

William Youdstone, *Labyrinths of Reason: Paradox, Puzzles, and the Frontiers of Knowledge*
New York: Anchor Press, 1988

2

INDUCTION

Hempel's Raven

THE BEST-KNOWN modern paradox of confirmation was proposed by German-born American philosopher Carl G. Hempel in 1946. Hempel's "paradox of the ravens" deals with induction, the drawing of generalizations. It is a mischievous reaction to those who think that science may be resolved into a cookbook scientific method.

Hempel imagined a birdwatcher trying to test the hypothesis "All ravens are black."¹ The conventional way of testing that theory is to

¹ Ornithological note: "Raven" usually means a single species, *Corvus corax*, found world-wide in the Northern Hemisphere. This is the raven of Poe's poem. Ravens are iridescent black with glints of green, purple, and blue predominating. Mexico and the American Southwest also have a smaller bird called the Chihuahuan raven (*Corvus cryptoleucus*). This bird is black with a white neck that is exposed when the bird crooks its head. I have failed to find

seek out ravens and check their color. Every black raven found confirms (provides evidence for) the hypothesis. On the other hand, a single raven of any color other than black disproves the hypothesis on the spot. Find even one red raven and you need look no further: The hypothesis is wrong.

All are agreed on the above. Hempel's paradox begins with the claim that the hypothesis may be restated as "All nonblack things are nonravens." Logic tells us that this is entirely equivalent to the original hypothesis. If all ravens are black, then certainly anything which isn't black can't be a raven. This rewording is known as a *contrapositive*, and the contrapositive of any statement is identical in meaning.

"All nonblack things are nonravens" is a lot easier to test. Every time you see something that *isn't* black, and it turns out *not* to be a raven, this restated hypothesis is confirmed. Instead of looking for ravens on damp, inaccessible moors, you need only look for non-black things that aren't ravens.

A blue jay is sighted. It's nonblack and it's not a raven. That confirms the contrapositive version of the hypothesis. So does a pink flamingo, a purple martin, and a green peacock. Of course, a nonblack thing doesn't even have to be a bird. A red herring, a gold ring, a blue lawn elf, and the white paper of this page also confirm the hypothesis. The birdwatcher does not have to stir from his easy chair to gather evidence that all ravens are black. Wherever you are right now, your visual field is filled with things that confirm "All ravens are black."

Now clearly this is ridiculous. There is yet a further absurdity. To see it, suppose that you want to dismiss the paradox by saying, all right, evidently the blue jay or the red herring *does* confirm "All ravens are black" to some infinitesimal degree. If you could sum-up a magic genie, capable of examining all the nonblack things in the world in the blink of an eye, and if that genie found that not one of those nonblack things was a raven, that assuredly would prove that there are no nonblack ravens—that all ravens are black. Maybe it is not so incredible that a red herring could confirm the hypothesis.

Don't get too cozy with this resolution. It is easy to see that that

any mention of albino or other distinctly nonblack ravens, but would not be surprised to learn that such birds exist.

None of this, of course, has anything to do with the present discussion. Aside from this footnote, I will assume that the color of ravens is perfectly well defined, and that no one has ever seen a raven that is any color other than black.

same red herring *also* confirms "All ravens are white." The contrapositive of the latter is "All nonwhite things are nonravens," and the herring, being a nonwhite thing, confirms it. An observation cannot confirm two mutually exclusive hypotheses. Once you admit such a patent contradiction, it is possible to "prove" anything. The red herring confirms that the color of all ravens is black, and also that that color is white; ergo:

Black is white. QED.

Reasonable assumptions have led to resounding contradiction.

To scientists, Hempel's paradox is more than a puzzle. Any hypothesis has a contrapositive, and confirming instances of the contrapositive are often very easy to find. Something is certainly wrong. But what?

Hempel's raven is a good introduction to the perils and puzzles of confirmation. Of all the major paradoxes to be discussed, it is among the most nearly resolved. It will be worthwhile to back up a bit before coming to the resolution, though, to discuss the background of the paradox.

Confirmation

To put it in as few words as possible, confirmation is the search for truth. It is the mainspring of science, and more than that, it is something we do every day of our lives.

Analyzing confirmation is almost like analyzing sneezing: We know what it is, but it is usually so automatic that it is hard to say exactly how it is done. The paradoxes of confirmation probably owe a lot to this shared set of subconscious expectations. These expectations can lead us astray.

As you remember from high school, there is a "scientific method" that goes roughly as follows. You form a hypothesis—a guess about how the world works. Then you try to test it through observation or experiment. The evidence you gather either confirms the hypothesis or refutes it. Like much of what you learn in high school, this is true while leaving important things unsaid.

Most useful hypotheses are generalizations. Hempel's paradox plays off a bit of common sense called "Nicod's criterion" after philosopher Jean Nicod. To put it in terms of black ravens, this says that (a) sighting a black raven makes the generalization "All ravens are black" more likely; (b) sighting a nonblack raven disproves the statement; and (c) observations of black nonravens and nonblack

nonravens are irrelevant. A black bowling ball or a blue lawn elf cannot tell us anything about the color of ravens. Nicod's criterion is behind all scientific inquiry, and if something is wrong with it, we are in serious trouble indeed.

The sighting of a black raven furnishes evidence in favor of the hypothesis that all ravens are black, but of course does not *prove* that the hypothesis is right. No single observation can do that. Sightings of black ravens, in the absence of ravens of any other color, increase your confidence (reasonably enough) that all ravens are black.

Confirmation is trickier than it appears. You might think that the more confirming evidence for a hypothesis, the more likely it is to be true. Not necessarily. It is possible for two confirming observations to prove a hypothesis *false*. That is what happens in the following thought experiment, inspired by philosopher Wesley Salmon.

Matter and Antimatter

Suppose that some of the planets in the universe are made of matter and some of antimatter (as has been speculated). Matter and antimatter look exactly alike. There is no way of telling, by examining a distant star in a telescope, whether it is made of matter or antimatter. Even the star's light gives nothing away, for the photon is its own antiparticle, and an antimatter star shines with the same kind of light as a regular star. The only thing is, when antimatter touches regular matter—BOOM!!! Both are annihilated in a tremendous explosion.

This unfortunate fact makes interstellar contact hazardous. A spaceship from planet X chances upon a spaceship from planet Y. They radio messages to each other (radio waves are made of photons too and are neither matter nor antimatter). Computers on board the ships decipher the alien languages and establish diplomatic relations. The two spaceships agree to dock and exchange goodwill ambassadors. Everything is fine until the last moment. Then the rockets make contact and BOOM!!!—or not, depending on the composition of planets X and Y. Whenever one is matter and the other antimatter, both ships are blown to smithereens. (There's no explosion if *both* spaceships are antimatter.)

One day, astronomers here on Earth report that they've sighted two tiny points of light that may be spaceships approaching each other. They're not sure the objects are spaceships, but on the basis of past experience the astronomers can say that for each point of

light there is a 30 percent chance it is a spaceship and a 70 percent chance it is an irrelevant natural phenomenon. It is also known from past experience that any pair of spaceships that approach closely always do dock. All the other species in the galaxy seem oblivious to the matter/antimatter problem, and have to learn the hard way.

So the big question is: Will they blow up or not? Oddsmakers in Las Vegas start accepting ghoulish bets on whether there will be an annihilation. The oddsmakers reason like this: It is known that two-thirds of the planets in the universe are made of matter and one-third are antimatter. Thus for each point of light there is a 70 percent chance it is a natural phenomenon of no interest here; a 20 percent chance it is a spaceship made of matter; and a 10 percent chance that it is an antimatter spaceship.

Call the two points of light A and B. An annihilation can occur in one of two mutually exclusive ways. Either object A is a matter spaceship and B is an antimatter spaceship, or A is an antimatter spaceship and B is a matter spaceship. The chance of the first case is 20 percent of 10 percent, or 2 percent. The odds of the second case is 10 percent of 20 percent; again, 2 percent. Since these two possibilities are mutually exclusive, the total chance of an annihilation is 2 percent plus 2 percent: 4 percent.

The oddsmakers set the payoffs to bettors based on this calculated probability. Now suppose that a space prospector, returning home to Earth, grazes object A in a trillion-to-one freak accident. The prospector learns that object A is a spaceship and is made of ordinary matter (this from the fact that there was no explosion). Arriving on Earth, the prospector finds out about the possible annihilation and the Las Vegas book on it.

The prospector would do well to exploit his "inside information" and bet on annihilation. He knows for a fact that object A is a spaceship, whereas everyone else thinks it is probably (70 percent chance) just an asteroid or some other natural body. Given that A is a spaceship of ordinary matter, there is a 10 percent chance of annihilation, since that is the chance that object B is a spaceship and made of antimatter. The oddsmakers have set the chance at 4 percent, but the prospector, with his more complete knowledge, can set the chance at 10 percent.

Fine. Now what if another space prospector had an identical accident with object B, and also determined that it is a spaceship made of matter? This second prospector could of course use the same reasoning to arrive at the same conclusion: that the chance of

annihilation has been boosted from 4 percent to 10 percent. *But* the combined information of the two prospectors actually rules out an annihilation entirely. They have determined that the spaceships are both made of the same kind of matter as Earth, and that means that the chance of annihilation is a big fat zero!

Absolute and Incremental Confirmation

Two items of evidence (the prospectors' collisions with the spaceships) each confirm the hypothesis that there will be an annihilation, even though the observations together refute it. I prefer to call this an irony rather than a paradox, for there is no doubt that such strange turns of affairs can exist. The probability calculations of the oddsmakers, of the prospectors, and of us, aware of both prospectors' experiences, are sound. These peculiar situations have been studied intensively by confirmation theorists.

The peculiarity is partly semantic. The verb "confirm" is used in two ways. In everyday speech, we almost always use "confirm" in the *absolute* sense, to mean that something is clinched, established beyond reasonable doubt. "The boss confirmed that Sandra got the raise" means that, whatever the doubts beforehand, it is now just about 100 percent certain that Sandra got the raise.

Hardly any experiment provides absolute confirmation of a hypothesis. Scientists and confirmation theorists often use "confirm" in the *incremental* sense. To confirm incrementally is to "provide evidence for" or "increase the probability of." We speak of probabilities because confirmation of a generalization is always tentative.

You can incrementally confirm a hypothesis that was, and still is, unlikely to be true. We would not say "The boss confirmed that Sandra got the raise" to mean that some equivocal comment of the boss's has upped the chance of Sandra's getting the raise from 15 percent to 18 percent. But that type of confirmation is typical of scientific research.

Incremental confirmation is common to ironic situations such as the spaceship annihilations. Each prospectors' information increments a low chance of annihilation (4 percent) to a greater but still low chance (10 percent). Taken together, their information decreases the chance to zero. It is reassuring that such flukes vanish when the chances are higher, when a hypothesis is close to being confirmed in the absolute sense.

You can demonstrate this by playing with the odds a little. Recast the situation, giving the oddsmakers a better handle on the real

state of affairs. For each object, the odds are now 10 percent that it is a natural phenomenon, 80 percent that it is a matter spaceship, and 10 percent that it is an antimatter spaceship. Then the oddsmakers must set the chance of annihilation at (80 percent of 10 percent) plus (10 percent of 80 percent)—or 16 percent. Each prospector, upon learning for certain that one of the objects is a matter spaceship, can figure (as before) that the chance of annihilation is 10 percent, the chance that the other object is an antimatter spaceship. Now each prospector's estimate is *less* than the oddsmakers'. This is as it should be, since they know more than the oddsmakers, and the actual chance happens to be zero.

Counterexamples

As this shows, confirmation is only half the story. Evidence may also refute or disconfirm a hypothesis. Philosophers of science, notably Sir Karl Popper, emphasize the role of refutation.

You might think that it's just a matter of saying a glass is half full or half empty. Actually, there is an asymmetry between confirmation and refutation. It is much easier to refute a generalization than to prove it.

A counterexample is an exception to a putative rule. A white raven is a counterexample to the hypothesis that all ravens are black. A white raven does not merely make the hypothesis less likely. It proves the hypothesis wrong in one fell swoop. Logicians call this *modus tollens*, or "denying the consequent."

Rarely is the situation so simple in practice. There have been many "counterexamples" to the hypothesis that there is no Loch Ness monster. Every alleged sighting is one. Yet most scientists continue to believe that there is no Loch Ness monster. It is evident that not all supposed counterexamples have enough weight to refute an otherwise confirmed hypothesis.

Most hypotheses on the edges of current knowledge can be tested only in situations where many "auxiliary" hypotheses are tested as well. Auxiliary hypotheses are background assumptions about how the main hypothesis fits into the general body of knowledge; how microscopes, telescopes, and other equipment necessary to test the hypothesis operate; and so on. These auxiliary hypotheses often rule out any quick use of *modus tollens*.

Wesley Salmon cited a neat case of two similar counterexamples leading to rejection of auxiliary and main hypotheses, respectively. Newton's theory of gravity makes predictions about the future mo-

tions of the planets. In the nineteenth century, these predictions for the orbit of Uranus were found to be slightly, but consistently, wrong.

Some astronomers wondered if the discrepancies might be due to an unknown planet beyond Uranus. Once this planet (Neptune) was discovered in 1846, Newton's theory was not only removed from doubt but strengthened. Neptune was further evidence for Newton's theory.

At about the same time, other irregularities were noted in Mercury's orbit. Astronomers also tried to find a planet near Mercury that might account for the deviation. French amateur astronomer D. Lescarbault reported seeing a planet within Mercury's orbit in 1859. The planet was accepted as real and named Vulcan by Urbain Jean Leverrier, co-discoverer of Neptune. Subsequent astronomers could not find the planet, though, and it was soon branded a mistake. Mercury continued to depart from its predicted orbit. The deviations were not haphazard but regular, and distinctly different from what Kepler's laws (founded on Newton's gravity) predicted.

In this case, the discrepancies were ultimately accepted as evidence that Newton's theory of gravity is wrong. Mercury's wobbling orbit was one of the earliest confirmations of Einstein's general relativity.

The history of Neptune and Vulcan demonstrates two features of counterexamples. First, a counterexample may refute an auxiliary hypothesis rather than the main one. It is important to find out which is at fault. There is usually such ample room for speculation that instant refutations are rare. Second, when a theory is thrown out, it is in favor of a broader theory that makes many of the same predictions as the original. Under typical conditions in the solar system, Einstein's general relativity predicts gravitational effects all but identical to those of Newton's simpler theory. The difference turns up only in very intense gravitational fields. Of the planets, Mercury, being closest to the sun, is most subject to these relativistic effects. It alone seems to be out of step with Newton's laws.

Crank Theories

Not only should a new theory account for the successful predictions of the theory it would replace. It should offer new, different predictions of its own. In Karl Popper's terms, the new theory must

have greater "empirical content." It must make more testable predictions in more realms of experience than the old theory.

A new theory should be *more* open to possible refutation, not less. If there is one thing that is a dead giveaway for a crank theory, it is that the theory has been modified to restrict its own refutation. An honest hypothesis is open to being disproven. It's one thing to say, there's a ghost that appears in the old Miller mansion at the crack of midnight whenever there's a full moon. That kind of hypothesis is worthy of attention provided there is any reasonable evidence to support it: say, testimony of a few reliable eyewitnesses. Far more typical are ghost stories that restrict refutation: A ghost appears, but never when skeptics are around.

These restrictions usually indicate that a hypothesis has failed the first stages of the confirmation process and is being kept alive by those who wish to believe it regardless of its truth. No one started out believing that

- channelers have such erratic recall of their past lives that you can't expect them to know checkable historic data (like the name of the contemporary pharaoh's wife); or
- UFOs purposely abduct people who won't be believed by the "establishment" so that the aliens' presence will remain unknown; or
- bigfoot remains disintegrate with extraordinary rapidity, so no skeletons are found (or bigfoots scrupulously bury their dead, like us humans); or
- the stars (of astrology) impel, not compel.

All these provisos were tacked on after confirmation failed to materialize. That doesn't automatically mean that the modified hypotheses are false, but it is hardly encouraging. If the process of modifying to restrict refutation continues long enough, the ultimate result is the type of hypothesis that Popper sardonically calls "irrefutable." This may sound good, but think about what it means. It is a hypothesis that cannot possibly be proven false—one so wishy-washy that no possible observation is incompatible with it. That kind of hypothesis doesn't really say anything.

The proposition that "ESP exists, but it is so iffy that even the best psychics may do no better than chance in controlled experiments"—which is essentially what some ESP apologists have said—is beyond refutation. You might ask, "How would the world be any different if ESP *didn't* exist?"

Why can't scientists give poorly supported hypotheses the benefit

of the doubt? The main reason is that many, many hypotheses can be devised to account for any fixed body of data. If we say, "Okay, ESP exists, because no experiment has ruled it out" (which is true), we would have to allow a multitude of equally unrefuted hypotheses. In the end it is a desire for simplicity that leads scientists to accept only those hypotheses that can be confirmed. Indeed, says Popper, the aim of science should be to try to eliminate as many hypotheses as possible with new data.

Contrapositives

The basics of confirmation in place, let's return to Hempel's paradox with this added perspective. The first thing that concerns most people hearing the paradox for the first time is this business about the contrapositive. "Nonblack things" and "nonravens" are awkward constructions. Is "All nonblack things are nonravens" really equivalent to "All ravens are black"? If it's *not*, there is no paradox.

Here is a good way to see that they *are* logically equivalent. Forget about our human and imperfect attempts at knowledge. Pretend that we have at our service a genie who can ascertain any and all *specific* facts instantly. In other words, the genie can determine any of Hume's "matters of fact"—the direct, sensory results of any observation, without any interpretation, interpolation, or editorial comment.

Also like Hume, the genie claims that it doesn't quite understand generalizations. So if you want to know whether a statement such as "All ravens are black" is true, you have to explain it to the genie as an aggregate of individual observations. You have to tell the genie exactly what he should do to determine if Hempel's hypothesis is right or wrong.

It may come as a surprise that observations of black ravens are virtually irrelevant to the *ultimate* truth or falsity of "All ravens are black." This flatly contradicts the foregoing discussion, but remember we are now talking about the genie and not humans. The genie is going to determine the final, cosmic truth of the statement, not merely find evidence to support it. Observations of black ravens can neither prove nor disprove the statement.

Suppose the genie found a black raven. Would that prove that all ravens are black? Of course not. Suppose the genie found a million black ravens. Would that prove it? No; there could still be ravens of other colors. The statement "All swans are white" was supported

by all available evidence until the discovery of Australia. There are black swans in Australia.

Suppose that the universe is infinite and there is an infinity of other planets so similar to Earth that they have black ravens on them and that the genie thereby finds an infinite number of black ravens. Would *that* prove it? No; for the same reason. At this point the genie would rightly get impatient with us, for evidently no amount of black ravens will settle anything. Looking for black ravens is a wild-goose chase.

Think about it, and you will realize that the crux of the matter is nonblack ravens. The only way Hempel's statement can be *wrong* is for there to be a raven somewhere that isn't black. The only way the statement can be *right* is for there to be no such raven. To decide ultimate truth or falsity, the genie must search for nonblack ravens. If he finds even one, the statement is irretrievably false. If he searches the entire universe—everywhere a nonblack raven could possibly be—and finds none, then Hempel's statement is unimpeachably true.

In a pragmatic sense, "All ravens are black" only *seems* to be talking about black ravens. When you translate it into an operational definition for the genie, it really says: "There is no such thing as a nonblack raven."

Now let's have the genie test the contrapositive statement, "All nonblack things are nonravens." This is another pie-in-the-sky generalization incomprehensible to the genie. We explain: "The only way 'All nonblack things are nonravens' can be wrong is for there to be at least one nonblack raven. The only way it can be right is by the complete absence of nonblack ravens everywhere."

This is just how we explained the original statement. What you must do to prove or refute "All ravens are black" is identical to what you must do to prove or refute "All nonblack things are nonravens." That is strong grounds for asserting that the two statements are equivalent.

You might object that there is one slight difference. Does not the truth of "All ravens are black" imply that there is at least one black raven?

Take the hypothesis "All centaurs are green." The genie, looking for nongreen centaurs, would find none and report the statement true. Of course, there are no centaurs of any description. It sounds funny to say the statement is true, then.

This point is again one of semantics. Logicians *do* allow that statements such as "All centaurs are green" and "If X is a centaur,

X is green" are true. For various reasons it is most convenient to do so. Hence, to a logician, there is no distinction whatsoever between a statement and its contrapositive.

You are free to dissent and insist that there must be at least one green centaur for the statement to be true. Doing so creates this slight asymmetry between Hempel's original hypothesis and the contrapositive: With the original, you have to tell the genie to make sure there is at least one black raven before reporting the statement to be true. With the contrapositive, the genie must find at least one nonblack nonraven (like a red herring). I do not think that this significantly alters the essential equivalence of the statements. Finding the obligatory black raven or red herring is but a formality; the genie's real task in either case is making sure there are no nonblack ravens.

Never Say Never

A "negative hypothesis" is one that claims that something doesn't exist. It is extremely difficult to prove negative hypotheses. ("Never say never.") It is one thing for a genie to check every place a nonblack raven could be and thus prove that there is no such thing. It is something else for us humans.

You set off on a raven-hunting expedition, see lots of black ravens, and don't find any nonblack ravens. At length you start to get sick of the whole business. All your friends say you'll *never* find a nonblack raven. When is it okay to call it quits and stop looking?

As a practical matter, you do quit sooner or later. Thereafter you feel pretty confident that there are no nonblack ravens. This does not begin to prove that all ravens are black in a logically rigorous sense, though. To do that, you would indeed have to check everywhere in the universe that a raven might be. That is obviously an unreasonable requirement.

Philosophers have a word for processes requiring an infinity of action: *supertasks*. Some philosophers think that when ascertaining something requires an infinity of actions, it cannot be known at all. Michael Dummett gave this example: "A city will never be built at the North Pole." To test this, you might hop in a time machine, set it for a given year, and travel to that year to see if a city exists at the North Pole. If not, you set the time machine for a different year, and try again. You could know whether a city will exist at the North Pole at any point in time, but knowing whether it will *ever* be

built is something else again. Knowing that requires knowing an infinity of facts; doing an infinite amount of research.

If the universe is infinite, then "There are no nonblack ravens" is another proposition requiring an infinity of observations. Our genie is capable of empirical supertasks, but we are not. This is really why we confirm from sightings of black ravens and not from failures to sight nonblack ravens. The number of black ravens seen is a way of "keeping score" while actually looking for a counterexample. The more black ravens we have seen, without seeing any nonblack ones, the more confident we feel that there are no nonblack ravens. Nicod's criterion says that black ravens are a better way of keeping score in the progress of confirmation than are nonblack nonravens. To resolve Hempel's paradox, we must decide why this is so.

Stream of Consciousness

Try a different tack. Categories like "nonravens" and "nonblack things" are unnatural. Most of the time you are first aware that a "thing" is a raven or a herring or a steak knife. You don't naturally experience objects as "nonravens" or "nonherings" or "non-steak knives." Only Hempel's original formulation ("All ravens are black") dovetails with how people really think.

Your train of thought is quite different with the two versions of the hypothesis. When you see a raven, your thoughts normally run like this:

- (a) Look, there's a raven.
- (b) And it's black.
- (c) So it confirms the statement "All ravens are black."

Connecting a red herring to Hempel's hypothesis requires a more roundabout stream of consciousness!

- (a) There's a herring.
- (b) It's red.
- (c) Oh, wait, how does that raven paradox go? Yeah, it's a "non-black thing" . . .
- (d) . . . and it's not a raven.
- (e) So it confirms the statement "All nonblack things are nonravens" . . .
- (f) . . . which is the same as "All ravens are black."

Between steps (a) and (b) in the original formulation—the instant after you realize that the object is a raven, but before you think

about its color—the hypothesis is at risk. In that split second, the raven could be some other color and disprove the statement. The statement "All nonblack things are nonravens" is never really at risk in the second formulation. By the time you reach (c), you have already realized that the object is red (you deduced it was nonblack from the knowledge that it is red) and that it is a herring (you probably knew that all along).

Why is "raven" a reasonable category and "nonraven" not one? Well, ravens share many attributes in common, whereas "nonraven" is just a catchall term for anything that doesn't qualify. One category is figure and the other is ground. It's like the joke about a sculptor chiseling away everything that doesn't look like his subject. Sculptors don't think that way, and neither do scientists.

There is also a staggering numerical imbalance between the categories. Let's have one more go at the original idea: that the paradox has something to do with the relative numbers of ravens and non-black things.

Infinitesimal Confirmation

Hempel's reasoning need *not* lead to a paradox when the number of objects under investigation is clearly finite. Suppose that all that existed in the universe was seven sealed boxes. Unknown to you, five of the boxes contain black ravens; one contains a white raven; and one contains a green crab apple. Then you could reasonably feel that opening a box and finding the crab apple confirms "All ravens are black." In fact, the speediest way to prove or refute the hypothesis would be to inspect all the nonblack things. There are only two nonblack things vs. six ravens. Of course, this model is artificial. It assumes prior knowledge of the number of things being investigated. You hardly ever know that, at least not at the start of the investigation.

More typical is the case where the original and not the contrapositive hypothesis talks of a knowably finite class of objects. The time, effort, and money required to establish "All ravens are black" is tied to the number of ravens (or the number of nonblack things). According to R. Todd Engstrom of Cornell's Laboratory of Ornithology, the world population of common ravens is something like half a million. More troublesome is the number of nonblack things. It is astronomical.

One day it is discovered that there is a monster in Loch Ness. There's just one monster; sonar equipment has established that

there are no more of its kind. You want to test the hypothesis "All Loch Ness monsters are green." You approach the monster in a submarine, switch on the searchlights, and look out the porthole. The monster is green. Since there are no more Loch Ness monsters, the statement "All Loch Ness monsters are green" is thereby proven.

Here a single test of a hypothesis holds a lot of weight. There is only one chance for a nongreen monster to disprove the hypothesis. Taking the contrapositive seems even more ridiculous here than with the ravens. The contrapositive is "All nongreen things are non-Loch Ness monsters." Imagine going around and assigning a number to every nongreen thing in the world. Nongreen thing #42,990,276 is a blue lawn elf. Is it a non-Loch Ness monster? Yes! It supports the hypothesis. . . .

This is a woefully roundabout approach. Still supposing that there is just one Loch Ness monster and thus one potential counterexample, the chance that that arbitrary nongreen thing #42,990,276 is going to disprove the hypothesis is no greater than $1/N$, where N is the number of nongreen things. There might be something like 10^{80} atoms in the observable universe (which is the number written by putting 80 zeros after a "1"). There are at least that many nongreen objects. You might even claim that abstractions like numbers qualify as nongreen objects. Then the number is infinite.

This reasoning, anticipated by Hempel in his original musings in the 1940s, is very tempting. Possibly, a red herring *does* confirm "All ravens are black"—but only to an infinitesimal degree, because there are so many nonblack things. Checking the color of ravens is simply a more efficient way of confirming the hypothesis. In this vein, philosopher Nicholas Rescher estimated the costs of examining a statistically significant sample of ravens and nonblack objects. Rescher put the research tab at \$10,100 for the ravens vs. \$200 *quadrillion* for nonblack objects!

There remains the dilemma of how a red herring can confirm "All ravens are black" and "All ravens are white." You can try to picture it as being something like the mathematics of infinitesimals. The confirmation provided by a red herring for "All ravens are black" is on the order of $1/\text{infinity}$. The "infinity" in the denominator refers to the infinity of nonblack objects, of which the red herring is one. Since the herring is also a nonwhite object, it should confirm "All ravens are white" to an identical degree of $1/\text{infinity}$.

One divided by an infinite quantity is defined to be an infinitesimal, a number greater than zero but smaller than any regular fraction.

Does infinitesimal confirmation make the conflict any more palatable? We would be saying that a red herring confirms both "All ravens are black" and "All ravens are white," but only to an infinitesimal degree.

A small truth is still a truth; a small lie is yet a lie; and a contradiction is still a contradiction, even on an infinitesimal scale. The only out is to admit that the confirmation in both cases is precisely zero—as plain horse sense demands. Why, then, isn't a confirming instance of a hypothesis a confirming instance of its contrapositive?

The Paradox of the 99-Foot Man

Sometimes one paradox suggests the resolution of another. Paul Berent's paradox of the 99-foot man is another demonstration of the fallibility of Nicod's criterion. Say you subscribe to the reasonable belief: "All human beings are less than 100 feet tall." Everyone you've ever seen is a confirming instance of this hypothesis. Then one day you go to the circus and see a 99-foot-tall man. Surely you leave the circus *less* confident that all people are less than 100 feet tall. Why? The 99-foot man is yet another confirming instance.

There are two sources of this paradox. First, we don't always say what we mean. Sometimes the words we use imperfectly express the (often vague) hypothesis in our heads.

Chances are, you meant that no human being attains fantastic height; height an order of magnitude or more greater than the average. The precise figure of 100 feet was not vital. It was pulled out of the air as an example of the great height that you thought was definitely out of the question.

Had you been using the metric system, you might have said, "All human beings are less than 30 meters tall." Thirty meters comes to 98.43 feet, so the 99-foot man *would* be a counterexample to the 30-meter hypothesis. One feels that what you meant by saying "All human beings are less than 100 feet tall" is partially violated by the 99-foot man. It is like obeying the letter but not the intent of the law.

There is another root of the paradox. Have the hypothesis be the substance of a running bet you have with a friend. If ever a 100-foot-or-taller person turns up, you lose and owe your friend dinner at a fancy restaurant. The hypothesis is propounded not out of intellectual curiosity but solely to formalize the bet. Only the exact terms

of the wager count. The 99-foot man is close but no cigar. He poses no threat whatsoever of deciding the wager against you.

You would still feel that the 99-foot man hurts the chances of your hypothesis being right. This is because you know many facts about human growth and variation that allow you to deduce an increased likelihood of a 100-foot person from the fact of the 99-foot man. Nearly every human attribute recurs (even to a greater degree) eventually. The 99-foot man demonstrates that it is genetically and physically possible for a person to attain a height of about 100 feet.

Now imagine that you find a way to test your hypothesis without acquiring any nonessential information. At the busiest part of Fifth Avenue, you place a sensor in the sidewalk that detects whenever anyone walks over it. A hundred feet above the sidewalk sensor is an electric eye. When someone steps on the sensor, the electric eye determines whether a beam of light 100 feet above the sidewalk has been broken by a tall pedestrian. A recording device keeps track of the total pedestrian traffic and the 100-foot-or-taller pedestrian traffic.

You check the meter to see the results. The readout is "0/310,628"—meaning that 310,628 pedestrians have passed, none (0) of whom was 100 feet tall. Each of the 310,628 pedestrians is a confirming instance of the hypothesis. Each confirms the hypothesis to a precisely equal degree. It would be ridiculous to say that some of the pedestrians provided more confirmation than others when all you know of the pedestrians is that they are shorter than 100 feet.

If it so happened that the 99-foot man crossed Fifth Avenue and was one of the people counted, he would confirm the hypothesis as much as anyone else, in your state of ignorance. Thanks to him, the meter reads "0/310,628" rather than "0/310,627," and you are slightly more confident for it.

Clearly, it is the additional information (that the man is 99-feet tall, and what you know about human variation) that converts a simple confirming instance into one that effectively disconfirms.

Philosopher Rudolf Carnap suggested that there is a "requirement of total evidence." In inductive reasoning, it is necessary that you use all available information. If you know nothing of the 99-foot man and only look at the meter readings, then he is a valid confirming instance. When you know more, he's not.

The requirement of total evidence has occasioned much soul-searching in the scientific community because it addresses much of the research arena of biochemistry, astronomy, physics, and other fields. The way we investigate genes or subatomic particles is more

akin to the pedestrian traffic meter than simple observation. We do not meet RNA or quarks face to face; rather, we pose an exact question and learn the answers from machines.

Nothing is wrong with this, provided we do not limit our knowledge-gathering unnecessarily. If we are ignorant of other factors, and necessarily so, then we can generalize only from the information that is available. However, the more complete the information gathered, the more effective we are in making generalizations.

Ravens and Total Evidence

Let's recap. Science deals mostly in generalizations: "All X's are Y." Only through generalization can we compress our sensory experience into manageable form.

Generalizations are concealed negative hypotheses: "There is no such thing as a non-Y X"; or "The above rule has no exceptions." A generalization's contrapositive corresponds to the identical negative hypothesis.

In an infinite universe, proving a negative hypothesis is a supertask. (If the universe is merely finite but very big, proving a negative hypothesis is a herculean labor so close to a supertask as to make no difference.) We are incapable of supertasks, and have reason to be suspicious of knowledge attainable only through supertasks anyway.

Instead we establish generalizations through confirming instances: "X's that are Y"; black ravens in Hempel's example. This can never rigorously prove a generalization; only disprove it (through a counterexample: a nonblack raven). Tallying sightings of black ravens is a way of keeping score on how well established the hypothesis is. We feel that each black raven represents another instance in which the hypothesis was truly at risk of being disproved and came through unscathed. We do not feel that nonblack nonravens (confirming instances of the contrapositive) hold the same—or any—weight. And the puzzle of the ravens is to give a legitimate reason for this empirical instinct.

The requirement of total evidence is the key to unlocking the puzzle. Were our knowledge of the universe so poor that black ravens, nonblack ravens, black nonravens, and nonblack nonravens were nothing but data points, then it would be proper to confirm as the paradox suggests.

We know too much about ravens to confirm that way. Someone finds an albino crow (a counterpart of the 99-foot man). It's a non-

black thing and a nonraven. Far from confirming the "ravens are black" theory, it would cast strong doubt on it. Crows are in the same genus as ravens. If crows are prone to albinism, then possibly ravens are too. This background information negates the confirmation.

More generally, we know that ravens bear many, many more similarities to related birds than they do to red herrings or blue lawn elves. When this totality of evidence is taken into perspective, we realize it is a waste of time to examine nonblack nonravens. Whether all ravens are black is an issue best decided by observations of ravens and their relatives and by studies of biological variability.

Arguments from number of ravens vs. nonblack things are perhaps misleading. Consider again the case where the universe consists of seven sealed boxes, a case where most agree it is proper to count nonblack nonravens as confirming instances. Is the deciding difference between this and the real world truly one of number?

Picture a universe containing, say, 10^{10} sealed boxes. Most of the boxes contain black ravens; a few contain green crab apples; and maybe there is a white raven or two somewhere. You have opened a great number of boxes and thus far found only black ravens and green crab apples. Opening a new box and finding yet another black raven confirms "All ravens are black"—to a slight degree, for you have already opened a lot of boxes and there are trillions yet unopened.

Would not opening a box and finding a green crab apple slightly confirm the hypothesis as well? For one thing, it means one less possible refutation to worry about. For another, it increases your confidence that the objects you find in the boxes have certain fixed colors. You might even explain your faith in the hypothesis like this: "Every raven I've seen has been black. In fact, whenever I've seen something that wasn't black, it was always a crab apple, never a raven. The crab apples are 'the exception that proves the rule.'"

In this universe of sealed boxes, there is no ornithology, no albinism, no biological variation. In short, there is no background information about the way the world works. Instead of containing real ravens or crab apples, the boxes might as well contain slips of paper bearing the words "black raven," "white raven," etc. Now it is completely reduced to a formal game. If you open a box and find it to contain a slip of paper saying "white crow," there is no way of seeing that it has any different bearing on the hypothesis than "green crab apple."

We all know instinctively that it is wrong to ignore background evidence, but (before Hempel) this important fact went unrecognized in discussions of scientific method. It is unnecessary (logicians say impossible) to deny the equivalence of the contrapositive. Hempel concluded simply that one must be wary of logical transformations of hypotheses. Yes, a contrapositive is equivalent, but confirmation does not always "recognize" logical transformations. The sundry ways in which consequences of inductive beliefs can mislead are the source of many paradoxes. A more troublesome paradox of induction is to come.