certain. The data which themselves support a genuine probability, must themselves be certainties.[11]

The disagreement between Lewis and Reichenbach reached its height in December 1951, at the forty-eighth meeting of the Eastern Division of the American Philosophical Association at Bryn Mawr. At that meeting there was a symposium on 'The Given', where Lewis, Reichenbach and Nelson Goodman read papers. Their contributions were published a year later in *The Philosophical Review*, and there we learn that Lewis sticks to his guns:

---

[8] As James Van Cleve has noted, Lewis's text appears to be ambiguous between two readings (Van Cleve 1977, 323-324). According to the first, Lewis says something like: 'The probability of q given p is x, and moreover p is certain'. In symbols: $P(q|p) = x$ and $P(p) = 1$. According to the second reading he says: 'It is certain that the probability of q given p is x', that is $P(P(q|p) = x) = 1$. It can however be proven that the two readings are equivalent, so this ambiguity is merely apparent. We will come back to this matter in Chapter 7.

[9] "Your very kind letter of July 29th has reached me, here at my summer address." The summer address was, by the way, Briar Hill in New Hampshire, close to Vermont.

[10] And apparently did not know quite what to do with them: "I find difficulty in understanding the ground from which they arise."

[11] Lewis 1946, 186.

> The supposition that the probability of anything whatever always depends on something else which is only probable itself, is flatly incompatible with the assignment of any probability at all.[12]

But Reichenbach, too, insisted on his own views. Already in his major epistemological work, *Experience and prediction*, he had found an apt metaphor for his anti-foundationalist position:

> All we have is an elastic net of probability relations, floating in open space.[13]

Fifteen years later Reichenbach still had the same conviction. He calls the claim of Lewis that probabilities must be grounded in certainties "just one of those fallacies in which probability theory is so rich".[14] In an attempt to understand the root of the fallacy he writes:

> We argue: if events are merely probable, the statement about their probability must be certain, because ... Because of what? I think there is tacitly a conception involved according to which knowledge is to be identified with certainty, and probable knowledge appears tolerable only if it is embedded in a framework of certainty. This is a remnant of rationalism.[15]

And being a rationalist would of course be a thorn in Reichenbach's logical-empiricist side. Lewis, in turn, rejects the accusation of being an old fashioned rationalist and replies that, on the contrary, he is trying to save empiricism from what he calls 'a modernized coherence theory' like that of his opponent. He writes:

> ...the probabilistic conception [of Reichenbach] strikes me as supposing that if enough probabilities can be got to lean against one another they can all be made to stand up. I suggest that, on the contrary, unless some of them can stand up alone, they will all fall flat.[16]

Who is right in this debate? Some authors, such as James Van Cleve and Richard Legum, have argued that it is Lewis.[17] To explain why we dissent, we will first spell out the argument that Lewis puts forward in support of his claim that probability judgements presuppose certainties. It is true that the negation of Lewis's claim leads to an infinite regress, but since not all regresses are vicious, an argument is required in order to show that this particular regress is of the unacceptable kind.

---

[12] Lewis 1952, 173.
[13] Reichenbach 1938, 192.
[14] Reichenbach 1952, 152.
[15] Ibid.
[16] Lewis 1952, 173.
[17] Van Cleve 1977; Legum 1980.

3.3 Lewis's Argument

As Mark Pastin correctly notes, the claim that probabilities presuppose certainties was repeated by Lewis "throughout his writings but [he] gave little attention to defending it".[18] The most extensive defence can be found in *Mind and the world order*, which contains the following argument:

> Nearly all the accepted probabilities rest upon more complex evidence than the usual formulations suggest; what are accepted as premises are themselves not certain but highly probable. Thus our probability judgement, if made explicit, would take the form: the probability that *A* is *B* is a/b, because if *C* is *D*, then the probability that *A* is *B* is m/n, and the probability of '*C* is *D*' is c/d (where $m/n \times c/d = a/b$). But this compound character of probable judgement offers no theoretical difficulty for their validity, provided only that the probability of the premises, when pushed back to what is more and more ultimate, somewhere comes to rest in something certain.[19]

In other words, Lewis says that the judgement

$$A\ is\ B \quad \text{is probable,} \tag{3.1}$$

is elliptical for

$$A\ is\ B \quad \text{is probable, given} \ \ C\ is\ D. \tag{3.2}$$

Since we are dealing with empirical knowledge, *C is D* is itself also only probable. The judgement '*C is D* is probable' is in turn elliptical for '*C is D* is probable, given *E is F*'. And so on.

We can formalize and quantify (3.1) and (3.2) by

$$P(A\ is\ B) = a/b \tag{3.3}$$

which is elliptical for

$$\begin{aligned} P(A\ is\ B) &= P(A\ is\ B|C\ is\ D) \times P(C\ is\ D) \\ &= m/n \times c/d \\ &= a/b, \end{aligned} \tag{3.4}$$

where a/b, m/n and c/d are probability values between 1 and 0. Now of course the probability that *C is D* may also be elliptical. If this series were to

[18] Pastin 1975, 410.

[19] Lewis 1929, 327-28. Here '*A* is *B*' means something like 'all *A*-things are *B*-things'. We have replaced Lewis's '*P is Q*' and '*p/q*' by '*C is D*' and '*c/d*'.

go on and on, then, because all the factors in the multiplication are probabilities and thus positive numbers less than one, the probability of the original proposition *A is B* would always tend to zero. But this is ridiculous, so the series of probability judgements must come to a stop in a statement that is certain. This is Lewis's argument for his claim that bestowing a probability value on a target presupposes the acceptance of a ground that is certain: without such a ground, the probability of the target will go to zero.

Lewis's argument is however simply mistaken. For $P(A \textit{ is } B)$ is not elliptical for the product $P(A \textit{ is } B|C \textit{ is } D) \times P(C \textit{ is } D)$, but for the following sum of products:

$$P(A \textit{ is } B) = P(A \textit{ is } B|C \textit{ is } D) \times P(C \textit{ is } D) \\ + P(A \textit{ is } B|\neg(C \textit{ is } D)) \times P(\neg(C \textit{ is } D)). \tag{3.5}$$

The first term of (3.5) coincides with (3.4), but (3.5) contains a second term, which Lewis forgets. He ignores the fact that, if the probability of *A* is *B* is conditioned by the probability of *C is D*, then you can only calculate the former probability if you also take into account what that probability is in case *C is D* is false.[20] Eq.(3.5) is an instance of the rule of total probability, which is a theorem of the calculus that Andrey Kolmogorov developed in his *Grundbegriffe der Wahrscheinlichkeitsrechnung*.

Kolmogorov published his *Grundbegriffe* in 1933, which might explain Lewis's mistake. The same can however not be said of Bertrand Russell. In 1948, nineteen years after *Mind and the world-order*, Russell published *Human knowledge: its scope and limits*. Part 5 of this book is devoted to the concept of probability, and there Russell criticizes several theories of probability, including Reichenbach's theory in his *Wahrscheinlichkeitslehre* of 1935. It is interesting that, quite independently of Lewis (for he does not mention him anywhere), Russell claims that attributing a probability value to a proposition presupposes a certainty. Moreover, he defends this claim with the same erroneous argument that Lewis had used. Russell writes:

> At the first level, we say that the probability that an *A* will be a *B* is m_1/n_1; at the second level, we assign to this statement a probability m_2/n_2, by making it one of some series of similar statements; at the third level, we assign a

[20] Mark Pastin seems to interpret Lewis as talking about the probability of the *conjunction* of the propositions '*A is B*' and '*C is D*' (Pastin 1975, 413). In this reading, Eq.(3.5) would be replaced by $P\big((A \textit{ is } B) \text{ and } (C \textit{ is } D)\big) = P(A \textit{ is } B|C \textit{ is } D) \times P(C \textit{ is } D)$, and this expression has no second term. However in this case there would not be a justificatory chain in which one proposition justifies the other. See footnote 31.

> probability m_3/n_3 to the statement that there is a probability m_2/n_2 in favour of our first probability m_1/n_1; and so we go on forever. If this endless regress could be carried out, the ultimate probability in favour of the rightness of our initial estimate m_1/n_1 would be an infinite product
>
> $$m_2/n_2 \times m_3/n_3 \times m_4/n_4 \ldots$$
>
> which *may be expected to be zero*.[21]

In other words, Russell argues that a series of statements like

$s_1 = A$ is B
$s_2 =$ The probability of s_1 is m_1/n_1
$s_3 =$ The probability of s_2 is m_2/n_2
...

implies that the probability of s_1 will tend to zero.[22] The argument is the same as that of Lewis: the probability of s_1 is the outcome of the multiplication of an infinite number of factors each of which is smaller than 1. It thus fails for precisely the same reason as does Lewis's argument. If a proposition

---

[21] Russell 1948, 434; our italics. Where Russell has α and β we have used A and B. It is assumed that $0 < m_i/n_i < 1$ for all i. Presumably Russell, a competent mathematician, wrote 'may be expected to be zero' because he knew that there exist infinite products of factors, all less than one, that converge (i.e. that yield well-defined, non-zero values). In this connection it is interesting that Quine, in his 1946 Lectures on David Hume's Philosophy (Quine 2008), indeed makes the point that such a product can be convergent: in fact he gives an explicit example. He fails, however, to note that the point is irrelevant, for the probabilities in question should not be multiplied together (because of the second term in (3.5)). Thanks to Sander Verhaegh for bringing Quine's lectures to our notice. We return to Quine's reasoning in Chapter 7.

[22] Note that Russell here speaks about higher-order probability statements rather than about the probability of a reference class in a conditional probability statement (see footnote 8 for the difference). Russell says that such a series of higher-order probability statements "leads (one is to suppose) to a limit-proposition, which alone we have a right to assert. But *I do not see how this limit-proposition is to be expressed*. The trouble is that, as regards all the members of the series before it, we have no reason ... to regard them as more likely to be true than to be false; they have, in fact, no probability that we can estimate." (Russell 1948, 435; our italics). In other words, Russell suggests that we cannot attribute a probability value to s_1 because we are unable to compute the limit of the series. This seems to be at odds with his earlier claim that the value of s_1 goes to zero, but we will not dwell on the matter here. In the next section we will rather specify the limit proposition that Russell was vainly trying to express.

with probability x is conditioned by a proposition with probability y, then the probability of the first proposition is not given by xy, as Russell says, but by $xy + x'(1-y)$, where x' is the probability that the first proposition is true if the second is false, and $(1-y)$ is the probability that the second proposition is indeed false. Just like Lewis, Russell forgets the second term in the rule of total probability, namely $x'(1-y)$.

Reichenbach notices that Russell makes the mistake, and points it out to him in a letter of March 28, 1949.[23] Russell clearly acknowledges his oversight, as we see from his reply three weeks later.[24] Lewis, on the other hand, seems to have persisted in his error, and Reichenbach confronts him with this fact in 1951, at the forty-eighth meeting of the American Philosophical Association at Bryn Mawr. Lewis appears however not to be impressed by Reichenbach's amendment:

> ...even if we accept the correction which Reichenbach urges here, I disbelieve that it will save his point. For that, I think he must prove that, where any regress of probability-values is involved, the progressively qualified fraction measuring the probability of the quaesitum will converge to some determinable value other than zero; and I question whether such a proof can be given.[25]

In other words, Lewis fails to see the relevance of the second term in (3.5): he simply does not believe that an infinite regress of probabilities can converge to some value other than zero. Even if we *do* take Reichenbach's amendment into account, Lewis still thinks that an infinite series of probability statements conditioned by probability statements will always converge to zero. And he defies Reichenbach to prove the contrary. As far as we know Reichenbach never took up the challenge. Perhaps he planned to, but never got around to it; or maybe he had difficulties finding what Russell called "the limit proposition" (see footnote 22); or perhaps he simply got tired of the debate. We will presumably never know, for in April 1953 Reichenbach died in California of a heart attack.

---

[23] The letter is printed in the volume with selected writings of Hans Reichenbach edited by Maria Reichenbach and Robert Cohen (Reichenbach and Cohen 1978, 405-411).

[24] "I perceive already that you are right as to the mathematical error that I committed on page 416" (letter from Russell to Reichenbach, April 22, 1949). Page 416 corresponds to page 434 in reprints of Russell's book. We are grateful to Mr. L. Lugar and Ms. B. Arden of the Pittsburgh Archive for sending us a copy of Russell's letter.

[25] Lewis 1952, 172.

In the next section we will take up Lewis's gauntlet by presenting a counterexample to his argument that a "regress of probability-values" always tends to zero. This counterexample involves an infinite iteration of the rule of total probability. Although this iteration produces a much more complicated regress than the simple product that Russell and Lewis had envisaged, it leads to a perfectly well-defined, and moreover nonzero probability for the target proposition. It thus also produces the "limit-proposition" that Russell was looking for.[26]

3.4 A Counterexample

Let our target proposition q be probabilistically justified by proposition A_1. We have seen that the unconditional probability of q, namely $P(q)$, can be calculated from the rule of total probability:

$$P(q) = P(q|A_1)P(A_1) + P(q|\neg A_1)P(\neg A_1). \tag{3.6}$$

To make contact with Lewis's argument, we can take q to be '*A is B*' and A_1 to be '*C is D*'. If A_1 is probabilistically justified by A_2, then $P(A_1)$ can be calculated from another instance of the rule,

$$P(A_1) = P(A_1|A_2)P(A_2) + P(A_1|\neg A_2)P(\neg A_2), \tag{3.7}$$

and if A_2 is in turn probabilistically justified by A_3 we have to repeat the rule again,

[26] Dennis Dieks put forward the possibility that Lewis might have been interested only in those probabilistic regresses in which the second term may be legitimately ignored (Dieks 2015). Dieks' suggestion is intriguing, but it causes difficulties. First, why did not Lewis make this explicit? In his debate with Reichenbach there appear to have been opportunities enough. Second, even if A_{n+1} has been called a reason for A_n, we should not overlook the fact that other propositions, contained in the negation of A_{n+1}, can well contribute to the justification of A_n. As Johan van Benthem phrases it: "[$P(A_n|\neg A_{n+1})$] measures intuitively the 'bonus' that A_n receives even if A_{n+1} were untrue. This inclusion might perhaps sound odd if we have just introduced A_{n+1} as reason for A_n — but we may, neither here nor in argumentation generally, ignore the fact that a postulated claim can already enjoy support without A_{n+1}" (Van Benthem 2015, 148, our translation from the Dutch; *cf.* Peijnenburg 2015, 205-206). In any case, if Dieks were correct this would considerably restrict the domain in which the Lewisian approach could apply, and it would appear to be inconsistent with the probability calculus.

$$P(A_2) = P(A_2|A_3)P(A_3) + P(A_2|\neg A_3)P(\neg A_3). \tag{3.8}$$

Can we continue this repetition, thus allowing for propositions being probabilistically justified by other propositions, being probabilistically justified by still other propositions, *ad infinitum*? It might look as though we cannot. How would we ever be able to calculate $P(q)$ if it is the outcome of an infinite regress of instances of the rule of total probability? The calculation seems at first sight to be too lengthy and too complicated for us to complete. After all, insertion of Eq.(3.7), together with

$$P(\neg A_1) = P(\neg A_1|A_2)P(A_2) + P(\neg A_1|\neg A_2)P(\neg A_2) \tag{3.9}$$

into the right-hand side of Eq.(3.6) leads to an expression with four terms, namely:

$$\begin{aligned} P(q) = {} & P(q|A_1)P(A_1|A_2)P(A_2) + P(q|\neg A_1)P(\neg A_1|A_2)P(A_2) + \\ & P(q|A_1)P(A_1|\neg A_2)P(\neg A_2) + P(q|\neg A_1)P(\neg A_1|\neg A_2)P(\neg A_2). \end{aligned} \tag{3.10}$$

A repetition of this manoeuvre to express $P(A_2)$ and $P(\neg A_2)$ in terms of $P(A_3)$ and $P(\neg A_3)$ would produce no less than eight terms. After $n+1$ steps, the number of steps is 2^{n+1}, yielding an ungainly expression that seems hard to evaluate in a simple, closed form.

There is however a way to reduce this complication of the rapidly increasing number of terms. In explaining this we first simplify the notation by abbreviating (3.6) by setting the two conditional probabilities, $P(q|A_1)$ and $P(q|\neg A_1)$, equal to α and β:

$$\alpha = P(q|A_1) \qquad \beta = P(q|\neg A_1). \tag{3.11}$$

Now $P(q)$ becomes:

$$\begin{aligned} P(q) &= \alpha P(A_1) + \beta P(\neg A_1) \\ &= \alpha P(A_1) + \beta[1 - P(A_1)] \\ &= \beta + (\alpha - \beta)P(A_1). \end{aligned} \tag{3.12}$$

Clearly, we can only compute $P(q)$ if we know $P(A_1)$. Of course, we also have to know the values of the conditional probabilities α and β. Their status is however rather different from that of the unconditional probabilities, and we will come back to this matter in detail in Chapter 4. At this juncture, we simply assume that α and β are given, and that they are the same from link to link (the latter assumption is dropped in the next section). But what *is* the value of $P(A_1)$? We do not know. However, we do know that A_1 is probabilistically justified by A_2, and so we can calculate $P(A_1)$ in terms of $P(A_2)$, and so on:

$$\begin{aligned}P(A_1) &= \beta + (\alpha - \beta)P(A_2)\\ P(A_2) &= \beta + (\alpha - \beta)P(A_3)\\ P(A_3) &= \beta + (\alpha - \beta)P(A_4)\,.\end{aligned}$$

We can now see how to get rid of the unknown unconditional probabilities, namely by nesting the formulas. Thus we can remove $P(A_1)$ by substituting its value into (3.12), so that we obtain:

$$\begin{aligned}P(q) &= \beta + (\alpha - \beta)P(A_1)\\ &= \beta + (\alpha - \beta)\big[\beta + (\alpha - \beta)P(A_2)\big]\\ &= \beta + \beta(\alpha - \beta) + (\alpha - \beta)^2 P(A_2)\,. \end{aligned} \tag{3.13}$$

Next, by inserting the value of $P(A_2)$ into (3.13) we attain

$$\begin{aligned}P(q) &= \beta + \beta(\alpha - \beta) + (\alpha - \beta)^2\big[\beta + (\alpha - \beta)P(A_3)\big]\\ &= \beta + \beta(\alpha - \beta) + \beta(\alpha - \beta)^2 + (\alpha - \beta)^3 P(A_3)\,, \end{aligned} \tag{3.14}$$

by which we got rid of $P(A_2)$. And so on. After a finite number m of steps we obtain the following formula:

$$P(q) = \beta + \beta(\alpha - \beta) + \beta(\alpha - \beta)^2 + \ldots + \beta(\alpha - \beta)^m + (\alpha - \beta)^{m+1}P(A_{m+1})\,. \tag{3.15}$$

Eq.(3.15) is the beginning of the “regress of probability-values” that Lewis is talking about. His argument is that, if this series is continued *ad infinitum*, $P(q)$ will always tend to zero, notwithstanding the fact that Reichenbach’s correction has been taken into account. This is presumably why Lewis comments: “I disbelieve that it [the addition of the second term] will save his point.” Let us see whether Lewis’s disbelief is justified.

There are two things that should be noted about (3.15). The first is that it contains only one factor of which the value is unknown. This is $P(A_{m+1})$, i.e. the probability of the first proposition, A_{m+1}, in this finite series. Since all the probabilities in the series are ultimately computed on the basis of this unconditional probability, it seems that we must know its value in order to be able to calculate $P(q)$. The second thing is that, as m gets bigger and bigger, so that the justificatory chain becomes longer and longer, $(\alpha - \beta)^{m+1}$ gets smaller and smaller without limit, finally converging to zero. But of course, if $(\alpha - \beta)^{m+1}$ converges to zero, then $(\alpha - \beta)^{m+1}P(A_{n+1})$ dwindles away to nothing too, for $P(A_{m+1})$ cannot be greater than 1. The right-hand side of Eq.(3.15) is a sum, and if a term in a sum goes to zero, it does not contribute in the limit. With an infinite number of steps, the terms that remain are

$$\begin{aligned} P(q) &= \beta + \beta(\alpha-\beta) + \beta(\alpha-\beta)^2 + \ldots \\ &= \beta\left[1 + (\alpha-\beta) + (\alpha-\beta)^2 + \ldots\right] \\ &= \beta \sum_{n=0}^{\infty} (\alpha-\beta)^n . \end{aligned} \tag{3.16}$$

Since $\alpha - \beta$ is less than one, the sum here is a convergent geometric series which we can evaluate:

$$P(q) = \frac{\beta}{1-\alpha+\beta} . \tag{3.17}$$

In general, (3.17) does not yield zero. For example, if α is $3/4$ and β is $3/8$, then $P(q)$ is $3/5$.[27]

We conclude that Lewis is mistaken. It is not the case that a "regress of probability values" always yields zero. We have just seen an example of such a series, consisting in a sum with an infinite number of terms, that yields a number other than zero. Since Lewis's statement is invalid, it cannot support his main claim that probability statements only make sense if they presuppose certainties.[28]

3.5 A Nonuniform Probabilistic Regress

The counterexample in the previous section is a very special case. For in demonstrating that a probabilistic regress makes sense, we have assumed

[27] Eq.(3.17) gives in fact the fixed point of a Markov process. The stochastic matrix governing the process is regular, and the iteration is guaranteed by Markov theory to converge to the solution of the fixed point, $p_* = \beta + (\alpha - \beta)p_*$. However, this quick route to (3.17) only works when the conditional probabilities are the same from step to step: in the general case that we consider in the next section Markov theory does not help, which is why we have not used it here. We shall discuss fixed points more fully in Sections 8.4 and Appendix D.

[28] This example shows that James Van Cleve's defence of Lewis, and thereby his attack on Reichenbach, is mistaken (Van Cleve 1977). Van Cleve argues that an infinite iteration of the rule of total probability must be vicious, because "we must complete it before we can determine any probability at all" (ibid., 328). But our counterexample to Lewis demonstrates that an infinite iteration may well be completable, in the sense that it is convergent and can be summed explicitly, yielding a definite value for $P(q)$.

that the conditional probabilities are uniform, i.e. that they remain the same throughout the entire justificatory chain. Such an assumption is of course rarely fulfilled. It is very uncommon that the degree to which proposition q is probabilistically supported by A_1 is the same as the degree to which A_1 is probabilistically supported by A_2, and so on.

However, it is possible to construct counterexamples without making the assumption that the conditional probabilities are uniform. The rule of total probability relating A_n to A_{n+1} is

$$P(A_n) = P(A_n|A_{n+1})P(A_{n+1}) + P(A_n|\neg A_{n+1})P(\neg A_{n+1}),$$

or, with the abbreviation of the conditional probabilities as α and β, as in the previous section:

$$P(A_n) = \alpha P(A_{n+1}) + \beta P(\neg A_{n+1}).$$

In the nonuniform case the conditional probabilities differ from one link to another, so we have to add an index n to α and β:

$$\begin{aligned} P(A_n) &= \alpha_n P(A_{n+1}) + \beta_n P(\neg A_{n+1}) \\ &= \beta_n + \gamma_n P(A_{n+1}), \end{aligned} \tag{3.18}$$

where α_n, β_n and γ_n are defined as follows:

$$\begin{aligned} \alpha_n &= P(A_n|A_{n+1}) \\ \beta_n &= P(A_n|\neg A_{n+1}) \\ \gamma_n &= \alpha_n - \beta_n. \end{aligned} \tag{3.19}$$

Imagine a finite probabilistic chain $A_0, A_1, \ldots, A_{m+1}$, where again A_0 is probabilistically supported by A_1, which is probabilistically supported by A_2, and so on. For notational convenience we have temporarily used A_0 for the target proposition q and A_{m+1} for the grounding proposition p. It is possible to concatenate all the instances of the rule of total probability to yield, for any $m \geq 0$,

$$P(A_0) = \beta_0 + \gamma_0\beta_1 + \gamma_0\gamma_1\beta_2 + \ldots + \gamma_0\gamma_1 \ldots \gamma_{m-1}\beta_m + \gamma_0\gamma_1 \ldots \gamma_m P(A_{m+1}). \tag{3.20}$$

Formula (3.20), of which a proof is given in Appendix A.1, is the nonuniform counterpart of formula (3.15) in the uniform case.

We have seen that, notwithstanding Lewis's opinion, the extension of the finite (3.15) to an infinite chain can be envisaged: in the uniform case the infinite extension is well-defined if the extreme values 0 and 1 for the conditional probabilities are excluded. Does it make sense to extend (3.20) to an infinite number of links? Can a probabilistic regress in the nonuniform case also be well-defined and moreover yield a nonzero value for the target? Again, one example is enough to refute Lewis's argument in this more general setting, and here it is:

$$\alpha_n = 1 - \frac{1}{n+2} + \frac{1}{n+3}\,; \qquad \beta_n = \frac{1}{n+3}\,; \qquad \gamma_n = 1 - \frac{1}{n+2}\,. \quad (3.21)$$

In (3.21) α_n and β_n depend nontrivially on n. The resulting infinite series is not a geometric series, as it was in the uniform case that was introduced in Section 3.4. Nevertheless, as is shown in Appendix A.5, when we insert the formulae (3.21) into (3.20) we can work out the sum explicitly, obtaining

$$P(A_0) = \tfrac{3}{4} - \tfrac{2m+5}{2(m+2)(m+3)} + \tfrac{1}{m+2}P(A_{m+1})\,. \quad (3.22)$$

In the limit that m goes to infinity, the second and the third terms on the right-hand side of (3.22), namely $\frac{2m+5}{2(m+2)(m+3)}$ and $\frac{1}{m+2}P(A_{m+1})$, both go to zero. Thus only the term $\frac{3}{4}$ survives in the limit, so that $P(A_0)$, that is the probability of the target, $P(q)$, equals $\frac{3}{4}$. Here then is a new and more general case that invalidates Lewis's argument that an infinite probabilistic regress must yield zero.

3.6 Usual and Exceptional Classes

The above examples not only illustrate that Lewis was mistaken, but also that a probabilistic regress can have a limit and in that sense be benign. But what are the conditions under which this is so? When exactly does a probabilistic regress yield a well-defined value for the target proposition?

In general there exist two conditions. Each of them is necessary, and together they are sufficient. Look again at our finite nonuniform chain, (3.20):

$$P(A_0) = \beta_0 + \gamma_0\beta_1 + \gamma_0\gamma_1\beta_2 + \ldots + \gamma_0\gamma_1\ldots\gamma_{m-1}\beta_m + \gamma_0\gamma_1\ldots\gamma_m P(A_{m+1})\,.$$

The right-hand side of this equation consists of two parts, namely the sum of conditional probabilities,

$$\beta_0 + \gamma_0\beta_1 + \gamma_0\gamma_1\beta_2 + \ldots + \gamma_0\gamma_1\ldots\gamma_{m-1}\beta_m\,,$$

and the remainder term,

$$\gamma_0\gamma_1\ldots\gamma_m P(A_{m+1})\,.$$

The first condition for a benign probabilistic regress is that the series of conditional probabilities converges in the limit. The second condition is that, as m is taken to infinity, the remainder term goes to zero.

As we prove in Appendix A.3, the first condition is always satisfied, given that we assume probabilistic support, i.e. the constraint $P(A_n|A_{n+1}) > P(A_n|\neg A_{n+1})$ for all n. No matter whether we are dealing with uniform or with nonuniform conditional probabilities, the infinite series

$$\beta_0 + \gamma_0\beta_1 + \gamma_0\gamma_1\beta_2 + \gamma_0\gamma_1\gamma_2\beta_3 + \ldots\,, \tag{3.23}$$

always converges. However, the matter is different as far as the second condition is concerned. This condition is satisfied in the uniform situation (with the restriction that α is not equal to one and β is not equal to zero), but it is not always satisfied in the nonuniform situation. We shall call the class of cases where both conditions are fulfilled *the usual class*.[29] In the usual class the probability of the target is equal to the following convergent series of terms, each of which is a function of the conditional probabilities only:

$$P(q) = \beta_0 + \gamma_0\beta_1 + \gamma_0\gamma_1\beta_2 + \gamma_0\gamma_1\gamma_2\beta_3 + \ldots\,. \tag{3.24}$$

The class of cases in which only the first requirement is fulfilled we will call *the exceptional class*. Regresses in the exceptional class do not furnish counterexamples to Lewis's conclusion; but those in the usual class, on the other hand, do so, on condition that at least one of the β_n is nonzero.

When does a nonuniform probabilistic regress fall within the exceptional class? For our purpose this question is of course important, since it creates the watershed between probabilistic regresses which are benign (in the sense that they yield an exact and well-defined value for the target) and those that are not (in the sense that they only yield such a number if they have a first

[29] In the usual class the infinite series (3.23) converges even if one relaxes the condition of probabilistic support. However, since we are interested in justification, of which probabilistic support is a necessary condition, this extension of the domain of convergence is not required for our purposes. Moreover, the condition of probabilistic support is needed for our conception of epistemic justification as a trade-off (see Chapter 5) as well as for convergence in the probabilistic networks discussed in Chapter 8.

member, a ground). Clearly the answer to this question depends on whether the remainder term vanishes in the limit. We have seen that this will be the case if the factor $\gamma_0\gamma_1 \ldots \gamma_m$ vanishes as m goes to infinity. For then the remainder term $\gamma_0\gamma_1 \ldots \gamma_m P(A_{m+1})$ will die out, since $P(A_{m+1})$, the probability of the grounding proposition, cannot be greater than one.

But when exactly does $\gamma_0\gamma_1 \ldots \gamma_m$ go to zero? That is the key question. As we show in Appendix A.4, the answer depends entirely on the asymptotic behaviours of α_n and β_n. The factor $\gamma_0\gamma_1 \ldots \gamma_m$ goes to zero if and only if α_n does not tend to one more quickly than $1/n$ tends to zero, *or* if β_n does not tend to zero more quickly than $1/n$ tends to zero. If at least one of these disjuncts applies, then the nonuniform probabilistic regress falls within the usual class. It then yields a unique probability value for the target proposition, A_0 or q, which does not depend on an inaccessible unconditional probability at infinity. That is, it does not depend on the value of $P(A_{m+1})$ — or $P(p)$ — in Eq.(3.20) in the limit that m goes to infinity.[30] A nonuniform probabilistic regress within this usual class constitutes a counterexample to Lewis's argument. A specific instance is provided by the example (3.21), for this lies in the usual class, since the remainder term in (3.22), $\frac{1}{m+2}P(A_{m+1})$, goes to zero as m goes to infinity. In this limit the right-hand side of (3.22) tends to $\frac{3}{4}$.

If, however, α_n goes to one very quickly *and* β_n goes to zero very quickly as n tends to infinity, more quickly in fact than $1/n$ tends to zero, then the nonuniform probabilistic regress belongs to the exceptional class. In this case the regress does not result in a unique, well-defined probability value for the target proposition, since the unknown probability of the ground still plays a significant role. The regress is now vicious in the sense that the probability of the target depends in part on the inaccessible ground, and it would not form a counterexample to Lewis's foundationalist argument.

An example of a regress in the exceptional class is as follows,

$$\beta_n = \frac{1}{(n+2)(n+3)} \qquad\qquad \gamma_n = 1 - \frac{1}{(n+2)^2}\,, \tag{3.25}$$

so that

$$\alpha_n = \beta_n + \gamma_n = 1 - \frac{1}{(n+2)^2(n+3)}\,.$$

Here $1 - \alpha_n$ and β_n both tend to zero as n tends to infinity more quickly than $\frac{1}{n}$ tends to zero, which shows that the example is indeed a member of

[30] That the resulting system is consistent, in the sense that there exists at least one assignment of probabilities for all possible conjunctions of the propositions A_n, has been demonstrated by Frederik Herzberg (Herzberg 2013).

the exceptional class. In Appendix A.6 we work out the expression for the probability of the target proposition, obtaining

$$P(A_0) = \tfrac{3}{8} - \tfrac{2m+5}{4(m+2)(m+3)} + \tfrac{1}{2}\tfrac{m+3}{m+2}P(A_{m+1})\,. \tag{3.26}$$

In this case the remainder term, $\frac{1}{2}\frac{m+3}{m+2}P(A_{m+1})$, does not vanish in the limit. It becomes formally one half times the limit of $P(A_{m+1})$ as m tends to infinity, which is ill-defined.

A probabilistic regress in the exceptional class is characterized by the fact that it is actually very close to a regress of entailments, i.e. to the 'classical' regress, in which A_{n+1} entails A_n for all n. It is therefore to be expected that a straightforward classical regress will also fail to provide us with a counterexample to Lewis's claim, and this is indeed the case. Here is how a classical regress looks in our probabilistic formalism. If A_{n+1} entails A_n for all n, then

$$\alpha_n = P(A_n|A_{n+1}) = 1\,;$$

and it is shown in Appendix A.7 that (3.20) reduces in this case to

$$P(\neg A_0) = \gamma_0\gamma_1\ldots\gamma_m P(\neg A_{m+1})\,, \tag{3.27}$$

for any m. We have to consider various possibilities for the behaviour of

$$\beta_n = P(A_n|\neg A_{n+1})$$

as n tends to infinity. If β_n were to tend to zero no more quickly than $1/n$ does, the product $\gamma_0\gamma_1\ldots\gamma_m$ in (3.27) would tend to zero as m tends to infinity, so $P(\neg A_0) = 0$, irrespective of the behaviour of $P(\neg A_{m+1})$. Moreover it follows also that $P(\neg A_n) = 0$ for all n, which means that β_n is not defined. This is inconsistent, so we conclude that after all β_n must tend to zero more quickly than $1/n$. But then the product $\gamma_0\gamma_1\ldots\gamma_m$ tends to some non-zero limit, and so $P(\neg A_0)$ is not uniquely determined, since $P(\neg A_{m+1})$ can be assigned no particular limit as m goes to infinity. The regress of entailments, or implications, is thus necessarily in the exceptional class.

A very special case is when

$$\beta_n = P(A_n|\neg A_{n+1}) = 0 \tag{3.28}$$

for all n. We have then $P(\neg A_0) = P(\neg A_n)$ for all n, so all the probabilities, $P(A_n)$, have the same, undetermined value. Eq.(3.28) implies that $P(\neg A_n|\neg A_{n+1}) = 1$, which is to say that $\neg A_{n+1}$ entails $\neg A_n$, which of course means that A_n entails A_{n+1} (up to measure zero). If $\alpha_n = 1$ and $\beta_n = 0$, then

A_n implies, and is implied by A_{n+1}: there is a regress of bi-implication all the way along the chain. All the probabilities are the same, but the value is undetermined by the regress. Such a regress of bi-implication is vicious in our sense, for here the truth value of the target cannot be determined in the absence of the truth value of the first member.

To summarize, the system of conditional probabilities belongs to the usual class if and only if $1-\alpha_n$ *or* β_n do not tend to zero more quickly than $1/n$ tends to zero. On the other hand, if $1-\alpha_n$ *and* β_n both tend to zero more quickly than $1/n$, then the system belongs to the exceptional class, and the unconditional probabilities of the propositions are not determined. The situation in which α_n is nearly one, and β_n is nearly zero, is close to the case of bi-implication. We therefore might call the exceptional class the case of *quasi-bi-implication*.

3.7 Barbara Bacterium

In this chapter we have introduced the concept of a probabilistic regress, that is an epistemic chain of the form

$$q \longleftarrow A_1 \longleftarrow A_2 \longleftarrow A_3 \longleftarrow A_4 \ldots$$

where the arrow is interpreted in terms of probabilistic support. We examined Lewis's view that such a regress is absurd, since it allegedly implies that the probability of q is zero. According to Lewis, the only way to avoid the absurdity was to stop at a proposition, p, which is certain:

$$q \longleftarrow A_1 \longleftarrow A_2 \longleftarrow A_3 \longleftarrow A_4 \ldots \longleftarrow p.$$

We have opposed Lewis's argument by giving counterexamples, i.e. probabilistic regresses which yield a unique, nonzero probability value for the target. Some of these regresses were based on uniform conditional probabilities, others on nonuniform ones.

All our counterexamples were abstract. This is somewhat unfortunate, since a familiar objection to infinite regresses is that they are not concrete and lack practical relevance. The objection becomes even more pressing if one distinguishes (as we did not do here but will do in later chapters) between propositions and beliefs. Propositions are abstract entities, but beliefs are propositional attitudes that people really have. Whereas the idea of an infinite propositional regress might sound not unreasonable, an infinite doxastic regress seems a contradiction in terms. Where could we ever find a doxastic series of infinite length?

In the next chapters we will come back to this objection, and then we will also discuss the distinction between a propositional and a doxastic regress. At this juncture we will restrict ourselves to showing that a probabilistic regress of propositions also is relevant to a real-life situation.

Imagine that we are trying to develop a new medicine to cure a disease. In this connection, we want to know whether a particular bacterium has a certain trait, T. Bacteria reproduce asexually, so one parent, the 'mother' bacterium, alone produces offspring. After having carried out many experiments, one day we take from a batch a particular bacterium, which we call Barbara. From our experiments we know that the probability that Barbara has T is considerably greater if her mother has T than if her mother lacks it. So if q is 'Barbara has T' and A_1 is 'Barbara's mother has T', then we can say that A_1 probabilistically supports q. It is not certain that Barbara has T if her mother has the trait, but on the other hand Barbara could have T even if her mother does not have it. Thus $1 > P(q|A_1) > P(q|\neg A_1) > 0$.

The unconditional probability of Barbara having T is given by

$$P(q) = P(q|A_1)P(A_1) + P(q|\neg A_1)P(\neg A_1).$$

Whereas the conditional probabilities in this equation, $P(q|A_1)$ and $P(q|\neg A_1)$, may be assumed to have been determined from our experiments, obtaining $P(A_1)$ is a problem. What is the probability that Barbara's mother has T? We know that it is given by

$$P(A_1) = P(A_1|A_2)P(A_2) + P(A_1|\neg A_2)P(\neg A_2),$$

where $P(A_2)$ is the probability that Barbara's grandmother has T, which in turn is conditioned by $P(A_3)$, the probability that Barbara's great-grandmother has T.[31]

It will be clear that we can only compute $P(q)$ if we know $P(A_3)$. And the situation remains the same, even if we add more and more instances of the rule of total probability, going further and further back in Barbara's ancestry. It seems we are only able to compute the probability that Barbara has T if we know what is the unconditional probability that her primordial mother had T. So at first sight it looks as though foundationalists are right: if q is probabilistically justified by A_1, which is probabilistically justified by A_2,

[31] In the reading of Pastin the probability intended by Lewis would be $P(q \wedge A_1)$, see footnote 20. But this is neither the probability of interest nor does it fit what is at stake in the debate between Lewis and Reichenbach.

et cetera, then we have to know for sure the probability of the grounding proposition in order to be able to calculate the probability of q.

This impression, intuitive as it may seem, is however incorrect, and we have already seen why. The chain $q \longleftarrow A_1 \longleftarrow A_2 \longleftarrow A_3$ leads to:

$$P(q) = \beta + \beta(\alpha - \beta) + \beta(\alpha - \beta)^2 + (\alpha - \beta)^3 P(A_3)\,,$$

see (3.14). Going infinitely far back into Barbara's ancestry, we obtain (3.16):

$$P(q) = \beta + \beta(\alpha - \beta) + \beta(\alpha - \beta)^2 + \ldots\,.$$

This does not have a grounding proposition p. A primordial mother of Barbara makes no contribution, yet we are able to calculate the probability that Barbara herself has T, and this probability, notwithstanding Lewis's opinion, is not zero.

Let A_n be the proposition: 'Barbara's ancestor in generation n has T'. Let the probability that a bacterium has T if her mother has T be 0.99, and the probability that a bacterium has T if her mother lacks it be 0.02. So $\alpha = P(A_n|A_{n+1}) = 0.99$, $\beta = P(A_n|\neg A_{n+1}) = 0.02$, and hence $\gamma = \alpha - \beta = 0.97$. Now (3.16) becomes:

$$P(q) = \frac{\beta}{1-\gamma} = \frac{\beta}{1-\alpha+\beta}\,,$$

in agreement with (3.17). With the numbers chosen for α and β, we can now calculate the probability that Barbara has T: it is $\frac{2}{3}$.

The foregoing example made use of uniform conditional probabilities. As an example of a nonuniform probabilistic regress, suppose that an effect of the increasing pollution of the nutrient, as a result of the growing mass of bacteria in it, is that the probability of a bacterium having T increases as time goes on, quite independently of whether the mother bacterium has T. For example, if $\alpha_n = P(A_n|A_{n+1}) = a + b^{n+1}$ and $\beta_n = P(A_n|\neg A_{n+1}) = b^{n+1}$, where a and b are positive numbers such that $a + b < 1$, then α_n and β_n are different from generation to generation, although $\gamma_n = a$ is constant. Note that, since b is less than one, the factor b^{n+1} increases as n decreases, so in Barbara's remote ancestry there was little pollution, but it increases from generation to generation until Barbara herself appears on the scene. Eq.(3.20) once more reduces to a finite geometric series that can be summed:

$$\begin{aligned} P(q) &= b\left[1 + ab + (ab)^2 + \ldots (ab)^m\right] + a^{m+1} P(A_{m+1}) \\ &= b\,\frac{1-(ab)^{m+1}}{1-ab} + a^{m+1} P(A_{m+1})\,. \end{aligned}$$

In the case of an infinite number of generations, since $(ab)^{m+1}$ and a^{m+1} both vanish in the limit of infinite m, we find

$$P(q) = \frac{b}{1-ab}\,. \tag{3.29}$$

For example, if $a = \frac{1}{3}$ and $b = \frac{3}{5}$, we find from (3.29) that $P(q) = \frac{3}{4}$.

One might object that our argument so far is still not very realistic, to put it mildly. For a start, the assumption that conditional probabilities are known as precise numbers is a travesty of what is attainable in scientific practice. In real experiments the conditional probabilities are imprecise, merely being known to lie within some specified interval, and as a result, the unconditional probability of the target, too, is subject to measurement error.

Fortunately, when the conditional probabilities are uniform, as for example in the case of Barbara, then it is relatively easy to determine the interval within which the target probability must lie. For suppose that $P(A_n|A_{n+1})$ is in the interval $[\alpha_m, \alpha_M]$, and $P(A_n|\neg A_{n+1})$ is in the interval $[\beta_m, \beta_M]$. It can be shown that expression (3.17) for $P(q)$ is an increasing function of both α and of β;[32] and this means that the uncertainty in $P(q)$ is given by

$$\frac{\beta_m}{1-\alpha_m+\beta_m} < P(q) < \frac{\beta_M}{1-\alpha_M+\beta_M}\,,$$

on condition that $\alpha_M - \beta_m < 1$.

In the more general case where the conditional probabilities are not uniform, the calculation of the uncertainty in the value of $P(q)$ is a little more intricate. However, since the condition of probabilistic support is in force, all the terms in Eq.(3.23) are positive, and it can be done without too much effort. One has to minimize and maximize each term, within the experimental error bounds, in order to obtain lower and upper bounds on $P(q)$.

Even so, one might still feel the urge to protest that we are not dealing with real life situations. No bacterium has an infinite number of ancestor bacteria, if only because of the fact of evolution from more primitive algal slime,

[32] The partial derivatives of $\frac{\beta}{1-\alpha+\beta}$ with respect to α and β are both positive:

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\frac{\beta}{1-\alpha+\beta} &= \frac{\beta}{(1-\alpha+\beta)^2} > 0 \\
\frac{\partial}{\partial\beta}\frac{\beta}{1-\alpha+\beta} &= \frac{1-\alpha}{(1-\alpha+\beta)^2} > 0\,.
\end{aligned}$$

which had grown out of earlier life forms, which sprang from inanimate matter, which originated in a supernova explosion, and so on.

This is of course true, and it makes short shrift of any remaining thought about a beginning in the form of a first bacterium.[33] For our approach, however, the issue is moot. The reason is that the further away a node in the chain is from the target, the smaller its influence on the target becomes. Applied to Barbara: long before we reach the stage where her ancestor bacteria evolve from more primeval life forms, they have become totally irrelevant to the question whether Barbara has T. This phenomenon we call 'fading foundations', and it is explained in the next chapter.

[33] Sanford 1975, 1984; Rescher 2010, 56.