Contents

Introduction:

Social Commentary

and I

vm

A young lady phoned me yesterday. She was writing a "profile" of me and needed a few questions answered. I invited her to ask them and she said: "Your writings seem to be very concerned with your views of social ethics. Even in your fiction, for instance, you have the Three Laws of Robotics, which deal with the manner in which robots and human beings should interact. You also have invented the science of 'psychohistory,' which makes it possible to foresee the future and leads people to direct history into desirable channels."

It was very flattering to be taken as a social philosopher, but truth is mighty and will prevail. I had to explain to her the facts as they were.

The Three Laws of Robotics were first invented by me when I was nineteen years old and psychohistory came along when I was twenty-one. My purpose was not

ix

at all social commentary, but merely the writing of interesting and different science fiction stories so that I could have the pleasure of seeing my name in print and (secondarily) so that I could earn enough money to pay my way through college.

Nor, in all the years since, have I ever seen myself as someone whose duty it was to lecture humanity and change the world.

What I have been doing in my writing, through the years of my maturity, has been (a) to continue to write stories and novels that interest me and, if possible, the editors and readers as well; (b) to explain science to the interested public; and (c) to express my personal views and opinions on any subject it occurs to me to do so.

Points a and b are duties I have set myself, and are the serious work of my life. Point c is my pleasure and it is my incredible luck to have found publishers who are willing to let me do this. It sometimes happens that my personal views and opinions sound like social commentary and are taken as such—but I don't write them with that in mind. I'm just pounding the table as anyone would like to.

I have a couple of examples of this in the collection you are now

holding in your hand.

For instance, I have labeled Part II of the collection "Humanity" for the obvious reason: That is what Part II deals with. The first two essays in the section deal with the evolution of humanity. "The Road to Humanity" carries the story from the origin of the Earth to the coming of the first hominid. "Standing lall" continues the story to the coming of Homo sapiens sapiens. Then in "The Longest River" I deal with early historic times.

It's in the fourth essay of the section, "Is Anyone Listening?", that I take up the human plight in the present and pound the table. I have been writing occasional articles on the dangers of overpopulation for over thirty

years now, and in that time the population of the Earth has nearly doubled and is still going up. I cannot understand how it is possible for people not to be aware of the terrible danger we are all in from this endless proliferation of human mouths and human needs. In each article I write on the subject, I stress the problems that inevitably arise out of population increase, and point out by how much everything has grown worse since I wrote my previous article.

And still human beings go forward blindly, pretending that nothing is happening or that if population is going up, it doesn't really matter, which is why I entitled the overpopulation essay in this book with a plaintive question. The United States, in fact, consistently refuses to help any nation or any international organization that is attempting to control population. Why? Because the nation is in the grip of the cavemen of conservatism who interpret all things in the murky light of an irrational ideology.

Once, in one of my earliest essays on the matter of overpopulation, I received a note from someone who said, "I'd say this was God's problem, wouldn't you?" It was the work of a moment to send back the answer, "God helps those who help themselves."

. . . And then, once in a while, I even feel it necessary to express my opinion on subjects concerning which (it is my uneasy feeling) I don't really have any expertise. For instance, modern poetry, for the most part, leaves me cold. I have this notion that modern poets are writing strictly for each other, and that there is no attempt to reach the ordinary intelligent nonpoet. Indeed, I even have the feeling if a poem appealed to nonpoets that would be taken as a certain sign that it was a bad poem.

I remember once reading an essay by a poet who explained that

modern poems were intensely autobio-

xi

graphic. As an example, she presented one she had written which made absolutely no sense to me, although I have this notion that I am able to interpret the English language with reasonable expertise. The poet offered to explain the poem by detailing the autobiographic background to anyone curious enough to write to her, but I didn't bother. I wasn't curious.

You see my feeling is that all writing is a device to transfer ideas from the brain of the writer to that of the reader, and that different types of writing are different modes of doing so. If a piece of writing does not succeed in making any such transfer at all, then it has failed. (But then, I may perhaps be influenced by my own lifetime obsession witfi making some of the difficult concepts of science accessible to nonscientists.)

It seemed to me that poetry could affect people and that some poems did affect them. Quite apart from whether such a poem is "good" or "bad" by the academic criteria of poets, ought not a poem achieve some sort of recognition because it is effective?

I thought it should and I decided I would write an essay on the matter. It is included in this book as the very last item, "A Sacred Poet." I must admit, I wrote it with a certain tremor of uncertainty. Would the Noble Editor sit still for my venturing so far out of my field? For that matter, would the readers tolerate it?

To my astonishment, the editor wrote me a very complimentary letter and I received a richer harvest of comments from my readers for this one essay than for any of the hundreds of other essays in the series. And not one letter condemned it.

I was delighted.

Part I Astronomy

XII

1

The Very Error of the Moon

I suppose I have seen more comments in print about my towering ego than almost anybody. The most recent case (at the present writing) is

in a review of a new edition of my essay collection The Roving Mind (Prometheus Books, 1983), which the reviewer found "exhilarating." He then couldn't resist referring to my opinion of myself, but added a saving clause. He said, "The egotistic Asimov, who has plenty to be egotistic about . . ."

Well, I'll accept that.

The truth is, though, that I am an easy mark. There is such an obvious self-assurance about me that everyone has the ambition to put me in my place, and a sizable percentage of them succeed, and that helps keep me humble. In fact, I am sometimes put in my place when it seems to me there is no chance of its happening. I remember a prize example of this . . .

It was in 1972, I think, when I had just joined the Gilbert and Sullivan Society and was waiting for the festivities to begin. I didn't know the gentleman at my right, who was a bit older than I was, and he clearly didn't know me. A young man came up to me and asked, very politely, for my autograph, which I was glad to give. There then followed this conversation between myself and the man beside me.

STRANGER (curiously): "Why did he ask you for your autograph?"

I (modestly): "I guess he recognized me."

STRANGER (naturally): "Who are you?"

I: "I'm Isaac Asimov."

STRANGER (at sea): "But why did he ask you for your autograph?"

I (sighing inwardly): "I'm a writer."

STRANGER (perking up amazingly): "My son is a writer. He has just published his second book. He has published two novels" (holding up two proud fingers) "on sports."

I: "Wonderful."

STRANGER: "What do you write?"

I (cautiously): "Different things."

STRANGER: "Do you write books?"

I (wishing he'd stop): "Yes."

STRANGER: "How many have you written?"

I (at my wit's end): "A few."

STRANGER: "Come on. How many?"

I (suddenly annoyed, and anxious to put an end to u): "As of now, one hundred and twenty."

STRANGER (totalty unfazed): "Any of them on sports?"

I: "No."

STRANGER (triumphant): "My son has written two novels. On sportsr

I (totaUy crushed): "Wonderful."

Something else that keeps me in my place is going back over my nearly three decades oiF&SF essays and taking note of those that show me to be something less than prescient.

That doesn't happen often, of course. In fact, sometimes I do pretty well. Thus,,in a recent issue of a magazine dealing with astronomy for the layman, a writer wrote about the distant "Oort cloud" of comets and said that "forward-looking scientists" now consider that such comets might someday become "stepping-stones to the stars."

As it happens, I was forward-looking enough to suggest that very thing in my F & SF essay of October 1960, twenty-seven years ago. And, what's more, I called the essay "Stepping-Stones to the Stars."*

Still, one lack of prescience somehow deflates any number of cases in which I was on the ball. Consider my essay "Just Mooning Around," which appeared in the May 1963 issue of F & SF.f In it, I talked about satellites in general, and when I got to Earth's Moon, I pointed out how different it was from other satellites (unusually large, unusually distant, and so on) and I admitted that I couldn't explain how it came to exist.

So let's go over the matter of the Moon in some detail, for now a solution has been thought of to the problem of how it comes to be there—but, to my great chagrin, not by me.

Of course, people haven't worried about this problem until recent times. On the fourth day of the biblical version of the beginnings of the Universe, God said, "Let there be lights in the firmament of the

heaven to divide the day from the night; and let them be for signs, and for seasons, and for days, and years: And let them be for

* See my book Fact and Fancy (Doubleday, 1962).

t See my book Of Time and Space and Other Things (Doubleday, 1965).

lights in the firmament of the heaven to give light upon the earth: and it was so. And God made two great lights; the greater light to rule the day, and the lesser light to rule the night: he made the stars also." (Genesis 1:14-16).

The Moon was the "lesser light" referred to in the verses above and I imagine that, in our Western past, the feeling was undoubtedly general that the Moon was merely a small, nearby lamp hung in the sky for the convenience of humanity, and that the reason it was there was that God put it there.

And yet, as long ago as 150 B.C. the Greek astronomer Hipparchus (190-120 B.C.) had worked out the distance of the Moon from the Earth by valid trigonometric methods and had found, correctly, that that distance was sixty times the Earth's radius.

The Greek scientist Eratosthenes (276-196 B.C.) had already calculated the Earth's circumference correctly, also by trigonometric methods, so that the Moon's real distance was known to Greek scholars as early as the second century B.C.

Modern measurements have somewhat refined the results, and the Moon's average distance from the Earth is now known to be 384,400 kilometers (238,900 miles). For the Moon to seem as large as it does in the sky from this distance means that it must be 3,480 kilometers (2,160 miles) in diameter.

It is not just a small lamp in the sky, then; it's a respectable world. In 1609, Galileo looked at the Moon through his telescope, and saw mountains, craters, and "seas," and in 1969, human beings stood on the Moon. It's a world, all right; it makes as much sense to doubt that as to doubt evolution.

Now the scientific game is to explain how the Moon happens to be in the sky, and to do so by making use of

the laws of nature as we understand them. That's not easy, but if it were easy, it wouldn't be fun, would it?

Among the earliest of those who made a stab at explaining the origin of the Earth without calling on the supernatural for help was Georges de Buffon (1707-88), a French naturalist who wrote a forty-four-volume encyclopedia of natural history. In the first volume, published in 1759, he took up the matter of origins.

A comet, he suggested, struck the Sun and sent some of its substance, together with some of the Sun's substance, flying into space. That flying matter cooled, condensed, and became the planets, including Earth. This, he said, had happened seventy-five thousand years earlier, for it would take that long for the Earth to cool to its present state.

Why a comet, by the way? In Buffon's time, no one knew what a comet actually was, but they sometimes looked very huge in the sky (though that hugeness consists of nothing more than a slightly thick vacuum) and they had orbits that brought them quite close to the Sun. Besides, in Buffon's time, comets were the "in" thing in astronomy, since Halley's prediction of the return of his comet had been fulfilled just before the book was published. Actually, we had best suppose that by "comet," Buffon merely meant "a massive body."

And the Moon? Buffon speculated that it was torn out of the Earth, as the Earth had been torn out of the Sun.

Don't think that Buffon got away with these daring suggestions, by the way. The creationists of the eighteenth century were in power and they did not look kindly on independent thought then, any more than they do today. Buffon was forced to take it all back and to say he had only been kidding.

The year after Buffon's death, however, the French Revolution took place and things eased up, at least as far as disagreeing with creationism was concerned.

Thus, as a result of the two centuries of observations and thought that have taken place since Buffon's time, astronomers are reasonably satisfied that they know how the Solar system started. It began as a vast cloud of dust and gas that may have existed for billions of years, and then suddenly began contracting—perhaps under the impulse of a shock wave from a nearby supernova.

Much of it collapsed toward what was eventually to become the Sun. Outside the forming Sun was a large disk of dust and gas—like those that have recently been found to be surrounding stars such as Vega and Beta

Pictoris.

In 1944, the German astronomer Carl Friedrich von Weizsacker (b. 1912) considered this outer disk of dust and gas, and presented reasons for supposing it to form eddies and subeddies,

These whirling eddies would carry material into collisions in the regions of intersection. As a result of these collisions, larger bits of matter would grow at the expense of the smaller ones. Eventually, the surviving bits would be large enough to be worth the name planetesimals ("small planets*'). With continuing collisions the larger planetesimals would sweep up the smaller ones until today's planets were formed. They would be separated by larger and larger distances as one went outward from the Sun, since the eddies themselves had been progressively larger with distance.

In the outer Solar system, where cooler temperatures allowed more of the very light and very plentiful elements hydrogen and helium to be collected, the planets grew large in consequence, and around them smaller eddies formed which gave rise to satellites.

The formation of the Solar system began, it is clear, about 4.6 billion years ago and it had reached essentially its present shape by 4 billion years ago.

Earlier versions of this condensing-nebula origin of the Solar system, some dating back to 1755, had come a cropper over the question of angular momentum (which is a measure of all the turning motions such as rotation about an axis and revolution about a center of gravity). Of the total angular momentum of the Solar system, the Sun (with 99.9 percent of the total mass of the system) has but 2 percent. The planets have the other 98 percent. Jupiter, alone, has 60 percent of the total.

Nobody could figure out how all that angular momentum could be crowded into the planets and for a long time astronomers had given up the condensing cloud bit. After von Weizsacker's new analysis, however, a Swedish astronomer, Hannes AlfVen (b. 1908), took the Sun's magnetic field into account. As the forming Sun whirled rapidly, its magnetic field twisted into a tight spiral and acted as a brake. The angular momentum couldn't disappear; it could only be transferred to the planets, which were forced into orbits that were farther from the Sun.

Even after the planets and satellites were just about formed there were still a few planetesimals to be swept up. On those worlds that lack atmospheres, we can still see the marks of those last impacts. The

craters on the Moon are most familiar to us, and in this era of rocket probes, we have found craters also on Mercury, Mars, Phobos, Deimos, Ganymede, Callisto, and other worlds.

Even today, there are objects such as comets, asteroids, and meteoroids that have orbits that make them potential dangers.

But let's get on with the Moon.

A large planet, such as the four outer gas giants, might form satellites as the Sun formed planets, so we expect Jupiter, Saturn, Uranus, and Neptune to have many satellites, some of them quite large—and ring systems, too.

But Earth? Earth is a small planet, so why should it have a satellite—and such a large one? Of the other small inner planets, Mars has two tiny satellites that are obviously captured asteroids, while Venus and Mercury have nothing at all. Why does Earth have one?

There would seem to be three alternative explanations:

1. Earth formed as a single body, but then split in two for some reason, forming the Moon.

2. Earth and Moon formed separately, but out of the same eddy of dust and gas. They have always been separate worlds, but the Moon has always been a satellite.

3. Earth and Moon formed separately, but out of different eddies so that Moon was once an independent planet, which was, however, captured by the Earth.

Alternative 2 must have happened at the very start. Alternatives 1 and 3 happened after the start but must have been catastrophic enough to wipe out any life that had gotten started. Life goes back uninterrupted for at least 3.5 billion years, so those alternatives must have happened, if they had happened at all, before then.

In 1879, the English astronomer George Howard Darwin (1845-1912)—the second son of Charles Darwin—attempted a rational explanation of the Moon's origin for the first time since Buffon. Darwin began with the following situation, which was already well-known in his time. The Moon sets up tides on the Earth, and the surface of the Earth, as the planet turns, moves progressively through the two tidal heaps of water on opposite

10

sides of the Earth (see "Time and Tide," F & SF, May 1966).$

As it does this, the water scrapes against the shallower sea bottoms and converts some of the energy of rotation into heat by friction. This slows the Earth's rotation to a very slight degree, lengthening the day by one second every 62,500 years.

This is not much, but it decreases the angular momentum of the Earth, which can't be destroyed and which must therefore be transferred to the Moon, which is being forced away from the Earth very slowly as a result.

Darwin pointed out that if one imagined the flow of time reversed, one could imagine the Moon to be slowly approaching the Earth, and angular momentum shifting from the Moon to the Earth, so that the Earth would be gaining speed little by little. As the Moon continued to approach the Earth, the tides would increase and the backward spin of time would see the Moon approach Earth and Earth gain speed more quickly. Finally, the Moon would reach and coalesce with the Earth, which would be spinning very rapidly indeed.

Now let time flow forward again. The Earth is spinning very rapidly and the result is an equatorial bulge much greater than the one Earth has now. Since the Earth would be warmer and softer in those early days, the bulge would be all the greater and a piece of it would finally break off and move away from the Earth. What was left of the Earth would have lost enough angular momentum to slow down markedly and it would be stable thereafter.

This would explain several things. The Moon has only three fifths the density of the Earth—but it pinched off the outer layers of the Earth (the rocky mantle), which

£ See my book From Earth to Heaven (Doubleday, 1966).

11

has just that low density. The high-density metal core of the Earth remained untouched.

Then, too, the Moon has just the width of the Pacific Ocean. Could it be that that was where it pinched away, leaving the basin behind, encircled by the "ring of fire" (the volcanoes and earthquakes that rim the Pacific) as the still-unhealed wound of that rupture?

It sounded very good at the time, but we now know that the Pacific

Ocean bit is all wrong. The ocean's shape and the ring of fire are explained by modern plate tectonics and have nothing to do with the Moon. What's more, if all the angular momentum of the Earth-Moon system were squeezed into the Earth alone, it wouldn't have enough spin to throw off the Moon. It wouldn't even be close. The total spin is only one fourth of that which would be required. Darwin's theory just won't work, therefore, and astronomers seem quite agreed that alternative 1 is out and that Earth and Moon were never a single body.

What about alternative 2? Might not Earth's eddy have had two nuclei so that two worlds developed, and done so far enough apart never to meet and coalesce? It might be much more usual for a single nucleus to collect the overwhelming amount of matter in its eddy, but unusual things happen sometimes, and the Earth-Moon system is certainly unusual.

After all, the four large satellites of Jupiter, taken all together, are only 1/5000 the mass of Jupiter. All of Saturn's satellites, taken together, are about 1/4000 the mass of Saturn. The Moon, on the other hand, is 1/80 the mass of the Earth, and perhaps that is the sign that we just happened to be the victim (or the beneficiary) of an unusual case in which there was a double nucleus.

In fact, we now know we aren't even the only case of

12

this. In 1978, Pluto's satellite, Charon, was discovered, and it turns out that Charon is about one tenth the mass of Pluto. To be sure, Pluto and Charon are much smaller than Earth and Moon are, and they are icy, in all likelihood, while we are rocky. It may be unsafe to draw comparisons. Still, it is possible that Pluto and Charon are another example of two nuclei in the same eddy.

Still, if that were so, Earth and Moon should have roughly the same composition. It's not reasonable to suppose that virtually all the iron in the cloud was on our side and practically none on the Moon's side. Yet the Earth has a large liquid-iron core, and the Moon has none. That is why the Moon has a density that is only three fifths of ours. Such a density is explained by alternative 1, but not by alternative 2, so the latter seems to go a-glimmering also.

What about alternative 3, that the Moon was originally formed in a different eddy?

Presumably it was formed in an eddy that was closer to the Sun than ours was. That would explain why the Moon seems to be covered with

glassy bits, although natural glass is very uncommon on Earth. It may be that the Moon was exposed to much more heat.

That would also account for the fact that the Moon is

lower in the content of volatile elements than the Earth

is. It's not only that it's short of carbon, hydrogen, and

. nitrogen, but also that it's short of metals like sodium,

^potassium, tin, and lead. Again, it has been exposed to

much more heat.

That might also account for the fact that it is so short of iron. Perhaps the eddy in which it was formed had less iron to begin with so that it ended up being formed almost entirely out of rock.

13

Actually, none of this is entirely compelling. Venus and Mercury have iron cores, so that those eddies closer to the Sun than ours obviously had plenty of iron. But if the Moon were formed in an eddy farther from the Sun than ours was, why doesn't it have volatiles—at least the

metallic ones?

Worse than all this is the fact that it is not easy for one body to capture another, particularly if the other is itself a large body. We might imagine the Moon to have a very elliptical orbit to begin with, swinging toward Mercury at one end and toward Earth at the other. This would be hard to explain, but assuming it to be so and supposing that the Moon were to approach the Earth rather closely, it would swing about it in a hyperbolic orbit and speed away. Its orbit would be changed but it would not be captured. Indeed, astronomers have tried to work out some set of circumstances whereby the Moon would be captured by the Earth and have failed to do so in any credible way.

As a result, alternative 3 doesn't look good, either.

This has frustrated astronomers in a way that reminds me of Othello's saying about the Moon under different circumstances: "It is the very error of the moon . . . / And makes men mad."

One astronomer is reported to have said, in total exasperation, "When we consider the various ways in which the Moon might have

been formed, and how unsatisfactory they all are, the only conclusion we can come to about the Moon is that it isn't there."

Well, then, what are we going to do? If only three alternatives are possible and if every one of the three is eliminated, are we forced back to creationism?

No, that pitch of desperation we have not reached. What we need is a fourth alternative. It may be that the

14

three I've mentioned are not, after all, the only ones possible.

Fortunately, as early as 1974, William K. Hartmann of the Planetary Science Institute in Tucson, Arizona, (along with some coworkers) did suggest a fourth alternative.

Suppose we go back to alternative 2. Let's suppose that as the planetesimals accreted into a planet in Earth's orbit, they did accrete into two bodies. The second, smaller body, however, was not the Moon. There's the point that everyone seems to have missed.

It was a second body just like the Earth in chemical composition, since it was formed out of the same eddy. It did have a metal iron core, just as Earth does, and it had the same volatile materials Earth had. What's more, it may not have been as small as the Moon. It may have been the size of Mars, or a bit larger, with a mass from one tenth to one seventh that of the Earth. We would then have been a truer double planet than even Pluto and Charon are.

But what happened to this companion of the Earth, which was not the Moon? Well, the two objects may have revolved about a common center of gravity, but in a quite elliptical way, which would mean a close approach each revolution. There were still somewhat smaller planetesimals about and both bodies may have been struck this way and that by them so that they underwent a kind of Brownian motion on a cosmic scale. That would give them both rather erratic orbits and the two worlds may have collided glancingly, at some time more than 4 billion years ago, at a mutual speed of eight to ten kilometers (five to six miles) per second.

In less than an hour, the deed was done and a portion of the outer layers of each object was smashed and sliced off, and shattered, and in part vaporized, and

15

launched into space. What was left of both worlds then coalesced to form the Earth as it now is.

Observe the consequences. The two metal iron cores remained put and when the two planets coalesced, they formed one core, so that Earth's present core is a combination of both original cores.

The smashed layers that were hurled into space might, to some extent, have eventually pattered back to Earth, or, in part, escaped permanently. That portion, however, which had vaporized could condense and eventually collect into a single world.

That new world would have been formed only out of the outer layers of the colliding worlds, out of the rocky mantles, and it would have no metal iron core worth mentioning. It would have a density of only three fifths that of Earth. What's more, the amount that coalesced would in no case be as large as the original companion. With so much of the interloper fusing with Earth and with so much of the sliced-off portion coming back to Earth or drifting away altogether, the Moon that finally formed would only be about one tenth the mass of the original proto-Moon.

Finally, the sliced-off portion of the outer layers would have been subjected to the heat produced by the collision, and when the vapors condensed, those of the volatile elements did so to an unusually small extent. That would explain why the Moon is short of volatile elements and long on glassy remnants.

In short, this alternative 4 avoids all the difficulties associated with the other three, and seems to introduce no major difficulties of its own.

Even so, Hartmann's 1974 suggestion was largely ignored. Scientists don't like catastrophic solutions that seem to depend on the happening of some low-probability event. Slow and inevitable evolutionary solutions

appeal to them much more.

16

After 1974, however, computer simulations were made of the situation and what showed on the computer screen seemed quite good. In 1984, when the idea was advanced again with computer simulations as backup, there was considerable enthusiasm. Pending a closer look at every stage of the supposed impact, astronomers now

think they have a way of accounting for the existence of the Moon.

And now I've got to explain my personal chagrin at falling short of prescience.

If you'l! think about it, alternative 4 is exactly Buffon's idea of two and a quarter centuries ago. He had Earth formed by a glancing collision with a Sun of a smaller but, nevertheless, massive body. He then suggested that the Moon was torn from the Earth, and one would have to assume that he was thinking of the same mechanism.

Well, then, the Moon may well have been formed in that way and, since Buffon had suggested it, and I knew about the suggestion, and it wasn't one of the three alternatives that I had brainwashed myself into thinking were the only ones possible—why didn't / see that Buffon was offering us all the fourth alternative, and suggested it twenty-seven years ago?

On the other hand, it makes me aware that there are limits to my "smartness," and that realization may be healthy for that supposedly swollen ego of mine.

17

necessary to look in various dictionaries, but I had anticipated that that was something that should be made clear. If you plan to explain science, you have to have a feel for asking the right question. I asked the right question once, many years ago.

Asking the Right Question

In the March 26, 1987, issue of New Scientist, a story is told of a chemist who was lecturing to a bunch of youngsters on the chemistry of matches. When he was done, he asked for questions, and one of the youngsters (and I'm willing to bet he was that nemesis of all lecturers, the bright twelve-year-old) said, "Why are matches called matches?" and the lecturer was instantly

stumped.

I laughed aloud at this, because I knew that the April 1987 issue of F & SF was on the stands that very moment, and that it contained my essay "The Light-Bringer."* In it, I happened to discuss matches, and I had, indeed, explained why matches are called matches.

It wasn't a very difficult thing to do, since it was only

* See my book The Relativity of Wrong {Doubleday, 1988).

18

Back in the October 1959 issue of F & SF, I had an essay entitled "The Height of Up."f In it, I discussed temperature. I explained the existence of absolute zero, at which motion reached its minimum point, and said that that was as low as temperature could go. I then asked whether there was a point that was as high as temperature could go?

I decided that a single proton, if one squeezed all the energy of the Universe into it, would have a velocity that would be the equivalent of a temperature that would be something like 3.6×10^{12} K (3.6 trillion degrees absolute). At that velocity, however, its mass would increase markedly and that would drive the temperature still higher. I ended by concluding that there was no upper limit to temperature.

My calculations were very primitive and I'm sure not valid, but apparently I had asked the right question, for a young man named Hong-Yee Chiu, who was studying at the Laboratory of Nuclear Studies at Cornell University, read the essay and it caught his imagination. He sent me a letter dated August 26, 1959, in which he tackled the question himself in a far more sophisticated manner than I could.

He concluded that the maximum temperature of the Universe was not infinite but merely enormously high. It was something, he said, like 10^{91} K.

However, Hong-Yee Chiu could not let go the prob-

t See my book View from a Height (Doubleday, 1963).

19

lem. He got his Ph.D. in elementary particle physics but found he kept thinking about the matter of how high a temperature we could have—not in the manner of supposing all the energy of the Universe to be squeezed into a single particle, but in real situations. In other words, if we searched through the Universe right now, as it is, what would be the highest temperature we are likely to find.

Obviously, the temperature at the center of a star is a lot higher than anything in our neighborhood. The central core of our Sun has a temperature of about 1.5×10^7 K (15 million degrees). TTiere are, however, stars more massive than the Sun, and the more massive a

star, the hotter its central core. What's more, as a star ages, the core gets still hotter. Therefore, the highest temperature must be at the center of a giant star that is so old and so hot that it explodes. Hong-Yee Chiu found himself asking what the temperature was at the core of a star at the moment it goes supernova.

He promptly switched fields of research and began to apply his knowledge of subatomic physics to the astrophysics of supernova. (He had no hesitation in placing the responsibility for the switch on me and it made me quite nervous, I assure you.)

He calculated the types of nuclear reactions that would take place as the temperature at the core got higher and higher. There were nuclear fusions, as small nuclei added to each other and grew larger, releasing energy in the form of photons of radiation and those little particles called neutrinos that go through matter as though it weren't there.

The neutrinos, naturally, streaked out of the core at the speed of light and left the star (even if it were a red giant) in a matter of minutes, but they carried off only a small fraction of the total energy being released, for most of the energy was carried by the photons. The pho-

20

tons were endlessly reabsorbed and reemitted and leaked out of the star very slowly indeed.

Hong-Yee Chiu found, however, that, according to his figures, a temperature was reached at which photons reacted with each other to produce neutrinos. For the first time, neutrinos became the dominant form of particle at the stellar center, and they all left at the speed of light, carrying the energy with diem. The central core's temperature plummeted and was no longer capable of keeping the star extended. The star collapsed and all the remaining hydrogen in the outer layers fused at once to produce a supernova.

Hong-Yee Chiu's calculations led him to believe that this took place at a temperature of 6×10^9 K (6 billion degrees), which is four hundred times the temperature of the Sun's core, and that this is the maximum temperature we are likely to find anywhere in the Universe today.

He sent me a letter dated November 14, 1961, describing his findings, which he published in Physical Reviews and in Annals of Physics, and I wrote about it in my essay "Hot Stuff," which appeared in the July 1962 issue of F&SF4

Clearly, this was a potentially important finding to have arisen out of my having asked the right question. Detecting a spurt of neutrinos from the sky might be a herald of a supernova about to blaze out, and from the neutrinos some of the details of the explosion might be worked out.

Unfortunately, it's not all that easy. Neutrinos are extremely difficult to detect. In order for one to be detected, it has to interact with some other particle, and

t See my book View from a Height (Doubleday, 1963).

21

neutrinos do that only very rarely indeed. As far as neutrinos are concerned, in fact, matter is just a high grade of vacuum. Only one out of many trillions of neutrinos manages to hit any other particle squarely enough to

interact.

Thus, though the existence of the neutrino was made theoretically plain in 1931 by the Austrian physicist Wolfgang Pauli (1900-58), it wasn't actually detected till 1956, twenty-five years later.

The detection was carried through by two American physicists, Clyde L. Cowan, Jr. (b. 1919) and Frederick Reines (b. 1918). They reasoned that the best chance of detecting a neutrino was to put their detecting device into the midst of a very dense stream of them. Such a stream would emerge from a nuclear fission reactor in operation. (A fission reactor releases antineutrinos rather than neutrinos, but that doesn't matter. If one exists, the other must.) The scheme worked.

Was it possible to detect neutrinos from the heavens, however? Whereas fission reactions release antineutrinos, fusion reactions release neutrinos, and there are fusion reactions going on at the core of every star. Every star is therefore a neutrino source.

The neutrinos are emitted by stars in every direction, and, as they travel outward, they spread out over the surface of an ever-enlarging imaginary sphere. From any particular star, the number of neutrinos that manages to pass through the space occupied by a detecting device decreases as the star is farther and farther away. What's more, it decreases as the square of the increasing distance.

Imagine two stars, A and B, with A ten times as far

away as B. If the two are releasing neutrinos at the same

rate, then the number of neutrinos arriving from A, the

more distant star, is only 1/100 of those arriving from B.

There is no chance, then, that the number of neutri-

22

nos being emitted by any normal star is large enough to deliver a useful number across light-years of space. Even the Alpha Centauri stars, which are only 4.3 light-years away, are too far away to deliver enough neutrinos to give us a reasonable chance of detecting even one.

This leaves us the Sun, which is only 1/250,000 the distance of Alpha Centauri. The Sun delivers about as many neutrinos as the Alpha Centauri stars do, but it is so close that we should get 625 million neutrinos from the Sun for every neutrino we get from Alpha Centauri.

Solar neutrinos were indeed detected and continued to be detected for some fifteen years, but only in about one third the number that physicists had expected. (This constitutes the "mystery of the missing neutrinos.")

Until 1987, then, neutrinos had been detected with origins in only two different bodies—the Earth and the Sun. No neutrinos originating anywhere else have been detected. Till now.

Back in 1961, remember, Hong-Yee Chiu estimated that as a star approached supernovahood, it ought to produce floods of neutrinos. He estimated, in fact, that a supernova should produce neutrinos at a rate of about a quadrillion times that of the Sun. If that were so, then it might be possible to detect neutrinos arising from a supernova that was not too far away. The only trouble is that waiting for a supernova is a thankless task. . . .

An essay of mine entitled "Super-Exploding Stars" appeared in the August 1987 issue of F & SF, * which was, of course, on the stands in July.

In the essay, I pointed out that in the last half century, astronomers had spotted and studied about four hundred supernovas, all of them in distant galaxies. (A

* See my book The Relativity of Wrong {Doubleday, 1988).

23

supernova, as I explained, is so bright that it can be seen as far as a galaxy can.) I also pointed out that no supernova had been spotted in our own Galaxy in almost four hundred years.

The last supernova visible in our own Galaxy blazed out in 1604, and was studied by Johannes Kepler (1571-1630), This was five years before the telescope was used for the first time to observe the sky. Since then, the nearest supernova to ourselves appeared in 1885 and it was in the Andromeda galaxy. It was so distant that it wasn't even visible to the naked eye. (Neither was any other supernova that has appeared since 1604.) I ended the essay thus:

"While no sane person would wish a supernova to erupt too near the Earth, we would be safe enough if one erupted, say, two thousand light-years away. In that case, astronomers would have a chance to study a supernova explosion in enormous detail, something they would dearly love to do.

"Astronomers are, therefore, waiting for such an event, but that's all they can do—wait. —And gnash their teeth, I suppose."

Those words were actually written on January 7,1987. Exactly forty-eight days later, on February 24,1987 (and nearly five months before the essay appeared in print), the astronomers got their supernova. It wasn't quite in our Galaxy, to be sure, but it was almost as good. Let me explain. . . .

In 1520, an expedition financed by Spain and led by the Portuguese explorer Ferdinand Magellan (1480-1521) was bumping its way down the Atlantic coast of South America. They were trying to find a way of reaching Asia by going west and they had to get past South America.

24

They found no pathway till they reached sub-Antarctic waters and made their way through what came to be called the Strait of Magellan. (What else?) In the process of reaching the strait, Magellan and his men studied the far southern skies, which contained stars and constellations, like the Southern Cross, never visible in European latitudes. Among these new features (to Europeans) were two cloudy patches that looked like detached portions of the Milky Way. These have ever since been known as the Magellanic Clouds. The larger is the Large Magellanic Cloud, the other the Small Magellanic Cloud.

Once the telescope was invented, it quickly turned out that the Milky Way was a mass of myriads of very faint stars—and the same turned out to be true of the Magellanic Clouds. When it was understood that our Sun was part of a huge, lens-shaped Galaxy, it was also understood that the Magellanic Clouds were both galaxies as well.

Galaxies have been discovered in uncounted number, many billions, but the two Magellanic Clouds are the closest of all to our own.

It was by studying the stars in the Small Magellanic Cloud, all more or less at the same distance from us, that the American astronomer Henrietta Swann Levitt (1868-1921) first noted the connection between the luminosity of Cepheid variables and their periods. This gave astronomers a new and extremely powerful way of judging distances. Thus, the main body of our Milky Way Galaxy stretches out over an extreme distance of] about 100,000 light-years.

The Large Magellanic Cloud is about 170,000 light-years from us, while the Small Magellanic Cloud is about 200,000 light-years away. These are not large galaxies like our own. Whereas our Milky Way Galaxy may contain something like 200 billion stars, the Great Mag-25

ellanic Cloud has no more than 20 billion and the Small Magellanic Cloud about 8 billion.

The advantage of the Magellanic Clouds is this: We can study the entire galaxies in greater detail than any others simply because they are closer. Most of our own Galaxy is hidden from us by dust clouds so that we know the Magellanic Clouds, as galaxies, better than we know our own.

An astronomer, lan Shelton, from the University of Toronto, was at an observatory in Chile taking long-exposure photographs of the Large Magellanic Cloud in order to study relatively faint objects in it. On Tuesday, February 24, 1987, he developed a photograph and found upon it a bright star that wasn't supposed to be there. At almost the same time one of his assistants, strolling in the night air, looked up and saw the bright star where none should be.

Soon afterward an astronomer in Australia saw it, and he alerted another astronomer who had happened to take a photo of that very portion of the Large Magellanic Cloud on February 23. At that time, the star in question was barely visible, so there was no question that the star had been spotted within hours of its explosion.

But the first indication came deep underground, in a tunnel under Mont Blanc in the Alps. Down there was a device designed to detect neutrinos and it was run by Italian and Soviet physicists. At 3 A.M. on Monday, February 23, 1987, five pulses of neutrinos were detected within a space of seven seconds.

There was great excitement, of course, since no one down there could imagine what had caused it. When, the very next night, news of the new supernova arrived, the connection was clear. It could not have been coincidence.

26

The crucial point about the supernova was that it was so close: only 170,000 light-years away.

The supernova of 1885 in the Andromeda galaxy which, till then, had been the closest since the invention of the telescope, was 2.3 million light-years away, nearly fourteen times as far as the supernova of 1987 in the Large Magellanic Cloud.

Even Kepler's supernova of 1604, which was the last supernova reported in our own Milky Way Galaxy, was 35,000 light-years away, so that the supernova of 1987 was less than five times as far away as that.

(To be sure, there were closer supernovas before 1604. A supernova which appeared in 1054 was only 6,500 light-years away from Earth. The very closest supernova we know of is one that left behind the vast Gum nebula. It may have been only 1,500 light-years away, but it exploded about thirty thousand years ago. That supernova must have shone with the light of the full Moon for several weeks, but there were only Stone Age men to watch and wonder.)

The neutrinos that were detected on February 23 were the first to have had their origin outside the Solar system. This was hailed by a number of scientists as the birth of neutrino astronomy, but I think that is wrong. As soon as neutrinos from the Sun were detected, that was the beginning of neutrino astronomy. The Sun is a thoroughly respectable star and it certainly qualifies as an astronomical object.

And even the Sun represented only the beginning of observational neutrino astronomy. If we want to include important theoretical work on the subject we ought to go back to Hong-Yee Chiu's work of 1961. After all, his prediction that a supernova would be heralded by a burst of neutrinos at the moment of collapse has been verified exactly. Yet in all I have read about the super-

27

nova so far, I have seen no mention of his name, which strikes me as a peculiar omission.f

There are two kinds of supernovas. Type I is a white-dwarf star in close association with a normal main-sequence star. After the white dwarf absorbs enough mass from its partner, it can blow apart. Type II is a giant star that suddenly gets hot enough to release a flood of neutrinos and collapse. Hong-Yee Chiu's calculations dealt with the Type II variety.

Astronomers studying old photographs of the Large Magellanic Cloud seem to think that the star that exploded is one that was about 30 times the mass of the Sun, 20 times its diameter, and 250,000 times its luminosity. If so, the supernova must be Type II.

Further evidence in favor of this is that the light of the supernova shows strong traces of hydrogen. Giant stars, even those that have aged at the center to the point of supernovahood, still have vast quantities of hydrogen in their outer layers, while white dwarfs have no hydrogen to speak of, but are rich in heavier atoms such as those of carbon, nitrogen, and oxygen.

The supernova is T^pe II, then, and the neutrino emissions are again in line with Hong-Yee Chiu's suggestions.

There is one sort of radiation that is even more elusive than neutrinos. That is gravitational waves, which are streams of speeding particles called gravitons. The existence of these was predicted by Einstein's general theory of relativity and physicists are, on the whole, absolutely convinced they exist.

The trouble is that gravitational waves are incredibly low in energy and, therefore, incredibly difficult to detect; far more difficult to detect even than neutrinos. An

t Hong-Yee Chm was also the first to shorten the phrase quaststellar object into quasar, a now universally used term for a very distant, very active galaxy

28

American physicist, Joseph Weber, made use of aluminum cylinders, five feet long and two feet thick, suspended in a vacuum chamber by a wire, as a detection device. Any gravitational wave washing over such

cylinders would distort it slightly by about the width of a proton. Such a wave from some distant event in outer space ought to be long enough to quiver the entire Earth, so to speak, so that the detecting cylinders in far different locations ought to record a wave at the same time.

In 1969, Weber thought he had detected such waves, but his results could not be repeated by others. What is needed are still more sensitive detectors, and some source of gravitational waves that is very powerful.

The sensitive detectors are being built, and the supernova ought to have released gravitational waves that would have reached Earth with far more intensity than anything else would have in the last few centuries. The only trouble is that none of the gravitational wave detectors are yet working full time, and none happened to be working at the time the supernova of 1987 exploded. Better luck next time, surely—but when will the next time be?

As the light of the supernova fades, its spectrum will be followed in full detail in every way possible, of course, to see what deductions can be made concerning the phenomenon. Still, even after it's all over, it won't be all over, A vast cloud (a supernova remnant) will be left behind, rather like the Crab Nebula, which is the remnant of the supernova of 1054. Tb be sure, the new remnant will be almost thirty times as far away as the Crab Nebula, but, on the other hand, it will be fresh and spanking new. We didn't get the chance to study the Crab Nebula in reasonable detail until it was about nine hundred years old.

Then, too, the supernova may leave behind a pulsar

29

(that is, a neutron star). The pulsar may not be sending its pulses in our direction and it will be much farther away than any other pulsar known, since all those we have so far discovered are in our own Galaxy. Still, if we are lucky and can detect anything at all concerning the pulsar, we will, for the first time, be able to study one that is freshly minted, so to speak.

And if we do not detect a neutron star, that might be because a black hole was formed. Perhaps there may be something there, or in the surrounding neighborhood, that will give us some information concerning such a newly born black hole. Almost anything of the sort would be terribly exciting.

But let's get back to the neutrinos.

In the February 1981 issue of F & SF, I had an essay entitled "Nothing and All"$ in which I discussed neutrinos. There are three types of neutrinos, I pointed out. There is the ordinary neutrino associated with electrons, which can be called an electron-neutrino. There are also moon-neutrinos and tauon-neutrinos, which are associated with muons and tauons respectively, Muons and tauons are like the electron in every respect except that muons are more massive, and tauons are still more massive.

These three types of neutrinos (for each one of which there exists also an antineutrino, of course) seem to be distinct from eackother, but physicists were at a loss to explain what the distinction actually was. All three had no mass so that all three moved constantly at the speed of light. All had no electric charge, all had the same spin, and all seemed to be identical in every measurable quantity.

* See my book Counting the Eons (Doubleday, 1983).

30

On the other hand, what if the neutrinos had a tiny mass, only a small fraction of that of an electron, one that had escaped detection. In that case, the neutrinos might differ very slightly in their masses and this would be the distinction. In such a case, each would travel at slightly less than the speed of light, and would "resonate," changing rapidly from one form to the other.

This meant that as neutrinos sped from the Sun to the Earth, even if the stream consisted of electron-neutrinos to begin with, they would arrive as a mixture of ah" three. The neutrino-detecting device on Earth, geared to detect only electron-neutrinos, would detect far fewer than expected and this would explain the mystery of the missing neutrinos.

Furthermore, since neutrinos are so common in the Universe, even a very tiny mass for each would mean that the Universe would be, in total, at least a hundred times as massive as had been thought. That would account for many puzzles—the manner in which galaxies rotate, the manner in which clusters of galaxies hold together, and so on. It would also mean that the Universe is "closed" and will someday stop expanding and begin to contract.

I was very enthusiastic about this possibility, and I hoped earnestly that the first rather tentative reports that neutrinos possessed mass would be confirmed. However, more than six years have passed since my essay was written, and the confirmation has not arrived. Neither

has it been definitely established that neutrinos do not have mass.

But now comes the supernova of 1987.

For the first time, physicists have picked up neutrinos coming not only from the Sun, but from a supernova that is 90 billion times as far away from us as the Sun is.

If the neutrinos have tiny masses, they must then travel at slightly less than the speed of light. Light

31

reaches us from the Sun in eight minutes, and neutrinos must reach us in slightly more time than that. However, the time difference may be too small to measure, especially since we don't know when particular solar neutrinos started their journey.

In the case of the supernova, however, we know that the neutrinos must have started their journey when the supernova exploded. The light, traveling at the speed of light, would reach us in 170,000 years. (Yes, that means that the supernova "really" exploded 170,000 years ago.) The neutrinos, traveling at slightly less than the speed of light, should reach us later. Even if they were traveling just one mile per second short of the speed of light, the neutrinos would arrive a year later than the light would. If the neutrinos were traveling one yard per second less quickly than light, they would still arrive five hours late.

But the neutrinos didn't arrive late. They arrived at just the time of explosion, as nearly as we can tell, or a little before. This strongly suggests they traveled at the speed of light and therefore had zero mass, or, in any case, too little mass to affect the Universe substantially.

What's more, if the neutrinos had mass, the more energetic ones would travel faster and arrive first. This also wasn't so. As nearly as could be estimated, all the neutrinos arrived at about the same time, regardless of energy. Again, this seems to support the zero-mass view.

That casts me down, but any theory, no matter how close to my heart, must give way in the face of adverse observation.

Out of the Everywhere

32

There is no secret about the fact that I do not view President Reagan's "Star Wars" fantasy in any favorable way. My own feeling is that it is the wish fulfillment dream of a shallow mind and that it cannot possibly work either technologically or politically.

Naturally, being a rational man, I know that it is conceivable that I may be wrong, but I don't think I am. I have written essays giving my reasons in detail, so I'm not going to do so again here. Instead, I will tell you a small incident that came to my mind a few days ago.

At a Nebula Awards banquet some two years since, I was accosted by a fellow science fiction writer of far-right persuasion.

"Hey, Asimov," he said belligerently, "why are you against the Strategic Defense Initiative?" (That's Penta-gonese for "Star Wars.")

33

I felt a little uneasy. The gentleman questioning me was larger than I, younger than I, stronger than I, rather drunk, and a well-known apostle of the righteousness of violence. However, I didn't see my way clear to denying my beliefs, so I said, as calmly as I could, "Because I don't think it will work."

Whereupon my friend rattled off the names of a number of scientists, and said, "These people all believe it will work. Do you doubt their expertise?"

"Not at all," I said. "What I doubt is their sanity."

That left him speechless for a few moments and, under cover of the silence, I slipped away.

It is always with relief, then, that I remember that there are many aspects of science that do not involve political rhetoric. For instance . . .

The story I am now about to tell you begins in the late 1700s, when a French physicist, Charles Augustin Coulomb (1736-1806), noted that if an electrically charged object was suspended by a silk thread it very slowly lost its charge. The charge could not very well leak away through the silk thread since silk is an excellent nonconductor of electricity. Coulomb thought, therefore, that it must quietly and slowly leak away into the air. He was correct, but he didn't know how it happened.

There was no answer to the question until radioactivity was

discovered in 1896. Radioactive atoms are sources of energetic radiation, and such energetic radiation (whether consisting of fast-moving particles or ul-trashort waves) would collide with atoms, forcing the transfer of electrons and thus producing atoms with a positive or negative electric charge (called ions). These ions could interact with an electrically charged object, neutralizing the charge.

Of course, one had to be certain as to whether the

34

ions were really produced by radioactive materials, present in traces almost everywhere in the soil, or were, somehow, produced naturally by the atmosphere.

Shortly after the turn of the century, a Scottish physicist, Charles Thomson Rees Wilson (1869-1959), tested the matter by suspending a charged object in a deep railroad tunnel. The object, he found, lost charge just as it would at the surface. That meant it was not likely that the ions arose as a result of some property in the atmosphere itself, since the vast bulk of the atmosphere was out of reach of the discharging object. The ionization had to result from the presence of trace radioactivity in the rocks all about.

In 1911, an Austrian physicist, Victor Franz Hess (1883-1964), thought the matter could be checked in the opposite sense. Instead of going deep underground to get away from the atmosphere and showing that charge leakage did not stop, why not go high above ground to get away from the soil and rock and show that charge leakage did then stop. There was no point in going up into the heights by climbing a mountain, of course, since then the ground rose with you. One would have to go up in a balloon.

Hess made ten balloon ascensions, therefore, five by day and five by night (and one of the daylight ascensions was carried out during a total eclipse of the Sun). The results he got were unequivocal—and totally unexpected. Although everyone was convinced radioactivity in the soil produced the atmospheric ions and the charge leakage, going up in the air and removing one's self from the soil by some miles actually led to an increase in the rate of charge leakage. The higher one went, the greater the rate of charge leakage.

The soil and its radioactivity might produce atmospheric ions, but so must some mysterious radiation present in the upper atmosphere.

35

Hess hadn't the slightest idea of what this radiation might be, so he called it simply Hohenstrahlung (German for "radiation of the heights")-

As the years went by, however, it became clear that this radiation of the heights came from beyond the atmosphere, from outer space. What's more, it came not from any specific direction, say from the Sun, but from all directions equally. It came from the everywhere into the here; it came from the Universe or Cosmos generally.

Recognizing this fact, the American physicist Robert Andrews Millikan (1868-1953) suggested, in 1925, that the radiation from outer space be called cosmic rays, and that suggestion caught on.

The next question is: What are cosmic rays?

To begin with, all that was known about cosmic rays was the fact that they were extraordinarily penetrating, but that, in itself, was not sufficient to define their nature.

In general, there are two kinds of radiation: (1) streams of particles, and (2) waves. Almost every form of radiation has had a particle-versus-wave controversy.

Sound and light turned out to consist of waves. Cathode rays and positive rays turned out to consist of particles; of electrically charged particles at that. Then came X rays and they were waves. Of the radioactive radiations, alpha rays were streams of positively charged particles, beta rays were streams of negatively charged particles, and gamma rays were waves.

These are not all independent phenomena. Light, X rays, and gamma rays are all examples of electromagnetic radiations (as are ultraviolet, infrared, and radio waves). Cathode rays and beta rays each consisted of streams of fast-moving electrons. Alpha rays and posi-

36

tive rays each consisted of streams of fast-moving atomic nuclei. Of these, gamma rays were the most penetrating. They consisted of electromagnetic waves that were exceedingly short and therefore of very high frequency.

Since cosmic rays were even more penetrating than gamma rays, might cosmic rays be waves of even shorter length and higher frequency? Or might it be that cosmic rays were particles more massive or speedier (or both), and, therefore, of higher kinetic energy

than any other particle streams known?

The problem of distinguishing between the two alternatives was made particularly difficult by the fact that, under careful observation, the difference between particles and waves blurred. In 1905, for instance, the German-born physicist Albert Einstein (1879-1955) showed that light waves had their particle aspects. As particles, they were called photons, from the Greek word for "light."

As time went on, it was found that every wave had its particle aspect, and every particle had its wave aspect. Nor was there any use in asking, "Which is it realty?" It is neither, really; it is both. However, any particular observation you make will demonstrate either the wave aspect or the particle aspect, never both. This is called the principle of complementarity and was advanced by the Danish physicist Niels Henrik David Bohr (1885-1962).

The more energetic a wave and the shorter its wavelength, the more prominent the particle aspect is. In 1923, the American physicist Arthur Holly Compton (1892-1962) showed that in the case of X rays, for instance, the particle aspect was much more prominent than it was for the less energetic photons of visible light. This was called the Compton effect and for it he received a share of the 1927 Nobel Prize in physics.

37

If, then, cosmic rays consisted of ultrashort waves, sufficiently ultrashort to account for their penetrability, then the particle aspect ought to be so prominent that casual experiments would only detect the particle aspect. How could one tell, then, whether cosmic rays were waves acting like particles, or were "real" particles?

There is, as it happens, one difference. All the particles known to science in the 1920s carried an electric charge as an integral characteristic, charges that were either positive or negative. (There are streams of uncharged panicles, too, as, for instance, neutrons and neutrinos, but they weren't known in the 1920s.) None of the waves known in the 1920s (or today, for that matter), whether electromagnetic waves or any other kind, carried any electric charge.

So it boiled down to this: Did cosmic rays carry an electric charge and were they therefore particles; or did they not and were they therefore waves?

As sometimes happens in the history of science, two scientists of roughly equal ability and reputation took up opposite sides of the

question and fought the matter

vigorously.

Millikan thought that cosmic rays were uncharged electromagnetic waves of unprecedentedly short wavelength and high frequency. He thought they were produced in the process of the creation of matter in the far reaches of the Universe and were, so to speak, the "birth cry" of matter.

Compton, on the other hand, was less dramatic, and simply thought that cosmic rays were streams of extraordinarily energetic charged particles.

Thoughts and opinions don't count, however. To settle the matter, evidence was needed; appropriate observations had to be made.

As it happens, Earth has a magnetic field. Electro-38

magnetic waves plunging out of space toward Earth's surface would pass through the field but be unaffected by it since they are uncharged. In that case, if cosmic rays were waves, all parts of Earth's surface would be bombarded by cosmic rays equally.

On the other hand, charged particles plunging out of space toward Earth's surface would be affected by the field in such a way that they would be made to curve toward the magnetic poles. To be sure, cosmic rays are so energetic and travel so quickly that their curvature in response to Earth's not-terribly-strong field would not be very much—but it ought to be measurable. And in that case, Earth's higher latitudes, north and south, ought to be subjected to a slightly more intense cosmic ray bombardment than would Earth's lower latitudes near the equator.

In the interest of research, Compton became a world traveler, measuring the intensity of cosmic ray bombardment at different latitudes. Before the end of the 1920s he showed that the "latitude effect" did exist and that the cosmic rays therefore must consist of electrically charged particles.

The latitude effect, in itself, did not distinguish between positive and negative charges. In 1930, however, the Italian physicist Bruno Benedetto Rossi (b. 1905) pointed out that positive charges ought to be deflected eastward and negative charges westward. The east-west distribution was studied and it became clear that cosmic ray particles were positively charged.

At that time, the only positively charged particles known were atomic nuclei. The proton was the nucleus of the hydrogen atom; the alpha particle, the nucleus of the helium atom; and more complex positively charged particles were the nuclei of more complex atoms. However, let's leave the actual makeup of cosmic ray parti-

39

cles for later on and turn, now, to the uses of cosmic rays.

In the 1930s and 1940s, cosmic ray particles were far more energetic than anything in the way of particles or waves that could be produced in the laboratory. This meant that for something like a quarter of a century, a deeper understanding of nuclear physics depended on cosmic ray observations.

In 1930, for instance, the English physicist Paul Adrien Maurice Dirac (1902-84), as a result of certain theoretical studies, suggested that subatomic particles might exist that were the mirror images (so to speak) of particles that were already known. For instance, the electron might have a mirror image, an antielectron, that would be identical to the electron in all its properties except that it would have a positive charge rather than a negative one.

No such antielectron was known in nature and Dirac's suggestion was not taken very seriously at first.

However, an American physicist, Carl David Ander-son (b. 1905), was at this time working with Millikan and studying cosmic rays at mountain heights, where they were particularly intense. Anderson was working with cloud chambers, devices that marked the path of charged particles by a line of tiny water droplets. Since the devices were placed in a magnetic field, the charged particles followed a curve. From the nature of the curve, from the density of the drops, and from other characteristics, a skilled observer tike Anderson could tell at a glance what was happening inside the chamber.

But cosmic ray particles themselves, and those they produced by collision with the molecules of the atmosphere, were so energetic and speedy that they produced paths that hardly curved at all.

40

Anderson therefore placed a lead bar across the middle of the cloud chamber. The particles associated with cosmic ray activity were energetic enough to smash into the lead bar and force their way through. In the process, however, they lost much of their energy.

When they emerged, therefore, their paths curved more sharply in response to the magnetic field.

One such curve, noted by Anderson in August 1932, was easily recognized as that of an electron—but one that curved in the wrong direction. It was a positively charged electron; one of Dirac's antielectrons. (Anderson, unfortunately, called it a positron and the name stuck, but, properly speaking, it is an antielectron.)

The antielectron indirectly demonstrated the existence of all other antiparticles and of "antimatter" itself.

Now Dirac's work suddenly gained enormous significance and he received the Nobel Prize for physics in 1933.

Here's another example. By 1932, it was known that, except for the simplest atomic nucleus of all, that of hydrogen-1, atomic nuclei were made up of a number of protons and neutrons, all squeezed together into a tiny object only a ten-trillionth of a centimeter across—a diameter only 1/100,000 that of an atom.

This created a prime puzzle. Neutrons, which had just been discovered, were much like protons, but they were electrically neutral (hence their name) and carried no charge. Neutrons did not attract each other and neither did they attract protons. Protons, on the other hand, all carrying positive charges, repelled each other violently.

If the particles within a nucleus showed no attractions among themselves, and did show repulsion, what kept them together?

41

Obviously, there had to be an attractive force present, and it had to be far stronger than the electromagnetic interaction that caused protons to repel each other. This attractive force came to be called the strong interaction, therefore, and proved to be over a hundred times as strong as the electromagnetic interaction.

But how did the strong interaction work? By that time, it was felt that interactions worked by means of exchange particles. That is, particles of certain types exchanged other particles constantly and rapidly so that the result was an attraction or, sometimes, a repulsion. Thus, the electromagnetic interaction, which could show either attraction or repulsion, was mediated by the rapid exchange of photons, while the gravitational interaction, which showed attraction only, was mediated by the rapid exchange of gravitons.

If an exchange particle had no intrinsic mass, then the result was a "long-range interaction." Thus, since photons and gravitons have no intrinsic mass, the electromagnetic and gravitational interactions declined in intensity only as the square of the distance and could make themselves felt over astronomical distances.

The Japanese physicist Hideki Yukawa (1907-81) tackled the problem of the atomic nucleus and its strong interaction. That strong interaction dropped so rapidly with distance that it could barely reach across the width of a nucleus and was not felt at all outside the nucleus. (That is why nuclei have to be so small.) In order for an interaction to be so short range, the exchange particle must have mass. Indeed, by 1935, Yukawa had estimated that the exchange particle must have a mass roughly two hundred times that of an electron, or one ninth that of a proton. Such a particle of intermediate mass came to be called a meson, from a Latin word for "intermediate." Again, no such intermediate-sized particle was

42

known, but almost at once, Anderson, still studying the tracks produced by cosmic rays, detected such a particle. (For this and for the earlier antielectron, Anderson got the Nobel Prize for physics in 1936.)

Anderson's particle was a meson, all right, in terms of mass, for it was 207 times as massive as an electron. It was not Yukawa's meson, however, since it didn't interact with protons and neutrons at all, and Yukawa's meson would have had to interact eagerly.

In 1947, however, an English physicist, Cecil Frank Powell (1903-69), discovered a slightly more massive meson, 273 times as massive as an electron, among the debris produced by cosmic ray particles, and that was Yukawa's meson. As a result, Yukawa received the Nobel Prize for physics in 1949, and Powell received it in 1950.

The two mesons, Anderson's and Powell's, were naturally given different names. Each was given a Greek-letter prefix. Anderson's became the mu-meson and Powell's the pi-meson.

The mu-meson turned out to be identical to the electron in every respect except for its greater mass. This puzzled physicists and still does, for there seems no reason for the mu-meson's existence. As the Austrian-American physicist Isidore Isaac Rabi (1898-1988) said, on considering the mu-meson, "Who ordered that?"

Indeed, such is the close identification of the mu-meson with the

electron, and so fundamentally different is it from the pi-meson and other mesons since discovered, that the mu-meson is no longer even called a meson. Its name has been condensed to muon and, of course, just as there is an antielectron, so is there an antimuon. And just as the electron and antielectron are closely associated with electron neutrinos and electron antineutrinos, so are muons and antimuons associated with muon neutrinos and muon antineutrinos.

43

Electron neutrinos and muon neutrinos are both without mass and without charge and, indeed, have no distinguishing characteristics we know of, but they do not substitute for each other in nuclear reactions so there must be some difference.

In 1977, a still more massive electron was discovered. It would nave been called the tau-meson at an earlier time, but now it is simply called the tauon. It is about 3,500 times as massive as an electron (and, therefore, twice as massive as a proton), but it still has all the electron's properties otherwise. There is also an an-titauon and, of course, a tauon neutrino and a tauon anti-neutrino.

The electron, muon, and tauon, with their neutrinos and antiparticles, make a total of twelve particles in all that are lumped together as leptons. Tin's is from a Greek word for "weak," because these particles are not subjected to the strong interaction, but to a much weaker, even shorter-range, force called the weak interaction,

These twelve leptons may be all there are of this variety of particle. They all seem to be fundamental particles hi that they don't appear to be made up of still simpler entities (as protons, neutrons, and pi-mesons are).

Now let's get back to cosmic rays.

Cosmic rays approaching Earth from outer space are speeding atomic nuclei. This is the primary radiation. The primary radiation doesn't reach us down here at the surface, however. It strikes the upper atmosphere, smashes into its atoms, and produces speeding secondary radiation. It is this secondary radiation that reaches us, and it is mostly in the form of muons.

Here we have a puzzle. The muon can be produced

44

by nuclear reactions in the laboratory, and we can note the length of

the path that a muon takes through a detecting device. After it travels a short distance, its path is converted into one that is typically that of an electron. The conclusion is that a muon is unstable and, after a short period of time, decays into an electron (which is stable). From the length of the muon's path, and its velocity, we calculate that its lifetime is about 2.2 mil-lionths of a second.

Now, how much distance can a muon travel before it is converted to an electron? That depends on its speed, but even if it travels at the speed of light, the fastest possible, it can only cover a distance of 660 meters (two fifths of a mile) before changing into an electron. Yet the muons are formed many miles high in the atmosphere. How can they possibly survive long enough to reach the surface?

That's where Einstein's theory of special relativity comes in. Einstein suggested that as velocities increase, lengths in the direction of velocity decrease. At ordinary velocities, which are only a small fraction of the speed of light, the decrease is immeasurably small. As velocities increase, the decrease in distance becomes noticeable, and at nearly the speed of light, distances become very short.

A meson in the laboratory moves comparatively slowly so that it travels only a very short distance before decaying. A meson hurled downward by a cosmic ray is traveling at very nearly the speed of light and the distance between itself and the Earth's surface shrinks to less than a hundred meters so it has plenty of time to reach the surface before decaying.

But that's the way it looks to the muon. To us, the distance seems to be many miles, so why is it we see the muon make it?

Well, another part of Einstein's theory says that when

45

an object is moving very rapidly relative to ourselves, the passage of time on that object seems to us to slow down. At speeds near the speed of light, time seems to creep.

Since the muons are traveling at nearly the speed of light, the rate of time passage for them seems to us to be very slow, and the allotted 2.2 millionths of a second stretches out a hundredfold and more, giving the muon ample time to reach the ground before its far-extended lifespan comes to an end.

The mere fact, then, that the secondary radiation of muons reaches us is a strong confirmation of Einstein's theory of relativity.

What's more, this business about a shortening distance and a slowing rate of time (and an increasing mass also, by the way) is unbreakably linked with the further consequence of the theory that the speed of light in a vacuum is an absolute maximum for any object possessing mass (objects such as ourselves and our spaceships).

Suppose, then, that someone says to you, "How do you know that we can't go faster than the speed of light? They broke the sound barrier and someday they'll break the light barrier."

In that case, you can answer, "The mere fact that the muons are formed high in the atmosphere and reach us unchanged here at the surface of the Earth demonstrates that we can't move faster than the speed of light in a vacuum." (Of course, you will then have to explain why one implies the other, and this might take time.)

This, however, did not wipe out interest in cosmic rays. The emphasis merely shifted. Instead of concentrating on the nuclear reactions that cosmic rays can induce, scientists began wondering about what cosmic rays could tell us of the outer Universe.

We pass, in other words, from the unimaginably small to the unimaginably large—and we do it in the next essay.

In the 1950s, physicists worked up particle accelerators that could produce speeding particles so energetic that there was no longer any necessity of turning to cosmic rays as the only phenomenon energetic enough to answer questions arising out of nuclear physics.

46

47

Into the Here

Every once in a while I overlook the obvious and I am always grateful at such times if my dear wife, Janet, saves the situation by not overlooking it.

Not very long ago (as I write this) I received a phone call, at a little before 5 P.M., from a newspaper that wanted me to do four hundred words for them on Reagan's "Star Wars" program so that they could run it on the editorial page.

"Sure," I said. "When do you want it by?"

"We need it phoned in by 3:30 P.M. tomorrow, though if you need more time, we can squeeze out another half hour, I suppose."

That faced me with a dilemma. Janet and I were about to leave for a banquet, to be followed by a theater show, and it was quite certain we wouldn't get back till well past 11 P.M. (which means, well past our bedtime).

48

The next day was Tuesday, when I make my rounds to various editorial offices, and I usually don't get back till pretty close to 3:30 P.M.

However, I hate to say no to any reasonable writing request, so I said, cautiously, "I'll try," and hung up.

Then I went to Janet and told her the sad tale. I said, "I'll have to do it either before I go to bed or immediately after I wake up and you will have to call them up and read them the essay sometime during the day."

She looked at me out of her cool blue eyes and said, "How long will it take you to do it?"

I thought a moment. To type four hundred words would take me four to five minutes. Add a little thinking here and there, plus a little editing . . .

I said, "Fifteen minutes at the outside."

She looked at her watch. "We've got almost an hour before we have to leave."

"That's right!" I said. (That hadn't occurred to me.)

So I sat down, knocked off the piece, called up the newspaper, read it to them while they typed it out, and then listened as they read it back to me. We got to the banquet in plenty of time and the essay appeared on the editorial page in due course.

When I was finished with the reading and the re-reading of the essay, by the way, I sighed with relief and said to the man at the other end of the phone, "Now you know my secret. I'm fast."

And he said, "That's no secret."

. . . These essays for F & SF are just ten times as long as the newspaper piece so, at the same rate, they should take me two and a half hours. However, I'll be frank with you—these essays take a little more research per word, a little more thinking per word, and a little more editing per word, so it takes me extra time.

But the essays are worth the extra time, because they're more fun, too.

49

In the previous essay, I discussed cosmic rays, which come out of the everywhere into the here. I ended the story with the 1950s, when human beings had invented particle accelerators that could produce particles every bit as energetic as most cosmic ray particles. That meant people no longer had to fool around with cosmic rays out in the field in order to get them to induce nuclear reactions that might manufacture new and unusual particles. That could be done, instead, in the comfort of the laboratory.

This is not to say that some cosmic ray particles are not more energetic than anything the accelerators of the 1950s—or of the 1980s—can produce. Indeed, the most energetic cosmic ray particles are more energetic than anything we can reasonably hope to produce in the foreseeable future—something I'll get back to later. However, the more energetic the cosmic ray particle, the rarer it is, and the less frequently it strikes the Earth.

It just wouldn't pay the nuclear physicist to wander about hoping that a superenergetic cosmic ray particle will strike his detecting device and do something extraordinary. If some superenergetics are occasionally detected by happenstance, fine, but for ordinary everyday work, it makes much more sense to deal with particles that are no more powerful than ordinary cosmic ray particles but that are produced by the trillions and can be made to strike at a known point in a known time and in a known way.

Which, of course, leaves cosmic rays still interesting in their own right.

For instance, exactly what are cosmic ray particles? I explained in the previous essay that in the early 1930s it had become quite clear that they were positively charged particles. The simplest positively charged parti-

50

cle known at that time was the proton. The proton was stable, so it could cross cosmic distances while retaining its identity, and it was quite massive, so that if it was going at nearly the speed of light, it would surely have the energy and the penetrating power characteristic of cosmic ray particles.

Why, then, look any further?

Well, though the proton was the simplest stable positively charged particle known to occur, it was not the only one. There are eighty-three elements known to possess stable isotopes, so that there are a couple of hundred stable isotopes altogether. Every one of those isotopes has a positively charged nucleus that is stable enough to survive travel across cosmic distances, and every one of them is more massive than the proton and is likely to be even more energetic as it speeds along. Some are over two hundred times as massive as the proton.

The various nuclei do not, however, occur in the Universe in equal amounts. Little by little, astronomers learned to determine, from light spectra, the ratio of elements in the Sun, in stars, in gaseous nebulas, and in galaxies. It became clear that by far the most common isotope in the Universe as a whole is hydrogen-1, the nucleus of which is a simple proton. The next most common is helium-4, the nucleus of which consists of two protons and two neutrons.

If we go by number of atoms, then about 90 percent of all the atoms in the Universe are hydrogen-1, 9 percent are helium-4, and everything else makes up the other 1 percent or so. Of course, the helium-4 nucleus is four times as massive as the hydrogen-1 nucleus, so that if we go by mass, roughly 75 percent of the mass of the Universe is hydrogen-1, 24 percent is helium-4, and, again, everything else makes up the remaining 1 percent.

51

It follows, then, that if cosmic ray particles are atomic nuclei, then the chances are that roughly 90 percent of them are hydrogen-1 nuclei (protons) and 9 percent are helium-4 nuclei, plus a thin scattering of a wide variety of more complicated nuclei. After all, there seems no reason to suppose that some rare isotope, say one of neodymium, ought, for some reason, be specially chosen to be fired out at great speeds while others aren't. Whatever does the firing should fire them all so that the various types of nuclei in cosmic rays ought to be, roughly at least, in proportion to natural occurrence.

But how can we be sure? It's all very well to reason and deduce, but

nothing beats actual observation. The fact is that from Earth's surface, it is difficult to observe cosmic ray particles directly. We observe chiefly the particles that result from the smashing of the cosmic ray particles into the atoms of the atmosphere. In the 1950s, however, we began shooting rockets above the atmosphere and a lot of them carried instruments designed to detect cosmic ray particles and to identify their nature.

It turned out that the reasoning was right. About 98 percent of the cosmic ray particles were atomic nuclei. The other 2 percent were high-speed electrons. There was also a trace of antielectrons and a smaller trace of antiprotons.

Of the 98 percent that are atomic nuclei, some 87 percent are hydrogen-1 nuclei, 12 percent are helium-4 nuclei, and 1 percent are all the other nuclei.

That certainly makes it look as though cosmic ray particles do indeed present a sampling of the Universe in general. But let's look a little more closely.

In the initial moments after the big bang, the temperature cooled to the point where the common subatomic

52

particles formed: protons, neutrons, and electrons. As the Universe continued to cool, protons and neutrons joined to form more complicated nuclei, and then electrons began to move into the neighborhood of the nuclei and formed intact atoms.

It would seem reasonable to suppose that the numbers formed of particular atoms would decline as the complexity of their nuclei grew. In a general way, this is true and the smaller nuclei are more common than the larger ones—but this is not an exact rule.

For instance, suppose you begin with protons, or hydrogen-1. A neutron combines with some of them to form hydrogen-2 nuclei. There are fewer of these than of hydrogen-1. Another neutron can add to a hydrogen-2 nucleus to form a hydrogen-3 (one proton and two neutrons), or a proton can add to a hydrogen-2 to form a helium-3 (two protons and one neutron). The hydrogen-3 is radioactive and spontaneously decays to helium-3, so we end with only helium-3. A neutron can then add to helium-3 to form helium-4 (two protons and two neutrons).

You would expect, then, that there would be a lot of hydrogen-1, less

hydrogen-2, still less helium-3, and just a trace of helium-4, but that's not the way it works. Hydrogen-2 and helium-3 sop up neutrons very readily so that, in effect, if you start with hydrogen-1, you slide right through the hydrogen-2 and helium-3 stages and end up with helium-4. As a result, you end with hydrogen-1 and helium-4 in a roughly ten-to-one ratio in atom numbers, while hydrogen-2 and helium-3 are present only in traces.

The helium-4 is so stable a nucleus and so reluctant to add on either a neutron or a proton that the nuclear buildup after the big bang stopped there. When the first stars formed they consisted onfy of hydrogen and helium. At the center of stars, however, conditions are

53

different than they are in space. At the center of stars, enormous pressures and densities combine with enormous temperatures to form nuclei far more complicated than that of helium. Some of these more complicated nuclei are eventually sprayed into space through supernova explosions so that later stars (like our Sun) are formed from materials that contain these complicated nuclei.

In the center of the stars, the concentration of various isotopes goes down as complexity goes up—but not perfectly. Elements 3, 4, and 5 (lithium, beryllium, and boron) form in quantity, but have a great tendency to indulge in further nuclear reactions and become elements 6, 7, and 8 (carbon, nitrogen, and oxygen). For that reason, there are many more nuclei of carbon, nitrogen, and oxygen in the Universe than of lithium, beryllium, and boron.

There is only one oxygen atom for every 1,500 hydrogen atoms in the Universe, but even so oxygen is the third-most-common element after hydrogen and helium. For every 660 million oxygen atoms there are 330 million carbon atoms and 90 million nitrogen atoms— but about 100 boron atoms, 11 beryllium atoms, and only 5 lithium atoms.

Among the cosmic ray particles, however, lithium, beryllium, and boron nuclei, while rare, are not so rare as they are in the Universe generally. These elements are anywhere from thirty thousand to 1 million times as common among the cosmic ray particles as they are in the Universe.

Why? The most likely reason is that as the cosmic ray particles travel across interstellar space, they occasionally collide with the sparse scattering of atoms and dust particles that are to be found there and, in so doing, they produce these rare light nuclei.

From the increase in concentration of these nuclei, it

54

is possible to make estimates as to just how dense the matter is in interstellar space. Apparently, the number of particles of matter that a cosmic ray particle would encounter, on the average, in its flight across space would be about 1 percent of the number of particles it would encounter in just passing through our atmosphere.

It is also possible to estimate how long the cosmic rays have been flying through space, and the best value seems to be, on the average, 20 million years. Since cosmic ray particles travel at nearly the speed of light, this means that the distance they have covered is nearly 20 million light-years.

If cosmic ray particles were traveling in a straight line, they would be originating at places that, on the average, would be eight or nine times as far away as the Andromeda galaxy. However, the cosmic ray particles are electrically charged so that their paths curve slightly in response to the electromagnetic fields of the various stars they pass and to the electromagnetic field of our Galaxy as a whole. They can therefore be seen as traveling around the Galaxy, just as the stars do. The particles make some two hundred circuits, on the average, before slamming into the Earth or some other similar object.

Energies on the subatomic level are measured in electron-volts. One electron-volt is the energy gained by a single electron accelerated through a potential difference of 1 volt. This is not a large amount of energy. Some 2,500 calories obtained from food would represent enough energy to keep a human being going for one day, and each of those calories is equal to 26 billion trillion electron-volts.

On the subatomic scale, though, energies in the electron-volt range are enough to hold electrons in an atom.

55

Chemical reactions, which involve transfers of electrons from atom to atom, give off or absorb energies in the range of several electron-volts.

Particles within the nucleus are much more massive than electrons are and are held together much more tightly. The energies involved are in the millions of electron-volts. Nuclear reactions are, for that reason, more energetic than chemical reactions are, so that when a nucleus breaks down, alpha particles can be shot out with an energy

of about 10 million electron-volts.

Cosmic ray particles are more energetic still. Even the lowest-energy cosmic ray particles have energies of nearly 1 billion electron-volts. About one thousand cosmic ray particles of such comparatively low energies strike every square meter of Earth's surface every second.

There are cosmic ray particles that are more energetic than this, but the more energetic the particles, the fewer of them there are. Apparently, each time you consider a tenfold increase in energy, the incidence of particles decreases by three hundred times. Thus, if you consider cosmic ray particles of 2 billion electron-volts, there will be only 3 of these striking every square meter of Earth's surface every second.

The most energetic particles yet detected have energies of 10 million trillion electron-volts, and these are so few that only three or four strike a given square kilometer of Earth's surface every year. When one of these hits an atom in Earth's atmosphere, it splatters it into a shower of a billion fragments that spray out over one hundred square kilometers of Earth's surface.

Can there be still more energetic particles in existence? Very likely, but the chance of detecting them is minuscule. A particle with an energy of 30 million trillion electron-volts, three times the maximum so far de-

56

tected, would strike at the rate of 1 per square kilometer per century.

The most energetic cosmic ray particles are really remarkable. The amount of energy concentrated in a single such particle is so great that if it could somehow be distributed as calories among the trillion trillion nuclei of our own body, it would keep us going for over half a minute.

The total energy of the cosmic ray particles traversing our Galaxy is surprisingly large. They represent an amount of energy equal to the light production of the Galaxy's stars. The question is how all that energy is concentrated into so comparatively small a mass.

• Every star emits constant streams of atomic nuclei | (mostly protons) in every direction. For our Sun, this is 1 the solar wind; for stars generally, it is the stellar wind. The speeding protons of the solar wind are ordinarily far less energetic than cosmic ray particles are. However, every once in a while there is an explosion on the Sun's

surface which produces a small and temporary solar flare that is far more energetic than the surface of the Sun is as a rule. The solar flare sends out a pulse of protons more energetic than the solar wind generally, and some of those protons attain the energies of weak, or "soft," cosmic ray particles.

We can assume that more massive, more turbulent, and more unstable stars than the Sun produce stellar winds that are much more rotense and energetic than that of the Sun, and that such active stars produce greater quantities of more energetic, or "harder," cosmic ray particles. Truly energetic events, such as super-{ novas, or the active centers of galaxies may produce still " harder cosmic ray particles.

However, even the most energetic supernova cannot

57

account for the upper ranges of cosmic ray particle energies.

That fact is not in itself troublesome. The situation is analogous to that of a rocket ship. To lift off Earth, a rocket ship needs to reach the escape velocity of seven miles a second, if we plan to do it in one big blast to begin with and just coast upward on the force of that blast.

However, it can also be done in stages. We can give the spaceship a push that will not, in itself, send it away from Earth, but will lift it above the thickest part of the atmosphere. There, where air resistance has become minor, the portion containing the fuel for that first push falls off, and a second push gives the ship (now with a much smaller mass) another jolt. Later, there can be a third jolt.

By doing it in stages, less fuel is required to send the ship into orbit, or on its way to the Moon. Once it is in space, if it is then powered by a fuel sufficiently high in specific impulse (an ion-drive, for instance), it can continue to accelerate until a speed is achieved that is a respectable fraction of the speed of light—a speed that could never have been attained in a single push under any reasonable conditions.

Very well, then, we needn't imagine cosmic ray particles coming out of some source and attaining, to begin with, several million trillion electron-volts. Suppose they come out merely with several billion electron-volts of fairly soft cosmic rays and are then accelerated.

But how do they accelerate? They have no rocket fuel.

They are electrically charged, however, and if they pass through a magnetic field, that will accelerate them just as magnetic fields

accelerate particles inside a human-made synchrocyclotron.

Certainly there are magnetic fields in space. The

58

Earth has a magnetic field, Jupiter has a much stronger one, and the Sun has one that is stronger still. Some stars have a stronger field than anything in our Solar system. The Galaxy itself has a general magnetic field of its own. (All this is true of every other galaxy as well, we can be sure.)

We therefore picture cosmic ray particles being produced by energetic stars and supernovas, streaking off through a largely empty space in paths that gently curve in response to the Galactic magnetic field. As they curve they accelerate and gain energy at the expense of the Galaxy generally. As they gain energy, their path straightens out, curving less under the influence of the magnetic field.

Every once in a while, a cosmic ray particle will pass near a star and, thanks to the star's magnetic field (which may sometimes be unusually strong), it will accelerate more sharply, bending in its path due to gravitational influence and then moving along as a considerably more energetic particle following a considerably less curved path.

Each particle accelerates in a more or less uneven fashion, but, as they all gain energy, they tend to move about the Galactic center in a generally expanding spiral until they happen to hit an object large enough and massive enough to absorb them. Those that happen not to hit anything for a sufficient number of millions of years gain so much energy that they scarcely curve at all in response to magnetic or gravitational fields. They move in paths that are sufficiently close to straight lines to carry them outside the Galaxy altogether and to shoot off through intergalactic space.

The really energetic cosmic ray particles that strike Earth travel in so nearly a straight line that they must have come from other galaxies in all likelihood, just as some of ours eventually reach galaxies other than our

59

own. Sooner or later, the energy withdrawn from galaxies and used for acceleration is returned to the galaxies through collision and absorption.

From the fact that cosmic ray particles are virtually entirely normal nuclei in nature, with only the barest trace of antinuclei, we can conclude, from that alone and with a fair degree of certainty, that our Galaxy is entirely matter. From the fact that even the most energetic cosmic ray particles seem to be positively charged and not negatively charged, we can suspect that virtually the entire Universe is matter.

This requires explanation since, under ordinary circumstances, particles of matter can't be formed without the accompanying formation of equal particles of antimatter. (I took up that matter in "The Crucial Asymmetry," F & SF, November 1981.)*

The question is, though, whether we can really expect cosmic ray particles to be accelerated by magnetic fields to a sufficient extent to attain the energy levels they do, in fact, attain. The Galactic field itself is not very intense, and the chance of approaching stars closely enough to get a more intense acceleration is not very great. In fact, it does not appear that cosmic ray particles produced and accelerated in ways that astronomers in the 1960s knew about could possibly attain the upper, reaches of energy that had been observed.

More powerful sources had to be found, or more powerful accelerations, or both.

A possible solution was reached in 1969, when pulsars were discovered and found to be reasonably common. These are condensed stars with the mass of an ordinary star but a diameter of some fifteen kilometers. They rotate in anywhere from several seconds down to several thousandths of a second. Their magnetic fields are

* See my book Counting the Eons (Doubleday, 1983).

60

as condensed as their gravitational fields, and are enormously intense.

A charged particle emerging from a pulsar's powerfully energetic surface is an energetic particle to begin with, and the incredibly intense magnetic field would accelerate it to enormous levels of cosmic ray particle energy at once.

However, if astronomers calculate the rate at which pulsars lose energy and slow their rotations (largely because they are radiating gravitational waves), things don't look so good. Even if all their loss of energy is put into the acceleration of cosmic ray particles, it wouldn't be sufficient to account for the upper range of particle energies.

But there are certain X-ray sources consisting of binary stars. One of the pair is a condensed star, either a pulsar or a black hole; the other is a normal star with from ten to thirty times the Sun's mass.

If the normal star is expanding toward the red giant stage, it tends to have some of its mass drawn into the intense gravitational field of the condensed star. This leaking mass spirals down into the condensed star, radiating X rays intensely as it does so.

The energy generated by this spiraling mass can be 100,000 times as intense as that delivered by our Sun.

At first, it was assumed that all this energy was radiating away as energetic photons (of X rays, for instance). Some is, of course, but, beginning in 1972, evidence began to accumulate that a good deal of the energy appeared in the form of cosmic ray particles. It is these particles, energetic enough to begin with, and enormously accelerated, that may be the source of the upper energy reaches of cosmic ray particles.

Perhaps!

61

Part II Humanity

The Road to Humanity

Last night (as I write this) there was an hour-long program on WABC-TV, and during the last ten minutes or so, various well-known personalities were asked to comment briefly on the subject matter of the program. Among the personalities was I.

As I watched, I could not help but ask myself, "Why am I included with all these people?" . . . You see, for all my reputation as a man of colossal ego, I have never gotten used to my present position as "celebrity." That position didn't come overnight, after all, or as a result of any single remarkable event. In fact, for most of my life, there seemed no sign that I would ever come to anything.

I made my first professional story sale in 1938, when I was eighteen, and by the time I got married, in 1942,1 had accumulated a bank account of $400 from my writ-

65

ing. My parents were in no position to help me out; I had no other relatives; and I certainly had no fairy godmother. My bride added $300 to the kitty, so I began married life with $700 in cash, and with a job that was going to pay me $2,600 a year, but would cease to exist when World War II ended or when I was drafted, whichever came first. My prospects were not bright.

By 1958, I was a little better off. I had a position as Associate Professor of Biochemistry at Boston University School of Medicine and my annual salary had reached the dizzying height of $6,500 a year. In addition, I was earning $15,000 a year through my writing.

However, I now had two children to support in addition to a wife, and I was so little a celebrity that the Director of the Medical School was annoyed with my neglect of research in favor of writing and fired me. By dint of hard fighting, I held on to my title, but my salary disappeared forever on June 30, 1958.

So there I was, thirty-eight and a half years old, clearly middle-aged, with a family to support, with 30 percent of my income suddenly gone, and with absolutely no status or reputation except with a few loyal science fiction readers. My prospects were still not bright.

Yet I made it. I'm not sure how it happened, or exactly when. One of the reasons I undertook to write my autobiography ten years ago, and then wrote it in great detail and in exact chronological order, going over my diary painstakingly from page to page, was in order that I might catch the moment when I suddenly emerged from my chrysalis.

It didn't help. I never found that moment. It had all happened so slowly, so gradually, so unnoticeably, that I was never aware of any change. By the time I came to the realization (with some disbelief) that I was a celeb-

66

rity, it turned out that everyone else had considered me to be one for some years.

I suspect that's a common state of affairs and can be applied to matters of much more moment than the life of individual human beings. For instance, when and how did humanity come into existence? What was the key event? To answer that question, let's start at the beginning and progress along the road to humanity in twenty evolutionary steps.

1. 4,600,000,000 years Before the Present (B.P.). The Solar system,

including the Sun and the Earth, have, at this time, formed out of a primordial cloud of dust and gas.

2. 3,600,000,000 B.P. The first indications of life appear in the form of prokaiyotic cells, tiny cells such as those of bacteria and cyanobacteria. (The bacteria are without chlorophyll, while the cyanobacteria possess it.) Such prokaryotic cells exist today and are not very different from the cells that first formed so long ago.

3.1,400,000,000 B.P. After more than 2 billion years in which prokaryotes remained the only form of life on Earth, eukaiyotic cells formed. These were single-celled organisms like the prokaryotes, but the eukaryotes are substantially larger and possess nuclei in which are concentrated the reproductive and hereditary functions of the cell. The eukaryotes may have formed through the combination of different prokaryotic cells that then, within an overall cell membrane, lived in symbiotic relationship with each other. Single-celled eukaryotes still live today—amoebas, paramecia, algae, and so on.

4. 800,000,000 B.P. At about this time, some eukaryotes went through the process of joining together to form multicellular organisms. All multicellular organisms (including human beings) are made up of eukary-

67

otic cells. The multicellular organisms evolved and diversified into numerous grand divisions called phyla (singular, phylum, from a Greek word for "tribe"), both plant and animal.

5. 550,000,000 B.P. Now the first animals belonging to the phylum Chordata appear. This is the last phylum to make its appearance, apparently, and it is to this phylum that human beings belong.

The first chordates were primitive* creatures that did not seem very different from worms. They apparently arose from another phylum called Echinodermata (Greek for "spiny-skins"), of which the best-known representatives today are the various starfish. In fact, the most primitive chordate living today, the balanoglossus, in its larval (that is, immature) state is so like echino derm larvae that it was first classified as an echino-derm.f

The chordates differ from all other phyla in three ways. First, they possess a notochord during at least some stage in their development. This is a stiffening rod that runs down the back, presaging the development of an efficient internal skeleton.

Second, they possess a hollow nerve cord down the back, just under the notochord. All other phyla have a solid nerve cord running down the abdomen. The chordate nerve cord eventually developed into a complex nervous system, superior to that of any other phylum.

Third, they possess gills, richly supplied with blood vessels, along which water passes and from which food can be strained and oxygen absorbed.

* Primitive and advanced are subjective words and, to the layman, represent the degree to which organisms resemble human beings in one respect or another. The greater the resemblance, the more "advanced" they are.

t A reader has written to tell me that the balanoglossus is no longer considered a chordate, because it does not really possess a notochord.

68

6. 510,000,000 B.P. From the primitive chordates, there now developed other chordates with additional characteristics. In place of the notochord, for instance, a line of vertebrae enclosed the nerve cord. They were separated so that the body could twist and the head could, eventually, turn (vertebrae is from a Latin word for "turn"). The first vertebrae were composed of cartilage, tough and flexible.

Chordates possessing vertebrae belong to the subphy-lum Vertebrata, and these now include all chordates (including human beings) except for some very primitive and out-of-the-way specimens like balanoglossus.

The earliest vertebrates of note were ostracoderms, (Greek for "tile-skins"), fishlike creatures without jaws. They were most notable for being the first to develop bone, which is to be found only in them and most of their descendants (including ourselves). The bone was most notably present as an outer casing that enclosed the head, which contained, after all, the chief sense organs and the nerve cord swelling we call the brain. The present-day organism most closely related to the ostracoderms is the lamprey, a jawless, eel-like animal.

7. 440,000,000 B.P. From the ostracoderms, there evolved the Acanthodii (Greek for "spiny," since they possessed spines at their fins). They were the first vertebrates with jaws—developed out of the first gill arch (the cartilaginous stiffening bars at each gill opening). From these organisms seem to have been developed the placoderms (Greek for "plate-skins"), which had not only jaws, but also two sets of paired fins, for steering. These represented the beginning of the four

limbs that all later vertebrates had, except in those cases where the two fore limbs (kiwis), or the two hind limbs (whales), or all four limbs (snakes) were reduced to vestigial remnants.

The placoderms had plates of bony armor over the

69

head and forepart of the body (hence, their name) and were the largest and most formidable creatures of their time when they were at their peak.

8. 400,000,000 B.P. From the placoderms, there evolved the Chondrichthyes (Greek for "cartilage-fish"). In them, the bone was lost and the internal skeleton was composed of cartilage. The chondrichthians thus lightened their bodies without sacrificing security too much. What they lost in invulnerability, they more than made up for in the gain of mobility. The most familiar chondrichthians that survive to this day are the

sharks.

At about the same time, from the acanthodians, there evolved the Osteichthyes (Greek for "bony fish"), which retained the bone but kept it inside the body where it made up the internal skeleton. The osteichthians and all their descendants (including human beings) retained the bony internal skeleton.

Not long after their appearance, the osteichthians divided into two branches. One was the Actinopterygii (Greek for "ray-fins"). Their fins were thin, with stiffening rays of cartilage, and were admirably adapted for swimming and steering.

The other branch was the Sarcopterygii (Greek for "flesh-fins"), who had two pairs of stubby, fleshy limbs with only a fringe of fin. Such fins were less good at swimming, but when a pool of water became brackish, muddy, or threatened to become dry, a sarcopterygian could stump across a stretch of dry land to another pool. The pattern of bones in the stubby sarcopterygian fins were retained in all their descendants (including human beings).

The chondrichthians and the actinopterygians flourished and have continued to flourish, as sharks and fish, to this day, but they proved dead ends. No startling new developments were derived from them. The sarcopter-

70

ygians, on the other hand, dwindled and all but died out. Only a few remnants are left, like the coelacanths (Greek for "hollow-spines"), which were discovered still living in the ocean in 1938. Yet it was to the descendants of the sarcopterygians that the future belonged and from whom human beings were to descend.

9. 350,000,000 B.P. About this time, some sarcopterygians had evolved into organisms that could, in adult life at least, remain out of water for extended periods. Their stubby fins had become legs, and they had simple lungs that made it possible for them to gulp air and obtain oxygen in that way, rather than depending solely on oxygen that was dissolved in rivers, lakes, or the sea. They were the first vertebrates with legs, and the legs were retained in almost all their descendants (including human beings).

These organisms had to lay their eggs in water and from those eggs there hatched larvae that were much like fish, lacking legs and possessing gills. For that reason, these organisms were placed in the class Amphibia (Greek for "double life") within the vertebrate subphy-lum. Familiar amphibians alive today are the frogs and toads.

The amphibia were by no means the first organisms to invade the land. Plants had colonized the land some 50 million years before the amphibia arrived. Following the plants, in comparatively short order, were such organisms as snails, spiders, and insects.

Amphibia, however, were the first land-living vertebrates and they were the largest animals of any kind that had yet appeared on land. Some forms, now extinct, were armored and were as large as modern crocodiles. The weakness of the amphibia, however, was that most were tied to water in early life, and this limited their control of the land.

10. 300,000,000 B.P. About this time, certain amphibia

71

developed an egg that was surrounded by a protective shell of thin limestone. The shell was permeable to air, but not to water. Air could reach the developing embryo inside, but water could not leave it. The embryo developed in a small reservoir of water inside the egg, with an elaborate series of adaptations allowing the embryo to tuck wastes into special membranes. With such eggs, organisms could remain on land indefinitely and were freed of the necessity of water life.

The organisms with such a land-based egg belong to the class Reptilia

(Greek for "creeping," since the most familiar reptiles in existence today are the snakes). The reptiles were able to colonize the land generally and became the dominant form of land life on Earth, at least in the sense that they were the largest. In fact, one reptile, now extinct, the brachiosaur, holds the all-time record as the most massive land animal that ever lived.

11. 270,000,000 B.P. The reptiles quickly diverged into a number of varieties, and at this time there developed the Theriodontia (Greek for "beast-toothed"). Their teeth were more differentiated than were those of other reptiles (more like ours than like those of crocodiles, in other words). Some among them may also have developed the capacity of maintaining a constant internal temperature (above that of the environment, usually) rather than taking on whatever the outside temperature might be. The theriodonts may thus have developed warm-bloodedness, where all other organisms that existed till then seem to have been cold-blooded. To cut down the loss of heat, some theriodonts may even have developed hair, a modification of the reptilian scale. (Later, birds evolved from other reptiles. They were also warm-blooded, and developed feathers, another modification of the reptilian scale, to conserve heat.)

Warm-bloodedness, a property which all the descendants of the theriodonts (including human beings) re-

72

tain, has the advantage of enabling an organism to remain active at all times, neither becoming torpid in the cold nor suffering sunstroke in the heat. The price to be paid, however, is that warm-blooded organisms must eat much more than cold-blooded organisms of the same size if they are to find the energy to maintain body heat.

12. 220,000,000 B.P. The theriodonts did not flourish and eventually died out, but before doing so they gave rise, at this time, to varieties that developed teeth, jawbones, inner ear structures, and other characteristics that resembled those of organisms like ourselves more and more closely. These were members of a new class, Mammalia (Greek for "breasts," since modern organisms of the class have breasts that produce milk for the feeding of the young). Human beings are obviously mammals.

The earliest mammals were small shrewlike organisms that managed to exist only with difficulty in a world dominated by reptiles, and survived only because they were small and could hide. They may well have laid eggs and may have had only primitive breasts, if any. Three species of egg-laying mammals still exist in Australia and New

Guinea. The duckbill platypus is most familiar.

13. 100,000,000 B.P. The primitive mammals gained a new advantage by developing reproductive mechanisms that offered increased protection to the young. Certain mammals at this time evolved the ability of allowing their eggs to hatch while still in the body. When the young finally emerged (still very undeveloped), they could make their way into a pouch within which they could attach themselves to nipples and feed on milk till they were much better developed. Such mammals are called marsupials (Latin for "pouch").

Other mammals evolving about this time went even further. Not only were the eggs hatched within the body, but they could remain within the body, nourished by a

73

placenta (Greek for "flat cake," because of its shape). Food could diffuse from the mother's bloodstream into the embryo's bloodstream across the placenta, while wastes diffused in the opposite direction. The embryo could be developed within the body until it was in a comparatively advanced state. Such mammals are placentals, and human beings are among them.

Among the placentals were an order of organisms known as Insectivora (Latin for "insect-eating"). The best-known modern insectivores are the shrews and hedgehogs. They are primitive organisms, with unspe-cialized limbs that retain the five digits to each paw that were to be found in the first amphibians.

Some of these insectivores had a rather large brain for their size, and the first digit of the paw could separate somewhat from the rest, so that it seemed to represent the beginnings of a thumb. The best modern example of such an insectivore is a tree shrew that lives in southeastern Asia and rather resembles a squirrel in appearance.

14. 70,000,000 B.P. At this time, certain insectivores had developed characteristics that placed them into a new order, Primates (Latin for "first," a bit of egotism, since the order includes human beings). The first primates may not have been very different fronj tree shrews. Indeed, there has been a tendency to consider modern tree shrews the most primitive primates rather than the most advanced insectivores.

15. 65,000,000 B.P. When the primates first appeared, the reptiles still dominated the land, and the mammals after 150 million years still had a most precarious existence. At this time, however, something happened. It may have been the collision of a small comet

with the Earth, or it may have been something else, but most of the large reptiles (along with many other kinds of organisms) died out rather suddenly.

74

Some of the mammals managed to survive the catastrophe, whatever it was (as did some reptiles, for that matter), and since they were now not competing with overwhelming numbers of large reptiles, they had the chance to evolve into a wide variety of spectacular organisms themselves. There were marsupials as large as modern hippopotamuses, and placentals four times as large as modern elephants. On the whole, placentals proved more formidable than the marsupials, and the latter have survived to this day, for the most part, only in Australia where, until the coming of human beings, placentals have not been well represented. Placentals dominated the rest of the world.

The giant mammals did not survive, however. The most successful mammals were those that were relatively small and agile. Mammals also evolved brains that were larger than those of other types of organisms of the same size and this seems to have helped in survival.

Included in this drive toward brains were the primates, which eventually did remarkably well in this respect. In the modern world, the only nonprimates to have brains that are larger than any primate brain are to be found in the order of Cetacea (Greek for "whale"), which includes the whales and dolphins, and in the order of Proboscidea (Greek for "to browse in front"), which includes the elephants, whose trunks make it possible for them to reach far forward for vegetation.

These larger brains, however, must handle a much larger body. Where the largest primate brain has a mass that is 2 percent of the body it is in, the largest elephant brain is only 0.1 percent the mass of its body, and the largest whale brain only 0.01 percent. It is apparently the size of the brain compared to the body that counts, and not the size of the brain alone.

Within any group of similar animals, the brain/body ratio tends to grow larger as the size grows smaller. A

75

dolphin, weighing no more than a man, has a brain that is as much as 2.5 percent the mass of its body. However, dolphins, living in water, must have a streamlined body so that they lack irregular projections

such as arms and hands. Furthermore, in the sea, it is impossible to deal with fire and that deprives the cetaceans of any chance of developing a technology. Their large brain therefore does them no good by primate standards.

Of course, small primates may have a higher brain/ body ratio than larger ones do. Some small primates have a brain that is more than 5 percent the mass of the body, as is true of some hummingbirds as well. The total weight of the small-primate brain, however, is too small for the kind of overall complexity required for high intelligence and the hummingbird brain is more minute still.

For a combination of large brain, relatively small body, and land life, nothing can surpass the largest primate brain—which is, as you may have guessed, our own.

The early primates are represented today by the lemurs (Latin for "ghosts," because they are mostly nocturnal, appearing dimly by night). They are found today, for the most part, in the island of Madagascar.»

16. 55,000,000 B.P. At about this time, a line of early primates evolved into the Tarsiiformes, The only living species of these animals is the tarsius (so called because of the unusually long bones in its ankles, or tarsus).

These organisms had their two eyes both front and close together, rather than one on either side of the head as in other early primates. That made possible the use of stereoscopic vision and increased the detail they could see. The additional information received put a premium on brain size and the tarsiiforms had larger brains than the other early primates. All the descen-

76

dants of the tarsiifonns (including human beings) have these forward-looking eyes.

17. 40,000,000 B.P. At about this time, primates belonging to the suborder Anthropoidea (Greek for "manlike") appeared. They did so, it is thought, from a branch of the tarsiiformes that are called Omomyidae, all of which are now extinct, though their descendants survive. These descendants are monkeys, apes, and human beings.

The anthropoidians can all sit up easily so they can use their forepaws for handling and manipulating with greater ease. Their fingers and toes have nails rather than claws, so that the softer, more sensitive

parts of the digits can be exposed for handling and manipulating.

The omomyids were found in both the Americas and in Eurasia and anthropoid species developed in both places. These are popularly differentiated as New World monkeys and Old World monkeys.

The New World monkeys have nostrils well separated and facing outward so they are called Platyrrhini (Greek for "flat noses"). They are relatively small, the largest weighing about twenty-two pounds, and have long tails. Some of the tails are prehensile and can be used as a fifth grasping device.

The Old World monkeys, which tend to be larger than the New World monkeys, have well-defined noses, with the nostrils close together and facing downward so that they are called Catarrhini (Greek for "downward noses"). Clearly, we are descended from the Old World monkeys.

Many Old World monkeys have tails, but those tails are never prehensile. As though to make up for the lack of prehensile tails, the Old World monkeys have hands and feet that are more efficient at grasping than are those of New World monkeys. The Old World monkeys have better thumbs and stronger grips. Since this in-

77

5 billton e f

4 blHion a p

3 billion B f>

2 billion B P

1 billion e P

900 million B P •

800 million B P

700 million B p

600 million B P

500 million B P

400 million B P

300 million e P

200 million a P 100 million a P 90 million B p 80 million B P 70 million B p 60 million e p 50 million B P 40 million B P 30 miHion B P 20 million B p 10 million BP

• Earth formed

• First prokaryotes

• First eukaryotes

•First multicellular organisms

- First chordates

- First vertebrates

- First jawed vertebrates

- First bony fish

- First amphibia

-First reptiles First thenodonts

- First mammals

- First ptacerrtals

• First pnmates

• End of giant reptiles

• First tarsiltormes

- First monKeys

•First apes

First great apes

• First homlnlds

-PRESENT

BP   BEFORE THE PRESENT Not* Chart Is not comptoWjf to scale

creased the flow of information, it further encouraged an increased size of brain.

18. 30,000,000 B.P. At about this time, the Old World monkeys developed a branch classified as the superfam-ily Hominoidea (Latin for "manlike"). This superfamily includes both apes and men.

Ape was originally used for a tailless monkey, like the Barbary ape which is found in North Africa and on Gibraltar. The organisms we now call apes are, for the most part, larger than the Barbary apes, and, indeed, include the largest primates who have ever lived. They also more closely resemble human beings than any primates outside the superfamily do, so they are sometimes called anthropoid apes to distinguish them from tailless monkeys.

19. 17,000,000 B.P. The early apes were small, rather like the gibbons of today, which are the least advanced of the apes. At about this time, however, the subfamily Ponginae (Congolese for "apes") evolved. They are commonly called the great apes.

The largest living great ape is the gorilla, which is over five feet tall and may weigh five hundred pounds or more. Still larger is a now-extinct gorilla-like ape, Gi-gantopithecus (Greek for "giant ape"), which was nine feet tall and may well have tipped the scale at a thousand pounds or more.

The great apes are the most intelligent of the nonhu-man primates and have the largest brains. Leaving the human brain to one side, the largest primate brain is that of the gorilla, which weighs up to nineteen ounces. The chimpanzee, which is a smaller ape, has a brain of about thirteen and a half ounces, while the brain of the orangutan is about twelve ounces.

20. 5,000,000 B.P. Any pongid that resembles the modern human being more closely than it does any of the apes, living or extinct, is called a hominid (from a Latin

79

word for "man") and it is at about this time that the first hominid appeared.

The chimpanzee is the closest of ail animals to the human in the genetic sense. Human genes and chimpanzee genes are so similar that the amazement is that humans and chimpanzees are as different as

they are. Very possibly, then, a common ancestor split into the two divergent lines about this time—a pongid from which the chimpanzee descended, and a hominid from which human beings descended.

The first hominids were comparatively small, perhaps four feet tall, no larger than the chimpanzees from whom they diverged, and probably more lightly built. The hominid brain may not have been more than fiftefcn ounces at first, scarcely more than that of a chimpanzee. However, the brain/body ratio hi the early hominid was perhaps twice that of a modern chimpanzee, and four times that of a modern gorilla.

Even the first hominids, then, may have been, at least marginally, the most intelligent land animals that had yet existed. Yet this is not the crucial point that made the hominid different, for it was only a small matter of degree. There was another difference that was much more important. v

The first hominid could walk upright exactly as we ourselves do. This is something no other primate could do, and no nonprimate either, in quite the same way.

I'll discuss the consequences of this in the next essay.

80

Standing Tall

When my beautiful, blue-eyed, blond-haired daughter, Robyn, was a little past her first birthday, it seemed to me that it was quite time she should be able to walk upright. Therefore, when I caught her propelling herself forward on her little legs, while hanging on to various articles of furniture, I very carefully and gently detached her arms from said articles in order to see what would happen.

What happened was that she promptly sat down with a plop.

I was chagrined and felt (as I do about all problems) that it only required a reasoned discussion of the matter.

"Walk, dear," I said to her. "Move your legs and don't hang on. Do like Daddy does. Here, watch Daddy. See? Like this."

81

It did no good. There was no reasoning with her at her age. (Nor pretty much at any age, I eventually found out.)

Then one day, shortly afterward, when I was sitting in the kitchen in the expectation of being fed lunch, Robyn walked in, and since I am not (and never have been) a noticing person, I simply said, "Hello, Robyn."

It was only after several seconds that the truth of the situation forced itself upon me and I said, in astonishment, "You're walking."

And so she was. Little Robyn had, in some dim way, discovered that it was easier to walk than to crawl and promptly began to walk. She never crawled again— which brings me to the point of this essay.

Human beings have always looked for some clear distinction between themselves and all other animals (out of self-importance and self-love, I presume).

Theologians found the perfect solution. Human beings are made in the image of God, while other animals are not. This brings on the difficulty that it limits Cod to imagine him as having any corporeal shape at all, let alone that of a man, so that statement is modified to "in the spiritual image of God." In other words, man has a soul and other animals do not. This is an irrefutable statement—and also an unprovable one. Therefore, those of us who find it difficult to rely on faith alone, but who want a difference to exist anyway, must find a physical and demonstrable one.

For instance, other animals have tails, but we don't. Other animals have body hair, but we don't. Other animals can't talk, but we can. Other animals have little brains, but we have big ones.

Somehow, though, it's not as simple as it seems. Bears, guinea pigs, and gorillas have no tails. Elephants,

82

hippopotamuses, and dolphins have no body hair. Animals may not speak English but they communicate. Elephants and whales have bigger brains than we do.

Of all the separating characteristics, however, bipedality—the ability to walk on two legs—seems the most attractive. The Greek philosopher Plato thought it was, but he had to eliminate birds, which were all bipedal. He therefore defined the human being as a "featherless biped."

Whereupon his fellow philosopher Diogenes plucked a chicken and

held it up, saying, "Here is Plato's man!"

This is a nonsensical counterargument, however. Just because a particular chicken has its feathers removed doesn't mean that the abstract concept "bird" doesn't have feathers. Diogenes might have brought in a kangaroo, or a jerboa, or a Tyrannosaurus rex, if these had been available, and that would have been a genuine refutation of the definition.

Still, Plato's feeling about bipedality was, in my opinion, correct. Let's think about it.

Vertebrates that are bipedal usually are restricted to two legs because the two others have been devoted to some other (and preferred) form of locomotion that does not involve legs primarily. Most birds are designed to be flyers and walking, running, or hopping is strictly secondary. The penguin is designed to be a swimmer, and walking is secondary.

But what about nonflying birds like ostriches, where walking or running is the only means of locomotion— and a good one since they can run as fast as horses when pressed? In such cases, the body is designed for it. The body is essentially horizontal, with as much sticking behind the legs as in front so that the center of gravity is above the legs.

This is also true of bipedal reptiles and mammals.

83

Think of the tyrannosaurus and the kangaroo. Each has a long tail for balance.

Suppose, though, there is no tail to act as a balance. In that case, the only way the body's center of gravity can be brought above the two hind legs is to tip the entire body into a vertical position.

Some tailless animals actually do this. Bears and chimpanzees can stand upright on their hind legs and can even walk about in this fashion, but they are clearly uncomfortable in doing so. They (like baby Robyn) feel much better if they allow their forelimbs to share the work. And, unlike Robyn, they never get to the stage where it becomes comfortable to use their hind legs only.

Plato would have done better, therefore, to define the human being as a "tailless, habitual biped." In that case, Diogenes would have found it much more difficult to find a counterexample. (He might have cut the

wings off a penguin—but penguins, even though they walk upright like human beings, are obviously clumsy at it and even without wings would prefer to belly whop on the ice if they can.)

What makes it possible for human beings to walk comfortably on two legs?

It is that the spinal column, just above the pelvis, bends backward in human beings. It assumes a shallow S-shape in us, and can therefore remain generally vertical without trouble. It adds a little spring and bounce to the human walk. No other organism has that backward bend to the spine in the small of the back, so that while some tailless animals can walk bipedally at need, none do so comfortably, let alone preferably.

How did the human spine develop that backward bend?

Presumably there is some advantage to getting on your hind legs. It lifts your head and major sense organs

84

higher so that you can spot food, or danger, at a greater distance. It also frees your forelimbs for temporary duty for something other than support, so that you can hold food, say, or a baby.

Various apelike creatures, some millions of years ago, would raise themselves to their hind legs temporarily for the advantages that would bring them, and those who could do so with reasonable comfort were, in the long run, better off. A particular species of ape experienced a random mutation that happened to make the spine a bit more bendable in the right place and that improved its chances of survival. Any further change in that direction would then be selected for and, eventually, you would have a tailless organism that could walk on two legs easily and comfortably.

Any such ape would then be closer to us in a key anatomical respect than it would be to any other ape, living or extinct. Such an organism would then be a hominid, or direct predecessor of man, rather than a pongid, or ape.

Now we can return to the point where we left off in the last essay, which, as you'll recall, was the twentieth step on the road to humanity.

20. 5,000,000 B.P. The earliest hominids were first identified by an Australian-South African anthropologist, Raymond Arthur Dart (b.

1893), to whom a skull, human-looking except for its extraordinarily small size, was brought from a South African limestone quarry in 1924. He recognized it as belonging to a primitive ancestor of humanity and, in 1925, suggested it be called Australopithecus, from classical words meaning "southern ape."

This is actually a bad name, for three reasons. First, it is a mixture of Lathi and Greek. Australo comes from the Latin ouster, meaning "south," and pithecus comes

85

from the Greek pithekos, meaning "ape." These Latin-Greek mixtures are frowned on by purists.

Then, the use of the Latin auster instead of the Greek notos somehow gives the impression that these early hominids lived in Australia (which is also named for "south," for obvious reasons) and that isn't so.

Finally, this primitive creature was not an ape but a hominid, and should have been called Notoanthropus, or something like that. However, it is hard to tell from a skull alone that an organism walks erect and was therefore a hominid. That knowledge came only after fragments of thighbones and pelvises were uncovered.

Since 1924, other remains have been found of such hominids and it is now believed that they existed in perhaps four different species, lumped together as aus-tralopithecines.

The best remains of the earliest of these species was found in 1974, when a large fraction of the skeleton of an australopithecine was located in east-central Africa by an American anthropologist, Donald Johanson. It seemed to be the skeleton of a woman so it was nicknamed Lucy. It was at least three million years old, and possibly four. We might speculate that the very first aus-tralopithecines, scarcely to be differentiated from the ancestral pongid, may have lived five million years ago.

Lucy is an example of Australopithecus afarensis, so named because Afars is the name of the territory where the remains were found. Apparently, east-central Africa was the cradle of humanity.

A. afarensis must have looked very much like a chimpanzee. The adults were no taller than a chimpanzee, and slighter in build. They seem to have ranged between three and four feet in height and weighed perhaps sixty-five pounds.

The brain, too, was no larger than a chimpanzee's, about 380 grams, or a quarter the size of our own. A.

86

afarensis probably lived much as chimpanzees do and, from its hipbones, toes, and fingers, there is a feeling it spent much of its time in trees. It certainly couldn't speak, and it must have been largely a vegetarian, though it may have scavenged meat from animals that had been killed and left over by the true carnivores.

However, since A, afarensis had a chimpanzee brain in a body half the weight of a chimpanzee, its brain/body ratio was twice that of a chimpanzee. The first hominid may already have been more intelligent than any ape. Even more important, A. afarensis walked on its hind legs as easily as we do, I have seen it suggested that this made it possible for it to scavenge. The females were not forced to remain near their helpless young, but could carry those young in one arm and run after the carnivores, ready to eat whatever they left over.

This would have been all the more important if A. afarensis was already beginning to lose body hair so that the young could not hold on to the hair, allowing the organism to run freely on all fours. (We don't know at what stage in hominid evolution body hair was lost.)

Also, it is possible that the young were routinely held in the left arm to have them nearer the beating of the heart. This would more closely resemble the environment in the uterus and experience might have showed that the child would remain more quiet there. It may be for this reason that hominids used their free right hand for other purposes and developed the overwhelming right-handedness that characterizes human beings (but not other animals) today.

21. 3,000,000 B.P. By 3 million years ago, A afarensis was definitely on the way out. It must surely have become extinct by 2.5 million years ago. Even so, it cannot be reckoned a failure. It is possible that the species survived for 2.5 million years and it seems extremely doubtful that our own species will do as well.

87

However, A. afarensis didn't disappear altogether. They had left descendants who, through the slow process of evolution, had become sufficiently different to be considered a new species. By 3 million years ago, Australopithecus africanus existed.

A. africanus was very much like A. afarensis. It may have been no taller, but it was a little more heavily built and it may have weighed as much as ninety pounds in some cases. The brain, too had increased in size and was now almost the size of that of a modern gorilla, say five hundred grams, or about a third the weight of our brain.

Most of the remains of A. africanus have been found in southeastern Africa. The first findings by Dart had been of A. africanus and they deserved the species name at the time, for they were the first hominid remains to be found in Africa. (It also took some of the curse off the prefix Australo-,)

22. 2,500,000 B.P. The evolution of the aus-tralopithecines is very hard to follow. No intact skeletons have been found, only scattered remains, and we can never tell whether a particular scrap happens to be typical of the particular australopithecines of that time and place, or if it happens to be of an individual that is atypical for some reason.

It had been thought, for instance, that A, africanus had given rise to a third australopithecine species which had in turn given rise to a fourth. But then, in the summer of 1986, a 2.5 million-year-old skull was found with a prominent ridge on top, to which, undoubtedly, powerful jaw muscles had once been attached.

Anthropologists are not certain what this find means. It is called the Black Skull from its color. It seems to be an australopithecine beyond doubt, and the best bet right now is that it descended from A. africanus and, in its turn, split up into the remaining two australopithe-

88

cine species simultaneously not long after 2.5 million years ago.

One of these species is Australopithecus robustus, which is so-called because it is larger than the earlier australopithecines and has thicker bones. Its height may have topped five feet at best and it weighed up to 110 pounds. Its brain showed another small increase in size, and may have weighed about 550 grams. Most of the remains of A. robustus have been found in southern Africa.

Somehow allied to A. robustus is another large australopithecine that appeared at about the same time. It may be just a variety of A. robustus or it may be a descendant, or it may have evolved along with A. robustus from the Black Skull. We don't know enough, yet, to be able to say.

This fourth and last australopithecine is Australopithecus boisei. Its remains were discovered in east-central Africa in 1959 by an expedition sponsored and funded by a British businessman named Charles Boise—which accounts for its species name.

A. boisei is the largest of the australopithecines and some may have been as large and as heavy as a modern human being of average size. Its brain was no larger than that of A. robustus, however.

Brain growth in these larger australopithecines may not have matched the body growth so they may have been less intelligent than the smaller ones. That may explain why they were an evolutionary dead end. They died out about a million years ago, and left no descendants. For about 4 million years there had been australopithecines in eastern and southern Africa, and then they were gone.

. . . But not entirely.

23. 2,000,000 B.P. About 2 million years ago, there was a hominid that we don't consider to be an australopithe-

89

cine, but that clearly evolved from one. We can't be sure exactly which one because we don't have any remains, as yet, of the intermediate steps.

Most anthropologists seem to think that A. africanus split into two lineages. One led to the Black Skull, to A. robustus, A. boisei, and the dead end. The other led to a nonaustralopithecine and to all the species that followed—including us. That may well be so, but we can use more evidence and someday we may have it.

The new hominid is sufficiently like us to be placed in the same genus with us—Homo. In other words, genus Homo, of which we are part, seems to have come into being 2 million years ago.

The full name of the earliest hominid of genus Homo is Homo habilis, where habilis is from a Latin word meaning "skillful." (The English word "able" is a descendant of habilis.)

H. habilis was not as large as the larger aus-tralopithecines. In fact, in the summer of 1986 a set of fossil remains of H. habilis were discovered that were some 1.8 million years old. It was the first time that both skull fragments and limb bones of the same individual had been located, and they seem to represent a small, light adult about

three and a half feet tall, and with arms that are surprisingly long. It was more like an australo-pithecine than had been thought, but it's hard to go by one specimen. It may have been an undernourished runt.

In any case, though H. habilis may have been small, he had a more rounded head and a larger brain, which may easily have weighed seven hundred grams, nearly half that of a present-day human being. He had thinner skull bones and possibly he possessed the beginnings of Broca's convolution in his brain (a section governing the power of speech), so that if he could not talk, he could at least make a greater variety of sounds than the aus-

90

tralopithecines could. His hands were more like our modern hands, and his feet seemed to be completely modern. His jaws were less massive so that his face looked less apelike. On the whole, you can see why he was Homo and not Australopithecus.

And why "skillful"?

Tool-using and tool-making are not a primarily human ability. Chimpanzees can use branches to threaten an enemy. They can strip the leaves from a twig and use it to probe for termites. They can crumple up leaves to use as sponges. Undoubtedly, the australopithecines could do anything chimpanzees could do, and more too. They may even have cracked rocks, on occasion, to make use of sharp edges.

It was H. habilis, however, who finally got to using his hands to their full potential. They had been freed when the first hominid became bipedal two or three million years before and they had been growing more useful ever since. Perhaps the necessity of dealing with only two limbs in locomotion had freed increasing volumes of the enlarging brain for the delicate control of the fingers.

In any case, H. habilis was the first organism of any kind to make a big thing out of chipping and flaking different kinds of rocks to make tools of various kinds for chopping, scraping, cutting, and so on. With H. habilis and its skillful hands, in other words, came the birth of technology.

As in the case of the australopithecines, the remains of H. habilis are to be found in east-central and southern Africa, so that both in space and time, it overlapped the larger australopithecines.

Homo habilis, with its rock tools and larger brain, was more formidable than the australopithecines. Indeed, H. habilis seemed to have been the first hominid to become a hunter rather than a scavenger, and the hunting

91

may have included the australopithecines. It may be that H. habilis and its immediate descendants finished off the last australopithecines so that for the last million years all hominids without exception have been part of genus Homo.

24. 1,600,000 8.P. By 1.6 million years ago, H. habilis was gone, even before the australopithecines were. The australopithecines really vanished, however, and became extinct. H. habilis had evolved.

It had become a new species, Homo erectus, which was about as large and as heavy as modern human beings. H. erectus was the first hominid to attain a height of as much as six feet and to weigh over 150 pounds. (That is why I distrust that small specimen of H. habilis. Can the body have expanded so much in a mere 200,000 years?)

The brain was larger, too, with a weight of 800 to 1100 grams. The upper limit is three fourths the size of the modern human brain.

H. erectus made much better stone tools than had been built before and was an enormously successful hunter, taking on the biggest animals it could find— even the mammoth.

Undoubtedly, the last australopithecines must have fallen prey to H. erectus. If a few specimens of relatively unchanged H, habilis remained, off they went, too. Homo erectus, between 1 million and 300,000 years ago, was the only hominid species in existence.

H. erectus made two particularly enormous advances.

In the first place, all the hominids for perhaps as much as four million years had been confined to Africa and to the southeastern half of that continent, at that. H. erectus was the first hominid to expand that range significantly. About 500,000 years ago (my guess), it moved off into the rest of Africa, into Europe, into Asia, and even into the Indonesian islands.

92

In fact, the first discoveries of remains of H. erectus were in Java, where the Dutch anthropologist Marie Eugene Dubois (1858-1940)

discovered a skullcap, a femur, and two teeth in 1894. No hominid with so small a brain had yet been discovered, and Dubois named it Pithecanthropus erectus, (Greek for "erect ape-man").

A similar find was made near Peking, beginning in 1927, by a Canadian anthropologist, Davidson Black (1884-1934). He named his find Sinanthropus pekinensis (Greek for "Chinaman from Peking").

Eventually, it was recognized that both sets of remains along with some others were all of the same species and deserved to belong to genus Homo. Dubois's term, erectus, was kept even though hominids had been walking erect for as long as 3 million years before H. erectus had evolved. This, however, was not known in Dubois's time.

The second great advance made by H. erectus was the use of fire. Traces of campfires have been found in settlements of H. erectus. It is possible that fire had been made use of, in a kind of casual and opportunistic way, before H. erectus. H. erectus, however, was the first to use it systematically. It was the greatest technological advance since the making of stone tools.

25. 300,000 B.P. Hominids who were recognizably H. erectus in characteristics may have lived as recently as 200,000 years ago, perhaps even longer, but they had been evolving in the direction of still larger brain size. By 300,000 years ago, hominids had been developed with body and brain size as large as ours.

The first trace of such hominids was located in 1856 in the Neander Valley (Neanderthal in German) in Germany. Such hominids were therefore called Neanderthal men. Their skulls were distinctly less human than our own. They had pronounced eyebrow ridges, large teeth, protruding jaws, smoothly receding chins—all

93

rather resembling H. erectus. They were stockier than we are, and more muscular. Their brains were as large as ours, or a few percent larger, but were differently proportioned, heavier in back, and lighter in front.

They were at first termed Homo neanderthalensis, but they were so like us everywhere but in a few details of the skull that they were finally recognized as being of our species: Homo sapiens ("wise man" in Latin). And why not? There may even be evidence of their having interbred with "modern man."

Still, they are thought of as a subspecies and they are now termed

Homo sapiens neanderthalensis. In their early years, H. sapiens n. must have overlapped in time and place with those H. erectus hominids who remained. If so, the Neanderthals must certainly have wiped them out. From about 200,000 years ago till about 50,000 years ago, H. sapiens n. were the only hominids alive, and existed all over Europe, Asia, and Africa.

26. 50,000 B.P. "Modern man" is Homo sapiens sapiens, presumably an offshoot of the Neanderthals. H. sapiens s. is taller, more slender, and less muscular than H. sapiens n. His brain is a tiny bit smaller but is larger in the forepart which, we are free to think (but don't really know), gives us an intellectual advantage and makes us better able to indulge in abstract thought and elaborate speech.

Between 50,000 and 30,000 years ago, //. sapiens n. and H. sapiens s. coexisted, but by the latter date, intermarriage and, probably, slaughter had put an end to the Neanderthals and for 30,000 years we have been the only variety of hominid that has existed.

About 25,000 years ago, H. sapiens s. extended the human range again, penetrating the Americas and Australia where, till then, no hominid had ever stepped foot. By 1,000 years ago, human beings were living on every substantial piece of land except for Antarctica.

94

5fnHHOne.p 4 mHHon B.P. 3 million B.P. 2 mttfion B.P. 1 mHlton a P.

500,000 B.P. 400,000 a.p.— 300,000 B.P.

200.000 B.P. 100,000 B.P.—

50,000 B.P 40,000 B P.— 30.000 B.P. 20,000 B.P —

10,000 B.P.

5,000 B.P.

3,500 B.P

220 B.P.

    t

— >

-/ -t

~t

-

— —t

-

-

-

—

—

—

(First austratopittiedne?) "Lucy," A. afarenais

A.africanus

—A atoensis extinct, black skuM roouslus; A bofee/ Homo habHls
(some tools)

—-H. erecfus

Australopitnecines extinct

H. erectus uses fire; expands out of Africa

Neanderthal man H onctus extinct

"Modem man"

Neanderthal man extinct

Modem man enters Americas, Australia

Civilization History Iron Age

Industrial Revolution ^PRESENT

B.R.   BEFORE THE PRESENT Note Chert le not completely to scato.

About 10,000 years ago, H. sapiens s. began to practice agriculture, to herd animals, and to build cities—the beginning of civilization.

About 5000 years ago, writing was invented by the Sumerians—the beginning of history. Metallurgy followed.

By 3,500 years ago, iron came to be smelted and the age of large empires began.

By 500 years ago, gunpowder artillery, and the printing press were in full swing so that the worst and best of modern times were upon us.

By 220 years ago, the steam engine was on the way, and with it the Industrial Revolution.

By 40 years ago, nuclear weapons came into being, and by 30 years ago, the space age had begun—and here we are, possibly at the beginning of a new and vast extension of range, and possibly at the point of ending the hominid story altogether, 5 million years after it began.

7

The Longest River

One way of achieving an act of creativity is to look at something in an unexpected way.

Thus, for thousands of years the hole in the needle was put at the blunt end so that the thread followed like a long tail after the needle had pierced the cloth. But when people tried to invent a sewing machine, nothing worked until Elias Howe had the brilliant turnabout idea of putting the hole near the point of the needle.

We who write science fiction find a particular necessity in looking at

things differently, for we must deal with societies other than those that exist. A society that looks at everything in the same way we do is not a different society. After nearly half a century of science fiction writing, that sort of sideways squint has therefore become second nature to me.

97

At a meeting over which I was presiding a couple of weeks ago, a member rose to introduce his two guests.

He said, "Let me introduce, first, Mr. John Doe, who is a brilliant lawyer and an absolute expert in bridge. Let me also introduce Dr. Richard Roe, who is a great psychiatrist and a past master at poker." He then smiled bashfully, and said, "So you see where my interest lies."

Whereupon I said, quite automatically, "Yes, in working up lawsuits against psychotics," and brought the house down.

But to get to the point . . .

More than twenty years ago, I wrote an essay on the great rivers of the world ("Old Man River," F & SF, November 1966).* Ever since then, I've had it in my mind to devote an entire essay to just one river. Naturally, it would have to be the largest river of them all, the one that drains the greatest territory, the one that delivers the most water to the sea, the one that is so mighty that all other rivers are merely rivulets compared to it. The river I speak of is, of course, the Amazon.

Now the time has come, and even as I sat down, with satisfaction, to write the essay, the kaleidoscope I call my mind suddenly heaved, rattled, and changed shape. I thought: Why should I be impressed merely by size, by gigantism? Why shouldn't I devote myself to a river that has done most for humanity?

And which should that be but the Nile.

In one respect, the Nile is an example of gigantism. It is much smaller than the Amazon in that it delivers far less water to the sea, but it is longer than the Amazon. It is, indeed, the longest river in the world, for it is 6,736 kilometers (4,187 miles) long, compared to a length of

* See my book Science, Numbers, and t (Doubleday, 1968).

98

about 6,400 kilometers (4,000 miles) for the Amazon, which is second longest.

The difference between them is that the Amazon flows west to east along the equator, through the largest rain forest in the world. It is constantly being rained on and has, in addition, a dozen tributaries that are mighty rivers in their own right. By the time it reaches the Atlantic, then, it is delivering some 200,000 cubic meters (7 million cubic feet) of water per second, and its outflow can be detected over 300 kilometers (200 miles) out into the sea. The Nile, on the other hand, flows south to north, beginning in tropical Africa, but with its northern half flowing through the Sahara Desert without tributaries, so that it receives no water at all, but merely evaporates. No wonder it finally discharges into the Mediterranean only a small fraction of the water discharged by the mighty Amazon.

But the Sahara was not always a desert region. Twenty thousand years ago, glaciers covered much of Europe and cool winds brought moisture to northern Africa. What is now desert was then a pleasant land with rivers and lakes, forests and grassland. Human beings, as yet uncivilized, roamed the area and left behind their stone tools.

Gradually, however, as the glaciers retreated and the cool winds drifted ever farther north year by year, the climate of north Africa grew hotter and drier. Droughts came and slowly grew worse. Plants died and animals retreated to regions that were still wet enough to support them. Human beings retreated also, many toward the Nile, which, in that long-distant time, was a wider river, one that snaked lazily through broad areas of marsh and swamp and delivered far more water to the Mediterranean. Indeed, the valley of the Nile was not at all an inviting place for human occupancy until after it had dried out somewhat.

99

When the Nile was still too wet and swampy to be entirely enticing, there was a lake that existed to its west, about 210 kilometers (130 miles) south of the Mediterranean. In later times, this body of water came to be called Lake Moeris by the Greeks. It existed as a last reminder of a northern Africa that had once been much better watered that it was in later times. There were hippopotamuses in Lake Moeris and other, smaller game. From 4500 to 4000 B.C., flourishing villages of the late Stone Age lined its shores.

The lake suffered, however, from the continued drying out of the land. As its level fell, and the animal life it supported grew sparser, the villages along its shore withered. At the same time, though,

population grew along the nearby Nile, which became more manageable.

By 3000 B.C., Lake Moeris could only exist in decent size if it were somehow connected with the Nile and was able to draw water from the river. It required increasing exertion, however, to keep the ditch between the two dredged and working.

The battle to do so was finally lost and the lake is now almost gone. In its place is a depression, mostly dry, at the bottom of which is a shallow body of water now called Birket Qarun. It is about 50 kilometers (30 miles) long west to east and 8 kilometers (5 miles) long north to south. Near the shores of this last remnant of old Lake Moeris is the city of El Fayum, which gives its name to the entire depression.

To go on to the next step requires a small digression . . .

In 8000 B.C., human beings the world over were hunters and gatherers, as they had been for ages. The total population of the Earth may then have been only 8 mil-

100

lion, or about as many people as there are in New York City today.

But at about that time, some people in what is now called the Middle East learned how to plan for the future where food was concerned.

Instead of hunting animals and killing them on the spot, human beings kept some alive, cared for them, encouraged them to breed, and killed a few—now and then—for food. They also got milk, eggs, wool, skins, and even work, out of them.

Again, instead of just gathering what plant food they came across, human beings learned to sow plants and care for them, so that eventually they could be harvested and eaten. Clearly, human beings could sow a much greater concentration of edible plants than they were likely to find in a state of nature.

By herding animals and farming plants, groups of human beings vastly increased their food supply, and their population grew rapidly. Increasing population meant that more plants could be grown and more animals cared for so that, in general, there was a surplus of food, something that never happened (except for brief periods immediately after a large kill) in the old hunting and gathering days.

This meant that not everyone had to labor at growing food. Some could make pottery and exchange it for food. Some could be metalworkers. Some could be tellers of tales. In short, people could begin to specialize and society began to gain variety and sophistication.

Of course, farming had its penalty. As long as one merely hunted and gathered, one could avoid conflict. If a stronger band encroached on a tribe's territory, it could prudently retreat to some safer place. Not much was lost in the process. The tribe only owned what it could carry and it would take that along.

Farmers, however, owned land, and that was immov-

101

able. If marauding bands, intent on stealing the farmers' food stores, swooped down, the farmers had no choice but to fight. To retreat and give up their farms would mean starvation since there were now too many of them to be supported by any means other than farming.

This meant that farmers had to band together, for in union there was strength. Their houses were built in clusters. They would choose some site with a good natural water supply, and surround their houses by a wall for security. They then had what we would today call a city (from the Latin civis). The inhabitants of cities are citizens, and the kind of social system in which cities are prevalent is called civilization.

In a city in which first hundreds and then thousands of human beings clustered, it would be difficult to live without people stepping all over each other. Rules of living had to be set up. Priests had to be appointed to make those rules, and kings to enforce them. Soldiers had to be trained to fight off marauders. (See how easily we recognize the coming of civilization.)

It is hard to tell now just exactly where agriculture got its very first start. Possibly this was on the borders of the modern nations of Iraq and Iran (the very border over which both nations recently fought a useless war for eight years).

One reason for supposing that area was the place where farming (or agriculture, as it is commonly called) begant is that barley and wheat grow wild there, and it is just those plants that lend themselves to cultivation.

There is a site called Jarmo in northern Iraq that was uncovered in 1948. The remains of an old city were found there, revealing the

foundations of houses built of thin walls of packed mud and divided into small rooms. The city may have held from one hundred to three hundred people. In the lowest and oldest layer, dating back

102

to 8000 B.C., evidence, of very early farming was uncovered.

Once discovered, of course, the techniques for agriculture spread out slowly from the original center.

What was needed for farming, first and foremost, was water. Jarmo is at the edge of a mountain range, where rising air cools and where the water vapor it holds condenses out as rain. However, even at best, rain can be unreliable and a dry year will mean a lean harvest and hunger, if not starvation.

A supply of water that is more dependable than rain is that which you get out of a river. For that reason, farms and cities grew up along the banks of rivers, and civilization began to center there.

The nearest rivers to the original farming communities are the twin rivers of the Tigris and Euphrates in what is modern Iraq, and this may therefore have been the site of the earliest large-scale civilization, but it soon spread westward to the Nile and by 5000 B.C. both areas were flourishing. (Agriculture also spread to the Indus. Some thousands of years later, it began independently in the Hwang-ho region of northern China. Some thousands of years later still, it began among the Mayans of North America and the Incas of South America.)

The crucial discovery of writing, which took place not long before 3000 B.C., was made by the Sumerians, who then lived along the lower reaches of the Tigris-Euphrates valley. Since the use of writing is the boundary line between prehistory and history, the Sumerians were the first people to have a history. The technique was quickly picked up by the Egyptians, however.

Living on a river may mean that farmers have an unfailing source of water whether it rains or not, but the water won't come to the farmer of its own accord. It

103

must be brought there. To do so in pails is clearly ineffective, so one must dig a ditch into which river water can run of its own accord and maintain that ditch to keep it from silting up. In the end, a whole

network of such irrigation ditches must be built up, with raised banks along them, and along the river, to prevent too-easy flooding.

Taking care of such an irrigation network requires a careful and well-coordinated community effort. This places a premium on good government and capable leadership. It also places a premium on cooperation between the various cities along a river, since if a city upstream is wasteful of water, or pollutes it, or allows flooding, that will harm all the cities downstream. There is a certain pressure, therefore, to develop a river-wide government, or what we would refer to as a nation.

Nationhood came first to Egypt, and the reason is the Nile.

The Nile is a placid river, not given at all to violent moods. This means that even primitive boats, inefficient in design and fragile in structure, can float on the Nile without trouble. There is no fear of storms.

What's more, the water flows northward and the wind usually blows southward. This means that one can hoist a simple sail if one wishes to be blown upriver (south) and then take it down if one wishes the current to carry one downriver (north). Thanks to the quality of the Nile, then, people and goods could easily move from city to city.

Such movement up and down the river ensured that the city-states would share a language and a culture and feel a certain economic interdependence and communal understanding.

As for the Sumerians, they had two rivers. One, the Tigris, was too turbulent to be navigable by simple means (hence, its name, tiger). The Euphrates is more

104

easily handled and the major Sumerian cities therefore lined its banks. Still, it was not quite the placid highway that the Nile was and the Sumerian cities felt more isolated than the Egyptians did and were therefore less prone to cooperation.

Furthermore, while the Nile was bounded on both sides by deserts that kept outsiders at bay, the Euphrates was less well protected and more open to raids and to settlement by surrounding peoples. This meant that the Tigris-Euphrates valley contained also Akkadians, Arameans, and other peoples whose language and culture were different from that of the Sumerians, while the population along the Nile, on the other hand, was quite uniform.

Consequently, it is not surprising that Egypt was unified before the peoples of die Tigris-Euphrates were. Somewhere around 2850 B.C., a ruler named Narmer (known as Menes to the Greeks) united the cities of the Nile under his rule and established the nation of Egypt. We don't know the details of how this was done, but it seems to have been a relatively peaceful process.

The Sumerian cities, however, fought each other viciously and the region was not unified till 2360 B.C., five centuries after the Egyptians had been. What's more, the Sumerians had fought themselves into war-weary weakness so that the union was brought about by a non-Sumerian, Sargon of Agade. He established his rule by harsh conquest and brought under his banner a variety of languages and cultures so that Sargon's unified kingdom was not a nation but, rather, an empire.

An empire tends to be less stable than a nation, as the dominated ethnic groups feel resentment against the dominating one. The Tigris-Euphrates valley therefore saw a succession of upsets as first one group and then another gained predominance, or as raiders from outside took advantage of internal disunion to establish

105

themselves. Egypt, in contrast, was an extraordinarily stable society for its first twelve centuries of nationhood.

Then there is the matter of the calendar.

Primitive people use the Moon for the purpose, since the Moon's phases repeat themselves every 291/2 days. That is a period that is short enough to handle and long enough to be useful. It gives us the lunar month, which can be 29 and 30 days long in alternation.

Eventually, it was noted that every twelve months or so the seasons went through their cycle. Twelve lunar months after sowing time, it was sowing time again, in other words. Of course, the seasons are not as reliable as the phases of the Moon. Springs can be cold and late, or mild and early. In the long run it was clear, however, that twelve lunar months (which have a total of 354 days) was not quite long enough to mark the cycle of the seasons. After two or three years, a lunar calendar of this sort would indicate sowing time so much earlier than it should be that it would lead to disaster.

For that reason a thirteenth month had to be added to the year every now and then if the lunar calendar was to be kept even with the

seasonal cycle. Eventually, a nineteen-year cycle was established within which twelve years had, in a certain fixed order, twelve lunar months, or 354 days, and seven had thirteen lunar months, or 383 days. This meant that, on the average, the year was 365 days long.

This calendar was awfully complicated, but it worked, and it spread to other peoples, including the Greeks and the Jews. The Jewish liturgical calendar to this day is the one developed by the people of the Tigris-Euphrates.

The early Egyptians were aware of and used the lunar months, but they were also aware of something else. The Nile (as we know, but they did not) rises among the

106

mountains of east-central Africa. When the rainy season comes to that distant region, water tumbles into the lakes and rivers and surges down the Nile.

The level of the Nile rises and the river floods over its banks for a period of time, leaving a deposit of rich, fertile silt behind. It is the Nile flood which ensures the harvest, and the Egyptians awaited it eagerly, for in the years when it was late, scanty, or both, they would see hard times.

The close attention Egyptians paid to the flooding of the Nile made them realize that it came, on the average, every 365 days, and it seemed to them that it was this period that was of overwhelming importance. They, therefore, adopted a solar calendar. They made every month 30 days long, so that twelve of them marked 360 days, and added five monthless holidays at the end before starting another cycle of twelve months. In this way the months were calendar months that were out of step with the Moon but in step with the seasons.

Actually, it was not quite in step with the seasons. The year is not 365 days long, but very close to 3651/4. The Egyptians could not help but understand this, for every year the Nile flood came six hours later (on the average) according to the Egyptian calendar. This meant that the date of the flood wandered through the entire calendar and returned to the original date only after 365 x 4, or 1,460 years.

This wandering could have been prevented by adding a 366th day to the year every four years, but the Egyptians never bothered to do this. However, when the Romans finally adopted the Egyptian calendar in 46 B.C., they spread those five extra days through the year, giving some months 31 days, and added an extra day every four years. That

(with very minor modifications) is the calendar the whole world uses today—for secular purposes, anyway.

107

The Nile flood sometimes wiped away the markings that separated the holdings of one family from those of another. Methods had to be devised to redetermine those boundaries. It is thought that this slowly gave rise to the methods of calculation that we know as geometry (from Greek words meaning "to measure the Earth").

Those same floods assured Egypt of so much food that it could afford to trade the surplus to surrounding peoples not blessed with the Nile and to get in exchange foreign artisanry. The Nile thus encouraged international trade.

What's more, with the large food surplus, it was not necessary to put every pair of hands to work growing food. There was an ample labor supply to be put to the task of what we would today caft public works. The prize example, of course, was the raising of the Pyramids between 2600 and 2450 B.C.

It may be that the Pyramids set the example of gigan-tism in architecture in the Western world. The latest manifestation of this I can see from my apartment windows—the total conversion of Manhattan into traffic-choking skyscrapers.

To my way of thinking, then, the Nile has given us one of the two earliest civilizations, boats, the first nation, the solar calendar, geometry, international trade, and public works. It has also given us a mystery that has intrigued human beings for thousands of years. Where does the Nile originate? What is its source?

The ancient world of western Asia and the Mediterranean knew of seven rivers with lengths of 1,900 kilometers (1,180 miles) or more. Leaving out the Nile, the other six, together with their lengths, are:

108

Euphrates—3,600 kilometers (2,235 miles) Indus—2,900 kilometers (1,800 miles) Danube—2,850 kilometers (1,770 miles) Oxus—2,540 kilometers (1,580 miles) Jaxartes—2,200 kilometers (1,370 miles) Tigris—1,900 kilometers (1,180 miles)

The Persian Empire included the Tigris and Euphrates in their totality. The Oxus and the Indus were at the eastern extremity of that Empire and the Jaxartes was just beyond the northeastern boundary.

The Danube formed the northern boundary of much of the European dominions of the Roman Empire. The source of each of these rivers was known as a matter of public knowledge, or, in the case of the Oxus and Jaxartes, from travelers' reports.

That left the Nile. It was the core of Egypt from the beginning and was included eventually in the Persian Empire and, still later, in the Roman Empire. The Nile, however, was twice as long (as we now know) as the longest of the other rivers and it extended outside the limits of civilization right down into modern times, so that in all that time no one knew where the source was.

The Egyptians were the first to wonder. About 1678 B.C., the land was invaded by Asians who were using the horse and chariot for warfare—something the Egyptians had not encountered before. The Egyptians finally managed to throw them out about 1570 B.C.

In reaction, the Egyptians invaded Asia in its turn and established the Egyptian Empire. For nearly four centuries, Egypt was the strongest power in the world.

Under the Empire, the Egyptians expanded up the Nile. The Nile has occasional sections of rough water (cataracts) that are numbered from the north to the south. The First Cataract is at the city known as Syene to the ancient Greeks and as Aswan to us today. This is

109

885 kilometers (550 miles) south of the Mediterranean. It was a navigation problem and Egypt proper did not extend south of the First Cataract. Even today, modern Egypt extends only about 225 kilometers (140 miles) south of the cataract.

South of the First Cataract was a nation called Nubia. Tbday it is called Sudan. Occasionally, strong Egyptian monarchs had attempted to extend their dominion beyond the First Cataract, and under the Empire that effort reached its maximum. The Empire's greatest conqueror, Thutmose III, penetrated, about 1460 B.C., to the Fourth Cataract, where the Nubian capital of Napata stood.

Napata is about 2,000 kilometers (1,250 miles) upstream from the mouth of the Nile and the river was still going strong, still mighty, showing no signs of dwindling to its source.

The later conquerors of Egypt—the Ptolemies, the Romans, and the Muslims—made no effort to extend their political control south of the First Cataract. If anyone explored southward, no coherent account of

their travels remains.

The first modern European to venture south of Aswan was a Scottish explorer, James Bruce (1730-94). In 1770, he traveled to Khartoum (the modern capital of Sudan), which is about 640 kilometers (400 miles) upstream from the ruins of Napata. There two rivers join to form the Nile. One (the Blue Nile) comes in from the southeast; the other (the White Nile) comes in from the southwest.

Bruce followed the Blue Nile upstream for something like 1,300 kilometers (800 miles) and finally came to Lake Tana in northwest Ethiopia. He felt that to be the source of the Nile, but he was wrong. The Blue Nile is merely a tributary. It is the White Nile that is the main stream.

Arab traders had brought back vague tales of great lakes in East Africa and some European explorers thought that those might well be the source of the White Nile. Two English explorers, Richard Francis Burton (1821-90) and John Manning Speke (1827-64), started from Zanzibar on the east African coast in 1857 and by February 1858 reached Lake Tanganyika, a long, narrow body of water 1000 kilometers (620 miles) from the African coast.

By then, Burton had had enough and left. Speke, however, moved northward on his own and, on July 30, 1858, reached Lake Victoria. This is 69,500 square kilometers (26,818 square miles) in area, so that it is a little larger than West Virginia. It is the largest lake in Africa and the only freshwater lake that is larger is Lake Superior, which has an area one fifth greater than that of Victoria.

The White Nile issues from the northern rim of Lake Victoria, which can thus represent the source of the river. However, the longest river that flows into Lake Victoria is the Luvironza, which is 1150 kilometers (715 miles) long, and flows into the lake from the west.

A drop of water from the headwaters of the Luvironza could flow into Lake Victoria and out again into the White Nile and from there to the Mediterranean, traveling 6,736 kilometers (4,187 miles).

The source of the Luvironza is, therefore, the source of the Nile and it is located in the modern nation of Burundi, about 55 kilometers (35 miles) east of Lake Tanganyika. When Burton broke away, he was almost at the source of the Nile.

But, then, how was he to know?

110

111

AFTERWORD

This essay is a rather quiet and noncontroversial one, but it approaches history from my own, somewhat unusual, point of view.

History, like mathematics, is something I love more than it loves me. When I was in college, as a matter of fact, I debated with myself whether to major in history or in chemistry. I decided on chemistry, because I felt that as a historian I would be condemned to the academic life, whereas as a chemist I might go out into industry or into a research institute.

This was unbelievably foolish of me, for when I finally became a chemist, I realized that industry was not for me and I remained in the academic life.

I have, however, never forgotten history, for I have written many history books as well as many scientific books, and even when I discuss science I tend to approach it historically. I am so grateful that my publishers tend to humor me and publish whatever I write so that I can indulge all my various penchants—chemistry and history (and everything else that catches my fancy as well).

8

Is Anyone Listening?

112

Everyone who has reached my level of late youth and has spent his time watching people and listening to them is bound to have become cynical. I, too, have become cynical. I have difficulty accepting things according to surface appearances, and have trouble believing promises and assurances.

And even so I get stuck on occasion.

It seems that a small plot of land on Manhattan's Upper West Side was going to waste. It was just a ravished lot. Some public-spirited citizens of the neighborhood managed to have it set aside for public use. A garden was planted, benches were introduced, and I received a phone call from a woman who asked me, as a prominent resident on the southern fringe of the Upper West Side, to come down and preside over the groundbreaking ceremony.

113

I said, "I'd love to, but the date you suggest is a Tuesday, and every Tuesday I make my rounds of my publishers and then preside over a weekly luncheon of an organization of which I am president."

A few days later she phoned a second time and said the date had been changed to Thursday, at 10 A.M. I apologized again, for I was slated to do a phone interview on that day from 10 to 11 A.M.

There then came a third phone call. The time had been changed to 11:30 A.M. and I said, "Good! I'll be there."

After that I received several letters, a pamphlet of detailed information on the garden, and, on Thursday morning, there came a final phone call to make sure that I was in good health and hadn't forgotten. I said, "Fear not. I will be there by eleven-thirty. In fact, I plan to be there a bit earlier so that you won't have cause to worry."

"That will be wonderful," she said.

As soon as my phone interview was over, therefore, I collected my dear wife, Janet, and we taxied to the garden. We were there at 11:20 A.M., ten minutes early, as I had promised, and, to my surprise, the festivities were over and done with, and everyone was departing.

I asked for the woman who had phoned me. She was pointed out to me. I approached her and said, "I'm Isaac Asimov and I'm here early," and showed her my watch.

She stared at it and at me for a moment as though trying to place me, and then she called out, "Isaac Asimov has just arrived. Come back." (As though I were late and they had given up on me.)

Some people came back, rather reluctantly. I was asked to say a few words. I spoke for about ten seconds and that seemed to be enough. The group left even faster, clearly annoyed at having been delayed.

114

Janet asked, "What was all that fuss about getting you here?"

"I don't know," I explained, and we walked to a favorite restaurant that happened to be not too far away and buried our sorrows in a good lunch.

Another and greater justification for cynicism is that people don't listen, even when warnings are explicit, and even when the outlook is threatening.

On October 27, 1987, the New York Times, in its weekly science section, ran a rather long item under the headline: "Indispensable Helium Is Routinely Squandered."

The article pointed out that three fourths of the helium produced in the United States (which has more than 90 percent of the world's supply) is allowed to escape into the atmosphere, from which it is all but impossible to retrieve it. From the atmosphere, it leaks into outer space, from which it is quite impossible to retrieve it. Yet helium is extremely important, and for some purposes, such as the continuing investigation of extremely low temperatures, it is indeed absolutely indispensable. There is and can be no substitute.

Do you suppose that people are now going to rise up and demand that helium be conserved?

Nonsense! This is not a new story. The New York Times may have just discovered this fact, but I once published an essay that mentioned the wastage of helium and strongly warned of the consequences, and of the necessity of conservation. It was in an essay entitled "The Element of Perfection" and it appeared in the November 1960 issue of F & SF a mere twenty-eight years ago!*

* See my book View from a Height (Doubleday, 1963).

115

Was anyone listening? Did anyone care? . . . Very few.

I have devoted at least two essays inF&SF to warning of Earth's growing population. In the May 1969 issue, for instance, there was my essay, "The Power of Progression, "f

At that time, Earth's population was 3.5 billion, as compared to about 2 billion at the time of my birth nearly half a century earlier. In that half century, it had increased by 75 percent.

In the May 1980 issue of F & SF, I published "More Crowded!"^ At that time, the Earth's population was 4.2 billion, so that in eleven years the number of human beings had increased by 700 million, which is very nearly the present population of India. In eleven years, in other words, we had added another India to the world, from the standpoint of numbers at least.

In "More Crowded!" I made the following statement: "It is quite likely that we will end the decade of the 1980s with a world population edging toward 5 billion."

As usual, I was conservative. We are not edging toward the 5 billion mark, we have passed it. And we have done this not by the end of the 1980s, but by the time the decade was only three fourths done. The Earth passed the 5 billion mark some time late in 1986 or early in 1987. (We can't possibly know the exact date of this accomplishment because so much of the world is so poorly censused.)

In the seven years after "More Crowded!" then, the Earth added 800 million people, 100 million more than it had added in the previous eleven years. In the eigh-

t See my book The Stars in Their Courses (Doubleday, 1971). $ See my book The Sun Shines Bright (Doubleday, 1981).

116

teen years between 1969 and 1987, the Earth's population grew by 1.5 billion people (as much as it had gained in the previously years) and that is equal to the population of two present-day Indias. What's more, since the birthrate in the poor and industrially undeveloped nations is far higher than in the long-industrialized ones, about 90 percent of the new mouths are born in the poor nations. We have therefore added two Indias not only in terms of numbers, but in terms of poverty as well.

And this has taken place despite the fact that the rate of increase has dropped from 2 percent a year in 1970 to 1.6 percent a year now, thanks chiefly to stern measures taken in China to reduce the birthrate.

Are we entitled to be relieved at the drop in birthrate? . . . No, for the increase in population more than compensates for that. An increase of 2 percent a year in 1969 when the population was 3.5 billion meant an increase of 70 million that year. An increase of 1.6 percent a year in 1987 when the population was 5 billion meant an increase of 80 million that year. So we're worse off now both in total numbers and in numbers of increase.

Let's take a closer look. An increase of 80 million people in one year means an additional Mexico in a year. That is equivalent to 220,000 new people every day, or one new Lima, Ohio, every time you wake up in the morning. It is also equivalent to 150 additional people every

minute or 5 additional people every two seconds. If we had a digital recording on which the Earth's population could be read off at each instant, the units figure would be flipping up new digits at more than twice the rate that the seconds figure would change on a digital watch.

Is anyone listening? Does anyone care?

117

In "The Power of Progression," I began with a world population of 3.5 billion, and a doubling rate of once every forty-seven years, and worked out an equation that would give me the world population at any time in the future, provided the doubling rate stayed constant.

I showed that by 2554 A.D., the world population would be 20 trillion, so that the average population density over the entire surface of the Earth, land and sea, would be equal to the average density, today, of Manhattan at noon.

I then assumed that every star in the Universe had ten habitable planets and that we could transfer people from planet to planet at will and instantaneously. By 6170 A.D., every planet in the Universe would be filled to Manhattan density.

I then imagined that all the mass in the Universe could be converted into human flesh and blood, and it turned out that if this could be done without limitation, then by 8700 A.D., the entire Universe would be nothing but a mass of humanity.

Since the birthrate has dropped since 1969, we can calculate the doubling rate right now at once every fifty years. This gives us a little more time. It won't be till 2585 A.D. that we achieve Manhattan density over all of Earth's surface, and not till 9050 A.D. that we convert the entire Universe into nothing but humanity.

Obviously, this is far from enough. There's no question of converting the entire Universe into nothing but human flesh and blood by 8700 A.D., and giving us another three hundred and fifty years to do it in is not going to help one iota. For that matter, we can't possibly live on an Earth that is all at Manhattan density by 2554, and giving us forty years extra won't help either. We can't continue multiplying at this rate for very

118

long, no matter what we do. It won't help us to advance technology by any conceivable amount. For instance, it won't help us to go out into

space at any conceivable rate. After all, since we're going to have 80 million people more in a year, when will we be able to put that many people in space in one year so as to stabilize the population? Do you want to be optimistic and say we can do that fifty years from now? . . . Well, by then we'll be gaining 160 million new mouths on Earth every year and the people in space will be multiplying, too.

Don't get me wrong. I'm not saying that we will maintain this population increase indefinitely, because we won't. We won't for the best and most insuperable reason in the world. We won't because we can't. The only question about population that we can ask is how we will stop the population increase.

And the answer to that is that either (a) the birthrate will continue to decrease, or (b) the death rate will increase, or (c) both will take place. There are no other alternatives. I've said that before and I'm now saying it again.

Is anyone listening? Does anyone care?

The feeling on the part of demographers is that by the year 2000 the population will begin to level off and that by 2100 it will reach stability, though by that time the population will have reached some 10 billion, or twice what it is now.

Is that a big sigh of relief I hear?

Then think! What kind of a world will it be by the time population stability is achieved?

The population of the Earth is not going up evenly. I said earlier that 90 percent of the population increase is in the underdeveloped nations. What's more, within those nations, the rural areas, as population multiplies,

119

are ground ever deeper into poverty. With land less and less available, the peasantry drifts into the cities in search of jobs, so that the cities of the underdeveloped nations are growing at a cancerous rate.

In "More Crowded!," written in 1980,[1] expressed my surprise that the second largest city in the world was Mexico City. Between 1967 and 1979, its population had gone from 3,193,000 to 8,628,074. In merely twelve years it had increased its population nearly threefold, going from the size of Chicago to more than the size of New York.

The latest figures I could find show that its population is now 13 million, and that is probably low. I have heard larger figures given. In any case, it is now the most populous city in the world.

Before World War II, London was the largest city in the world, with a population of 8 million, and New York City was second, with 7 million. New York has kept its population at that height (with its suburbs growing rapidly, of course) and London has actually shrunk.

New York is now, according to the latest statistics I can find in 1988, only the fourteenth-largest city in the world and London is the sixteenth-. Here are the figures, which I imagine err on the low side:

Mexico City, Mexico   13,000,000

Sao Paulo, Brazil 12,600,000

Shanghai, China 12,000,000

Cairo, Egypt   12,000,000

Seoul, South Korea   9,600,000

Peking, China   9,300,000

Calcutta, India   9,200,000

Rio de Janeiro, Brazil   9,000,000

Tokyo, Japan   8,500,000

Bombay, India   8,300,000

Moscow, U.S.S.R.   8,000,000

120

Tientsin, China Jakarta, Indonesia New York, U.S.A. Canton, China London, U.K.

7,900,000

7,700,000

7,200,000

6,800,000

6,700,000

Of the thirteen largest cities in the world, one is in Africa, three are in Latin America, and eight are in Asia. Only one is in Europe, and that is Moscow.

None of this alters the fact that the richest very large city remains New York, and this is significant. Size does not necessarily mean wealth. In fact, the very large cities in the nonindustrial countries tend to contain square mile upon square mile of hovels, shacks, and shanties deprived of any of the amenities that an average dweller in a large city in an industrialized nation takes for granted.

And this will only get worse. Fast though the world's population is growing as a whole, and still faster though the world's underdeveloped population is growing, the fastest growth rate is in the cities of the underdeveloped nations. By 2000, even though the population will begin to be moving into its stabilizing period, the cities of the underdeveloped nations may still be expanding and may have collapsed into rotting nightmares.

Consider, too, that the terrible need for agricultural land forced by the population increase, together with the need for firewood (which is the most important fuel in many underdeveloped areas), is already resulting in the slaughter of the forests, particularly the rain forests, which are being hacked down at a fearful rate. Almost 15 million acres of forest are being cleared each year and, by the year 2000, half the present forests of Earth may be gone.

Remember that forests aren't just pretty trees taking up land that might better be used by human beings.

121

Forests have root systems that conserve the soil and prevent the violent runoff of excess water. The trees give off water into the air instead, cooling and moistening it in this way. Forests also produce oxygen at a rate higher than will any form of vegetation replacing them.

The soil in which rain forests grow is not very good and will be soon leached of nutrients by crops planted in them, while rain runoff will gully and destroy the soil altogether. Far from supplying us with agricultural land, the vanishing rain forests will yield to deserts.

The deserts are indeed expanding as a result of forest destruction,

overfarming, and general human mishandling, and, by the year 2000, the area of new desert will be perhaps one and a half times the area of the United States. Arid the fact that there will be less and less good land to cultivate will send more and more people into the overcrowded, festering, fetid cities.

The forests, too, are the habitat of myriads of species of plants and animals, a couple of million of which (mostly insects, to be sure) have not yet even been classified. The destruction of the forests destroys habitats and about a fifth of the animal and plant species now living will be extinct by the year 2000.

This is not something to be dismissed lightly. Such extinctions will upset the ecological balance and wreak havoc far beyond the actual extinctions themselves. There is also the question of what compounds of important medicinal and industrial value might exist in the plants and animals we have not yet investigated, and which will vanish forever together with whatever good they might have done us.

Then, too, the more people there are, the greater the rate at which we must consume the Earth's finite resources. Worse yet is the fact that the more people there are, the greater the rate at which we must produce waste products, many of them toxic.

122

Usable fresh water supplies will decrease, since larger and larger portions of them will be polluted to the point where they will be undrinkable without expensive treatment that many regions will not be able to afford. Nor will life be able to thrive in polluted water. Acid rain will grow worse and kill more lakes and more fish.

Even the ocean rim, which is the richest portion of the sea, is being increasingly polluted (and remember that microscopic forms of plant life in the uppermost layers of the ocean produce 80 percent of the oxygen that we cannot do without).

The atmosphere, too, is being increasingly polluted, and cities are becoming more and more smog-bound.

Even carbon dioxide, which is itself a rather benign and relatively nontoxic substance, is a deadly danger. The fuels we burn for energy at an ever-increasing rate are producing carbon dioxide at a rate greater than the Earth's vegetation can utilize it and the Earth's ocean dissolve it. The result is that the percentage of carbon dioxide in the air (quite low today—only 0.035 percent) is slowly, but steadily, increasing from year to year.

By 2000 A.D., the carbon dioxide content of the air may have increased by one third as compared with today's content. This won't interfere with our breathing noticeably, but it will conserve more of the heat the Earth receives from the Sun so that Earth's average temperature will go up somewhat. This will change the weather pattern, probably for the worse, and increase the rate at which the polar ice caps melt, raising the sea level noticeably and allowing coastal areas to suffer more from high tides and storms.

Other forms of pollution are just as slowly and just as surely destroying the ozone layer in Earth's upper atmosphere. This will increase the intensity of ultraviolet light from the Sun at Earth's surface. Usually, the warning here is that skin cancers will increase, and so they

123

will, but there may be worse. We don't know what the additional ultraviolet will do to the microscopic forms of life living in the soil and in the uppermost layers of the ocean. If these are badly damaged, the very viability of the Earth as a planet may be decreased markedly.

To be sure, Earth's resources may be made more efficient use of and wastes may be more rationally disposed of, if we make the social and technological effort, but there is a limit to what can be done if we continue to pour tens of millions of new human beings onto Earth's surface each year.

And as the population increases, as people crowd together more closely, as people find they can only get a smaller and smaller part of a pie that does not increase as the numbers do (but decreases in many ways), there will be increasing alienation, increasing refuge in drugs, increasing crime, increasing chance of war. In short, the world will become an ever-more-violent place.

Every one of these changes, which comes about more or less directly because of the ever-increasing population, will serve to raise the death rate. There will be increasing starvation and bodies weakened by undernourishment will be more prone to disease. There will be more deaths by violence. In short, the Four Horsemen of the Apocalypse (Famine, Pestilence, War, and Death) will ride the Earth.

This might seem a natural way to make overpopulation self-limiting. It will seem an automatic cure—but what a horrible cure it will be. Surely, the alternative of a deliberate effort to lower the birthrate is far preferable.

But is anyone listening? Does anyone care?

Some people, far from listening and caring, actually advocate a rise in the birthrate. The unbelievable Pat

124

Robertson, who attempted to secure the Republican presidential nomination, is in favor of such a rise. He shares the feeling that the more hands and brains there are, the more that can be done and thought. Moreover, he feels that we need more young people who actually do the work, since with a low birthrate, the elderly multiply and they are a drain on society.

Surely, the reverse is true. In an overpopulated society, people grow old and weak at a faster rate, and the young are relatively weaker, too. Fertility and population growth past a certain point are obviously not strengthening. India, Bangladesh, and the nations of Africa and Latin America prove the weakening effect of overpopulation all too well.

A nation that is not crowded can deal with its problems. An advanced medical technology that is not overwhelmed by more problems than it can handle does more than merely keep the aging alive. It keeps them stronger, healthier, and more mentally alert. The aging are not necessarily so great a drain on society in that case.

To look at it in reverse, what good is the increasing number of hands, if the hands are weaker? What good is the increasing number of brains, if those are damaged by undernourishment? Do you want lots of stuff, or good stuff? You can't have both.

Some people are constantly afraid of "race suicide." (That's rather like constantly fearing a drought during a downpour that has been continuing without a break for centuries.)

Some people fear that a neighboring country with a /higher birthrate will outbreed their own country and Jlake them over. Illegal immigrants from Mexico, for in-* stance, flood into the American Southwest in unstop-|pable numbers. They don't do it, however, to "take us vover." They are fleeing a land that can't feed them be-

125

cause of its steady increase hi population—and they are coming into a land which is willing to offer them jobs at wages lower than native-born Americans will accept. If Mexicans remain at the lowest rungs of

the economic ladder, they can't "take us over," whatever their numbers. If they climb the ladder, they will be as useful to society as other immigrants have been.

If Mexico were to stabilize its population, the pressure for emigration would be greatly reduced, for people, generally, given a chance would rather remain with their own society.

We. Most nations now have soldiers that are never used for any purpose but to kill their own countrymen in case of unrest. For this, is it necessary to increase population and destroy the world?pumuon

But is anyone listening? Does anyone care?

Is a high birthrate necessary to supply cannon fodder? In the days before World War II, France's birthrate was down while Nazi Germany offered prizes to women with many children. The Nazis were openly rewarding women for producing cannon fodder, reducing them to the status of brood mares.

To be sure, Germany then overwhelmed France in 1940, but it wasn't a matter of more men; it was a matter of better organization, better technology, and better generals. The proof is, first, that Italy also offered prizes for children in the 1930s and it did them no good whatever. And that, between 1750 and 1850, Great Britain overwhelmed India, though India was far larger than Great Britain in both area and population.

Nowadays, mere numbers in wars amount to less and less. What counts are, on the one hand, trained guerrillas, relatively few in number, and, on the other, advanced machines and relatively small numbers of specialists who can handle high tech. And, in the end, of course, there are nuclear bombs that are quite able to kill virtually everyone without distinction, thus making numbers irrelevant.

In short, wars in the old-fashioned sense are impossi-

126

127

Part Radiation

9

The Unrecognized Danger

To possess the power of concentration is to have a useful tool.

When, as a teenager, I began to write seriously, I was living in the bosom of a family who were crowded together in an inconvenient apartment and who were, one and all, not the least concerned that I was writing.

I had no choice, therefore, but to work with blaring radios and arguing voices and, since I lived on the second floor in a building on a busy Brooklyn street, there was also the obligato of booming traffic and the shrieks and shouts of playing children.

I had to learn to ignore it all and, to this day, I can work unperturbed in the midst of miscellaneous activity. When I am perforce diverted from my work, no "mood" has been broken. I simply tend to the diversion in a

131

more or less absent-minded manner and, without trouble, pick up my work at the point where I had left off.

Very useful, indeed.

But it does mean that I tend to be unaware of things that go on about me, even when I am not working. After all, when I am not engaged in doing something, I am thinking about doing something so that the rest of the world recedes. This is a process that is not without serious risk, especially when I am treading the streets of New York.

Someone who was aware of my absent-mindedness was my beautiful, blond-haired, blue-eyed daughter, Robyn, who, from an early age, would discuss my peculiarities with her friends, but who loves me anyway. She once said to me, "I've spent my whole life laughing," which is a good way to spend your life, I think. I hope that she has spent much of her life laughing with me, for I do a lot of laughing myself, but I suspect she has spent some of her life laughing at me.

One example of my preoccupation took place about a dozen years ago when I was giving a talk at Boston University. I was perfectly well aware that Robyn was attending Boston University at the time and I rather expected to see her in the audience. However, the hall was crowded, and I didn't spy her, and once I began to give my talk with my usual concentration on the task at hand, I forgot about her.

After the talk, a number of students crowded about me to ask questions and I was answering with great vivacity, as I usually do, and

with only the vaguest awareness of my surroundings.

Very casually, I noted a beautiful, blond-haired, blue-eyed young woman standing nearby, but my eyes slid over her without pause. This happened several times, until a vague feeling of having missed something pervaded me. I turned back to the young woman, stared a

132

while as I gathered my otherwise-busy perceptions, and finally said, with a distinct question mark in my voice, "Robyn?"

And Robyn, for it was she, turned to a friend next to her, held out her arms helplessly, and said, "See! He finally recognized me. How many minutes did it take?"

But other things go unrecognized, too; not only by me, but by everyone. It's one of those unrecognized things I'm going to discuss in this essay.

One of the remarkable chemical discoveries in the 1890s was that of a group of gaseous elements whose existence had, until then, been entirely unsuspected. They were relatively rare, existed in the atmosphere in percentages that varied from small to tiny, and were most notable for being almost totally inert. They existed as single atoms that did not combine among themselves, or with others (with a few exceptions first noted in the 1960s).

As a group, these were called the inert gases, though in the last quarter century the phrase noble gases has come into fashion.

Between 1895 and 1898, five of these gases were discovered in the following order: argon, helium, neon, krypton, and xenon. The names are derived, respectively, from the Greek words for "inert," "sun," "new," "hidden," and "stranger."

It is conventional in chemistry to give nonmetallic elements names that end in on (as boron), en (as oxygen), or ine (as chlorine). Exceptions are those elements named before the convention was established, as in the case of sulfur and phosphorus.

Metallic elements have names that end in urn (as aluminum) or iwn (as sodium). Again, there are exceptions

133

for those named preconventionally, as gold, copper, and

lead.

All the elements named after 1800 adhere to the convention, except for helium.

Helium is a nonmetal; in fact, it is the most pronouncedly nonmetallic of all the elements. The trouble is, though, that the first indication of its existence came indirectly, back in 1868, through some lines in the spectrum of the Sun's corona. Since nothing could then be told about its chemical nature, and since a majority of the elements were metallic, it seemed safe to call it helium.

Once the element was actually located on Earth, studied chemically, and its nonmetallic nature understood, it should have been renamed hehon, but it wasn't. I presume the chemists who made the Earthbound finding felt it was important to preserve the remarkable priority of its discovery in the Sun, and not mask that by changing the name.

At the time the five inert gases were discovered, the concept of atomic number had not yet been worked out (see "The Nobel Prize That Wasn't," F & SF, April 1970).* This was a pity, for had it been worked out, chemists would have known at once that a sixth noble gas had to exist.

The atomic numbers of the five noble gases are: helium 2, neon 10, argon 18, krypton 36, and xenon 54. The five numbers are then 2, 10, 18, 36, 54.

Suppose we imagine ourselves starting with 0, and working out the amount by which we must increase each atomic number to get the next one. We must increase 0 by 2 to get 2; then increase that by 8 to get 10; that by another 8 to get 18; that by 18 to get 36, and that by 18 again to get 54.

* See my book The Stars in Their Courses (Doubleday, 1971)

134

If we list these new numbers, they are 2, 8, 8, 18, 18. Perhaps you see that these numbers are the series of square numbers multiplied by 2. This 2 is I2 x 2, 8 is 22x2, and 18 is 32x2. Following this system you can then add two numbers that are each twice 42, then two numbers that are each twice 52, and so on. This would give you a number series like this: 2, 8, 8, 18, 18, 32, 32, 50, 50, 72, 72, and so on.

If you start from 0 and add these numbers in succession, you get 2,10,18, 36, 54, 86,118,168, 218, 290, 362, and so on. This would give you a series of atomic numbers for an infinite number of inert gases. In the 1890s, the element with the highest-known

-   atomic weight was uranium, and its atomic number turned out to be 92. Even as of today, nearly a century later, we have driven the atomic number up to only a shaky 106. There is therefore no use considering the atomic numbers of 118 and beyond.

What about atomic number 86, however? That falls well within the realm of possibility, since the fairly common metals thorium and uranium have atomic numbers

,   of 90 and 92 respectively. However, in the 1890s, no

"   element of atomic number 86 was known and, without the concept of atomic number to guide them, scientists

'   didn't even know that such an element ought to be searched for.

>   So let's change the subject slightly.

A The noble gases would have been the chemical find of 3" the decade, had it not been that, in that very same de-i cade, radioactivity was discovered.

*   The noble gases were new elements that fit into the

-, already established periodic table of elements neatly. |*r Unexpected though they were, they merely served to I round out the advances of the 1870s.

!   135

Radioactivity, however, did not just add on to what was known. It was a revolutionary finding that led to a remarkable change in our conceptions of what the basic constituents of matter were.

Radioactivity, however, was not an easy thing to untangle.

The original discovery, in 1896, was that the otherwise unremarkable metallic element uranium gave off strange radiations. In 1898, it was discovered that thorium did the same. But what was it that happened to uranium and thorium after they had given off those radiations?

We now know that radioactivity is a phenomenon that changes

uranium and thorium into other elements (some of them hitherto unknown) that are also radioactive, and that change into still other elements, until finally nonradioactive elements are formed.

Realizing that this was what was happening and demonstrating it were, however, difficult indeed. The daughter elements that were formed appeared in excessively small quantities and could be isolated and studied only after heroic endeavors.

If only some of these daughter elements would isolate themselves and make themselves obligingly and easily available for study, the nature of the radioactive series might be understood at once and be placed beyond argument—an argument that might otherwise consume scientific thought and effort for years, or even decades. It would not have seemed, offhand, that such an obliging event could possibly take place, but consider . . .

Uranium has an atomic number of 92, and thorium one of 90. Both of them decay to lead, which has an atomic number of 82. In passing from 90 or 92 to 82, the chances are almost certain that the radioactive series

136

will have a member at atomic number 86—which would be a noble gas.

To us, in the brilliant light of hindsight, that is plain, but to the experimenters of the late 1890s, who did not know of atomic numbers, nothing of the son would occur to them.

Just the same, in 1899, Marie Curie (1867-1934) and her husband, Pierre (1859-1906), noticed that substances that happened to be near a radium preparation themselves began to show signs of radioactivity, even when they were then carried away from the radium. This induced radioactivity might be the result of the impingement of radiations upon the substance. Or else, some radioactive material might somehow have traveled from the radium to the substance and stuck there.

In that same year, an American physicist, Robert Bowie Owens (1870-1940), noticed that there were changes in the radioactivity of thorium if currents of air impinged upon it. The current of air couldn't very well blow the radioactive radiations about, since those radiations were moving too quickly and energetically to be affected. However, if there were such a thing as a radioactive gas, that might be blown about.

Owens happened to be working in the laboratory of Ernest Rutherford (1871-1937) in Montreal, and Rutherford took over the problem. By 1900, he had demonstrated that a radioactive gas was indeed formed in the course of thorium radioactivity. He called it thorium emanation.

That same year, a German physicist, Friedrich Ernst Dorn (1848-1916), showed that radium also produced such a gas, radium emanation. It was this gas that must have produced the induced radioactivity noted by the Curies.

In 1903, a French chemist, Andre Louis Debierne (1874-1949), who had discovered the element actinium

137

(atomic number 89), found that it, too, produced a radioactive gas, actinium emanation.

It became clear as these gases were studied that they were inert and must be related to the argon family. At first, it was assumed that they were three different radioactive gases, since each broke down at a different rate. There was a tendency, therefore, to call them tho~ ron, radon, and actinon, after the parent substances.

However, once atomic numbers were understood, it became clear that all three gases had the same atomic number, 86 (the one that would have been predictable if atomic numbers had been known twenty years earlier). By then, furthermore, it was understood that an element with a given atomic number might exist in several varieties called isotopes. There was therefore a tendency to consider the three gases as isotopes of a single element that might be called emanon from "emanation." The name niton was also suggested, from the Latin word meaning "to shine," because a sample of the gas in a glass container made the glass fluoresce through the radioactive radiations.

Of the three isotopes, radon has a nucleus made up of 86 protons and 136 neutrons. The total number of nuclear particles is 222, so it might be called radon-222. It has a half-life of 3.823 days.

Thoron has a nucleus made up of 86 protons (the number of protons in various isotopes of a given element is always the same) and 134 neutrons, so it is tho-ron-220. It has a half-life of 52 seconds.

Actinon has a nucleus made up of 86 protons and 133 neutrons, so it

is actinon-219. It has a half-life of 3.92

seconds.

These are the three isotopes that occur naturally in tiny traces (since they break down so rapidly). There are many other isotopes that have been formed in the labo-

138

ratory, but none have a half-life of more than 15 hours, and none occur naturally.

Radon, then, which has the longest half-life by far, outweighs all other isotopes of the element and, in 1923, it was decided to make radon the official name of the element, so that the three naturally occurring isotopes are radon-222, radon-220, and radon-219. When I speak of radon in the remainder of the article, however, I mean radon-222.

Radon fits in very well with the noble gases, for its radioactivity does not interfere with its ordinary properties. Thus, the boiling point of the noble gases goes up steadily with atomic number. The most massive of the stable noble gases, xenon, has a boiling point of 166.0 K (-107.1 C) and that of radon is 211.3 K (-61.8 C).

(If it were conceivable that we were to manufacture an element with atomic number 168, it would be a noble gas that was liquid at more or less ordinary temperatures).

Radon occurs naturally because it is constantly being

produced by uranium atoms breaking down in the soil.

Wherever uranium exists, and it is very widespread in

small quantities, radon is produced. Solid isotopes pro-

:" duced by uranium breakdown stay with the uranium, of

i course, but radon percolates up through the soil and

"' into the atmosphere.

>:   How much radon is to be found in a particular por-| tion of the atmosphere depends on how much uranium

•i there is in the local soil, how porous the ground hap-I pens to be,

whether the ground is wet or dry, how high

•' above the ground the measurement is taken, how much

fuel is burned in the locality, and so on. \, Over the oceans, far away from uranium deposits, the . quantity of radon in the air may be as little as 64 bil-

139

lionths of a gram in a cubic mile. (This is the result of scribbled calculations on my part, and I don't swear to the absolute accuracy—correct me if I'm off.) Over cities it may be as high as 20 millionths of a gram per cubic mile. In the atmosphere as a whole, I calculate there may be 100 grams altogether, or less than 4 ounces.

This tiny quantity may have its uses. We all know that rain is essential to life, but it isn't that easy to get raindrops started. A nucleus is required around which the molecules of water can gather and increase in number until the whole is heavy enough to fall. Dust particles are useful in this process, and there are some who think the most effective are those that result from the constant bombardment of our Earth by uncounted numbers of micrometeorites. In other words, the fact that space about us is dusty helps support life.

However, it is also possible that radioactive radiations produce ions in the atmosphere by knocking electrons off atoms, and that these ions act as nuclei. Thus, the constant dribble of radiations from radon in the atmosphere may contribute to rainfall as well.

If a tiny quantity of radon is mixed with beryllium powder, the radon radiations knock neutrons out of the beryllium, and you have a steady source of such neutrons that will last for days. This can be used in cancer

therapy.

Radon can be detected with great delicacy, so that by putting a tiny quantity of radon into the air or into the ground here and then testing for it there, it is possible to measure wind action or underground water transport,

and so on.

An even more exotic use is this. Any change in the porosity of the soil will introduce a sudden rise (or fall) in radon over some particular site. There are tiny changes in geologic faults prior to an earthquake

that could affect porosity and thus be reflected in a sudden

change in radon concentration. If radon helps us to detect a soon-to-come major earthquake with sufficient certainty, and in sufficient time, to allow for evacuations and other safeguards, that would be a blessing indeed. However, radon also has its dangers, dangers that went unrecognized until just a couple of years ago.

Everything all about us has its traces of radioactive

substances—not only uranium and thorium, but rare

isotopes of potassium, rubidium, and so on. For the

"* most part the radioactive substances stay where they

are, and it is only the radiations that strike us.

One exception is carbon-14, which is found in trace amounts in the carbon dioxide in the atmosphere. It is absorbed by plants, incorporated into plant tissue, and from there, it finds its way into animal tissue, including our own. It can do damage there (see "The Enemy Within," F & SF, September 1986).f

Another exception is, of course, radon, which manages to percolate into the atmosphere.

Now radioactivity has existed as long as the Earth has. In fact, since the formation of the Earth, that major part of radioactivity that originates with uranium has declined to merely half of what it was at the start.

In any case, life has lived with radioactivity and has airvived and flourished. Indeed, it can be argued that radioactive radiations are one of the factors that bring aboul mutations, and that they therefore serve as part of ^ttte engine that drives evolution and has produced us from the original primitive bacterial cell. >t Since we live with carbon-14 and the damage it inevitably does to us, it might seem that we could certainly

five with radon. To begin with, there is only 1 atom of

&

j t See my book The Relativity of Wrong (Doubleday, 1988).

141

radon for every 200,000 atoms of carbon-14 in the atmo-sphere (my own calculation).

This means that though there are always a number of carbon-14 atoms in the airspace of our lungs, bombarding the delicate lung tissue with radiations that may conceivably do serious damage, there are far fewer radon atoms doing the same thing. Furthermore, carbon-14 can enter the body and be incorporated into our very genes. Radon, however, is an inert gas. It goes into the lungs and out of the lungs and, it would seem at first glance, that's all.

Why, then, worry about radon? There are three reasons:

1. Radon breaks down far faster than carbon-14. The former has a half-life of 3.823 days, while the latter has a half-life of 5,568 years. Radon atoms are, therefore, much more likely to produce radiation in a given short period of time than carbon-14 atoms are. Indeed, even though there are 200,000 carbon-14 atoms for every radon atom in our lungs, the radon atoms produce twice as much radiation per unit time as the carbon-14 atoms do.

2. Carbon-14 produces light particles of beta radiation. Radon produces the much more massive and harmful alpha radiation.

3. When carbon-14 breaks down, it changes into harmless, stable nitrogen atoms. Radon, however, breaks down into other radioactive atoms, including several that produce alpha particles—such as polonium-218, astatine-218, polonium-214, and polonium-210. The first three have very short half-lives and produce particularly energetic and dangerous alpha particles What's more, they, unlike radon itself, are not inert but are quite likely to combine with atoms in the lung tissue and remain there till they break down.

142

The result is that radon is far more likely to cause lung cancer than carbon-14.

f And things are even worse than they seem so far. Although human technology has not created radioactivity on Earth, it has tended to concentrate it in spots.

Here and there, human beings have engaged in the task of processing and concentrating radioactive materials for use in bombs, power

plants, and so on. Inevitably, some of the radioactivity gets into the soil of the region and stays there. The soil is then very likely to be a long-term source of radon in higher-than-normal concentrations. What's more, there is always a danger that such radioactivity may get into the ground water and spread more widely.

In New Jersey, for instance, a large quantity of such contaminated soil has been collected from the yards of homeowners who had no way of knowing that they were living close to danger. Now the problem is where to put that soil. Not surprisingly, no one wants it in his vicinity.

This sort of thing can take place almost anywhere. Houses can be built in areas where the radioactivity level (either through geological or technological processes) is higher than average. Again, houses can be built of brick or concrete that just happens to be drawn from a section of soil in which the radioactivity level is higher than average.

The quantity of radioactivity, either in the soil or in "the building materials, is not likely to be overly danger-

s in itself, as long as it stays in the soil or in the

ilding materials.

However, from the soil or the building materials, ra-leaks into the interior of the house and may build to concentrations higher than would be found out-

le the house in the open air.

This has become an increasingly dangerous possibility

143

in recent years. In older times, houses were poorly fitted, and full of chinks and drafts. In our own energy-conscious times, however, we tend to labor to make our houses and apartments airtight so as to minimize leaks of heat based on increasingly expensive fuel.

Then, too, whereas in summertime, at least, windows used to be thrown open to allow for ventilation that would somewhat ameliorate the summer heat and humidity, the coming of air-conditioning has made it certain that we close our windows tightly to conserve the coolness.

In short, we are making our dwelling places airtight with respect to

the atmosphere, but we don't bother making it airtight with respect to the ground underneath. The result is that radon leaks into the house from the ground and the walls, and then can't get out—so it builds up.

Consequently, a brand-new activity of the average householder is to get his dwelling place tested for radon content. If it tests high, then one might try stopping all leaks in the basement floor and in the foundations, and, at the same time, try opening the windows whenever possible.

(I have recently read a report that denies that ventilation, or the lack of it, affects the concentration of radon within a house, but I find that difficult to believe. However, these are early days for the investigation of radon danger, and I'll await further work.)

In any case, it is suspected that radon is now the number two cause of lung cancer and that it is responsible for anything from 5,000 to 30,000 deaths from that cause each year.

The number one cause of lung cancer deaths, by a wide margin, is, of course, tobacco smoking. I view with a certain sardonic amusement, therefore, the fact that a

surrounding himself and everyone

be sure

^S£^££*&?JF»*^

uwwiujg piace is a little above average.

144

145

10

The Radiation That Wasn't

The late, great science fiction editor John Campbell was fascinated by all sorts of fallacious devices that purported to do something in defiance of the well-understood laws of nature.

One of these devices was the "Hieronymus machine," and the one thing I remember about it was that one stroked a surface while turning a dial. At certain settings of the dial, the surface was supposed to turn sticky.

I was visiting Campbell back in the 1950s, and he trotted out his Hieronymus machine and, since I was a notorious skeptic, he insisted I try it and see for myself that it worked. I desperately didn't want to, but I was submitting a novel to him and I wanted him to take it and pay me several thousand dollars, so I didn't want to

offend him. I had to go through the motions, therefore. He turned

146

the dial and I stroked. I earnestly tried to feel stickiness but there was none—absolutely none. In fact, as the palms of my hand grew moist with perspiration, because of my discomfort and nervousness, the surface began to feel not sticky, but slippery.

At that point, Campbell said, "Well, Isaac, did you detect a change just then?"

To which I replied in hangdog fashion, "It just turned slippery, Mr. Campbell."

"Aha," said Campbell with deep satisfaction, "negative stickiness."

He insisted that proved the worth of the machine. When I tried, rather diffidently, to bring up the matter of perspiration, Campbell dismissed it as irrelevant.*

Now Campbell was a formidably intelligent man, so

what made him act so foolishly? The only answer I

'•could think of was that the drive to believe what one

wants to believe can be so overpowering that it beats

down everything else.

1 This can happen in serious science, too, and can pro-

fiduce annoying snarls. Nor need there be any question of

funny stuff. Mistakes can be made by capable and ut-

•terly honest scientists entirely because of the Campbel-

•-iesque will to believe.

J- In the late 1960s, for instance, an American physicist, Jfoseph Weber (b. 1919) reported the detection of gravi-Jational waves. These waves must exist according to the general theory of relativity and there seemed nothing rong with the claim from the theoretical standpoint. Nor did there seem to be anything wrong from the serimental standpoint. Weber was an able physicist

* He bought my novel, by the way, though I don't suppose there was Ty direct connection between that and my having agreed to try the marine.

147

and had set up very careful and elaborate devices to detect those waves.

The trouble was that no other physicist was able to detect the waves, no matter how carefully he duplicated Weber's work, and without such confirmation by others, a finding doesn't count.

But why was it that Weber could see what others couldn't? Again, the only explanation would seem to be that a scientist is human and, if he is very eager to make an important finding, and if that finding involves an observation that is just at the barest edge of sense perception, that scientist is liable, in all honesty, to see what he desperately wants to see.

The error was not, in this case, a dreadful one. Weber was on the right track, but his technique was not quite sensitive enough for the task. Physicists are busily engaged hi devising more sensitive techniques and someday, they are confident, gravitational waves will be indisputably detected.

Then there was the case of the American astronomer Percival Lowell (1855-1916), who saw canals on Mars through his telescope. He saw them in considerable detail, made careful drawings of his observations, wrote books on them, and placed the canals very firmly into half a century of science fiction stories. . . . But, just the same, the canals did not exist.

Lowell was an honest man, and a careful, hardworking astronomer, but he was trying to see things on Mars that were on the very edge of what could be seen. He was victimized partly by optical illusion and partly by the ardent desire to see what he thoroughly believed to be tfyere.

Sometimes an observation falls apart at once, but not soon enough to keep me from committing myself to it in an essay and then finding myself forced to backtrack later.

148

Thus, in the March 1986 issue of F & SF, I published "Superstar,"t in which I discussed stars that were far more massive than it had been thought, till then, that stars could possibly be. Alas, even before the article appeared, astronomers (confound them) had changed their minds and decided that superstars did not exist.

Earlier, in the June 1985 issue of F & SF, I had published "The Rule of Numerous Small."$ Almost all that I said in that essay is correct, but I had been inspired to write it by the discovery of what was called a brown dwarf, an object too small to shine by ordinary nuclear fusion but large enough to shine dimly by other processes. Naturally, it was assumed that there must be many brown dwarfs in the Universe.

However, the brown dwarf simply disappeared. Attempts to detect it where it had earlier been detected failed completely. What's more, a search for other possible brown dwarfs turned up none at all.

I repeat that I am not talking about fraudulent work by scientists who, for one reason or another, are lured out of the paths of rectitude. Such scientists exist but their crimes are not interesting; merely shameful.

I am talking about honest and skillful scientists, doing honest and skillful work, whose own all-too-human eagerness to find, and unwillingness to let go, lead them into error, embarrassment, and, sometimes, into destroyed careers.

In this connection, there may be nothing sadder than the case of a French physicist named Rene Prosper Blondlot (1849-1930). Blondlot was born, lived, and died in Nancy, a provincial French town 280 kilometers (175 miles) due east of Paris. He was the son of a well-

t See my book Far as Human Eye Could See (Doubleday, 1987) i Ibid.

149

known chemist and he himself taught physics at the local university.

He would probably have done much better if he had located himself in Paris, but he apparently loved the town of Nancy and made no attempt to leave it. Even so, he didn't do badly. He was a topflight

experimentalist and for his work he won three prestigious prizes from the Paris Academy of Sciences.

In 1875, for instance, a Scottish physicist, John Kerr (1824-1907), had shown that glass and other substances could be made to exhibit double refraction in an intense electric field. This was called the Kerr effect. Blondlot set up a very ingenious and delicate experimental procedure that would measure the time it took for the double refraction to appear after the intense electric field had come into being. He showed that it appeared in less than i/40,000 of a second,

He used a similar experimental technique to check the speed of the electrical impulse. By Maxwell's equations, it made sense to suppose that the electrical impulse traveled at the speed of light, but it always helps to make an actual measurement. Blondlot sent simultaneous electrical charges through two wires, one of which was 1,800 meters (1.11 miles) longer than the other, and was able to show that the speed of propagation of an electrical impulse was very close to the speed of light. In other words, Blondlot was a very good scientist. But then, in 1895, the German physicist Wilhelm Konrad Roentgen (1845-1923) discovered X rays (see "X Stands for Unknown" in the August 1982 F & SF)* This initiated a rapid-fire series of discoveries that totally revolutionized physics and, in 1901, when the Nobel Prizes were set up, Roentgen got the first Nobel Prize in physics.

* See my book X Stands for Unknown (Doubleday, 1984)

150

The world of physics was dazzled by the prospect of new and hitherto unknown forms of radiation that offered a highway to scientific fame. It was not just X rays. That had been preceded by the discovery of radio waves and cathode rays, and it was to be rapidly followed by the discovery of alpha rays, beta rays, and gamma rays.

Almost every physicist in the world turned toward the study of these new radiations, but it seems to me that Blondlot had a special drive pushing him forward. This is only speculation on my part, but consider . . .

The town of Nancy, which it would seem Blondlot strongly loved, was the capital of Lorraine, which for a period of nearly a thousand years was an independent or semi-independent duchy. It was French in language and culture, but it did not become an integral part of the French kingdom till 1766.

In 1870, however, when Blondlot was twenty-one, France was badly

beaten by Prussia in the Franco-Prussian war. Prussia combined with other German-speaking regions to form the German Empire, which at once became the strongest power in continental Europe. As part of the spoils of war, the German Empire forced ' France to cede to it a portion of its eastern territories called Alsace-Lorraine.

Alsace was, indeed, to some extent a German-speaking province, but Lorraine was entirely French. To be sure, Germany took only eastern Lorraine, including the important city of Metz, and left western Lorraine, including Nancy, to France. Nevertheless, Nancy was now only sixteen kilometers (ten miles) from the new German border.

For some forty years or more, France refused to be reconciled to the loss of the provinces. It viewed Ger-many with intense hostility and waited only for revenge (which it finally got in World War I at far too high a price). Surely, the feeling must have run particularly

151

high in Nancy and Blondlot could not have been immune to it. He must have wanted, with all his heart, to match the work of the German scientist Roentgen, and, if possible, to surpass him.

After the discovery of X rays, the immediate'contro-versy was over the nature of the new radiation. Were X rays a wave form, or were they a stream of speeding particles? Either alternative might have been correct. Radio waves and visible light were clearly waves, while cathode rays, alpha rays, and beta rays were streams of speeding particles.

All the particle streams then known consisted of particles that carried an electric charge, and these could be deflected if they passed through an electric field. If, however, the particles were moving very quickly, the deflection might be unnoticeably small.

Blondlot decided to tackle it from the other end. If X rays were waves, they could be polarized when passing through an electric field and made to wave in one particular plane. This would be a phenomenon not shown by particles. The Kerr effect had involved polarization, so Blondlot felt he was thoroughly expert in this field, Blondlot used a detector made of two sharply pointed wires with an electric spark leaping across the gap between them. He reasoned that if the spark were hi the plane of polarization it would be more energetic and would brighten. If it brightened when it was placed in one direction but not in others then the X rays were polarized and would be proved to be waves.

He tried the experiment and it seemed to him that it worked. The spark appeared to brighten and he felt that he had proved that X rays were waves (which they indeed are, by the way).

But then came trouble. When Blondlot passed the X rays through a quartz prism, he had to change the orientation of his detector as though the plane of polariza-

152

tion had shifted. However, there seemed no reason to suppose that quartz would affect the plane of polarization of X rays. Something was wrong.

But then, Blondlot reasoned this way. Something was brightening the spark. If it wasn't X rays, it had to be some other form of radiation that perhaps accompanied the X rays but was completely different in nature.

Blondlot was taking a terrific chance here. When Roentgen discovered X rays, he detected them by the fact that they made a certain chemical fluoresce brightly. The difference between the absence and presence of X rays was a difference between total darkness and a bright fluorescence. There was no chance of mistake.

What Blondlot was detecting, on the other hand, was a tiny further brightening of an already bright spark, a brightening that was not very noticeable at all. Blondlot did notice that brightening—there is no chance at all that he was faking—but he wanted to notice that bright-

- ening, and it was a case of honestly seeing what he

wanted to see. Once he got the idea he had a new kind

'• of radiation, he wanted to see brightening more than

* ever and so he saw it.

; Unquestionably, Blondlot would ordinarily have been $ enough of a scientist to check the matter over and over ^until he was sure; and to maintain an air of healthy 'skepticism till he considered the evidence certain. Un-^doubtedly, he would have tried to find a method of detection of the new radiation that would be reasonably " certain. However, the excitement of doing what Roentgen had done and of matching a German discovery with |n possibly greater French discovery must have been too Heat for him. He was entirely too eager

to have the new Radiation be real.

To be sure, he did try to rely on something more than

i   153

simply gazing at the spark and deciding that it had or had not grown brighter.

Thus, he had the sharply pointed wires and the spark that flashed from one to the other enclosed in a cardboard box to keep out ordinary light. Beneath the spark was a piece of ground glass that diffused the light. Under the ground glass was a photographic plate that recorded the fuzzy light of the diffused spark. An alternative was to place a fluorescent chemical underneath the ground glass.

This gave the illusion of removing the subjective nature of the determination, but that was only an illusion. The photographic plates and the fluorescent substance would, indeed, show a brightening that would not be influenced by subjective desire, but what brightening there was was still very small. In the end, the human eye had to be called upon to decide whether one photograph showed a brighter fuzzy spot than another, or whether a fluorescent material glowed more brightly at one time or another. And the human eye might easily be victimized by a human brain that knew the answer it wanted and insisted on having it.

In 1903, Biondlot could wait no longer. He announced his discovery to the world. He had discovered a new radiation totally different from anything of the sort that was already known and that might therefore open a new frontier of physics. He called the radiation N rays and the N, as you can guess, stood for "Nancy."

At once, other scientists, particularly Frenchmen who rejoiced in this French discovery, jumped on the bandwagon. They all began to set up detecting devices to see small bits of brightening under this circumstance and that, and to determine new facts about N rays.

For instance, what were the sources of the new radiation? Biondlot had first detected them in connection

154

with his work on X rays, which were produced by cathode-ray tubes, so that was an obvious source.

Heated metals and certain oxides, when heated, emitted them, since

the spark was reported to have brightened when exposed to these substances. The Sun emitted N rays, Biondlot reported. Others found that the human body was a source, whether it was living or dead, H and that individual protein molecules were a source, too.

Almost everything was transparent to N rays. To put it : another way, N rays could pass through almost every-1 thing. About the only substances that were opaque to N

rays were water and rock salt. However, even when N rays passed through certain

substances, they might still be affected in some ways.

Just as glass refracts light rays, so do substances such as

. aluminum refract N rays (it was reported). Biondlot de-

' vised lenses and prisms made out of aluminum. These

> would act to concentrate N rays and make their effects

I more noticeable.

s   All this made such a splash that, in 1904, Biondlot r received a prize of fifty thousand francs for his work. It ; was for all his work, to be sure, and not for his discovery £ of N rays specifically, although that was mentioned. %

I   Of course, there were dissenting voices, particularly I outside France, where there were no patriotic reasons to support Blondlot's views. Physicists in Germany, Great .Britain, and the United States repeated Blondlot's ex-jtperiments as closely as possible and reported being un-f able to detect any sign of N rays. Two of those who ^couldn't were two topflight British physicists, Lord iKelvin (1824-1907) and William Crookes (1832-1919). Such disagreements did not dismay Biondlot and his sllow enthusiasts. Disagreements were, after all, to be

155

expected, and they were easily explained by assuming that those who disagreed were doing the experiments improperly or were using inferior equipment,

(Responses of this sort were usual. When Lowell was busy mapping the canals of Mars, there were oiher astronomers who reported never

being able to see the things. Loweli's confident response was that he had a better telescope and better viewing conditions.)

There was, however, an American physicist, Robert William Wood (1868-1955), who was a professor of physics at Johns Hopkins University and who specialized in optical work. He was interested in the new radiations, particularly in the mysterious N rays. Eagerly, he tried to repeat Blondlot's work and failed totally. He got nothing and was both chagrined and disappointed.

Wood, feeling that he might have done something wrong, traveled to Nancy in 1904 (a far more onerous trip in those days than it would be now), in order to witness experiments as conducted by Blondlot himself. Blondlot was delighted to see him, was unreservedly cooperative, and willingly ran a whole series of experiments for the American's benefit.

For one thing, Blondlot said, if Wood were to place his hand in the path of the N rays between the source and the spark, some of the N rays would be stopped or scattered by his hand, and the spark would grow dimmer. (There was never any worry in those days about possible dangerous physiological effects of energetic radiation. People had to learn the hard way. Marie Curie herself died of radiation-induced leukemia.)

Wood placed his hand in the path of the N rays and Blondlot and his group immediately pointed out that the spark had grown dimmer. Wood could see absolutely no change, however, and said so. He was told that his eyes were insufficiently sensitive. Wood then suggested that his hand be hidden and

156

fr that he move it in and out of the path at irregular intervals. The N-ray group could then tell him when the radiation was blocked and when it was unblocked by saying when the spark dimmed and when it brightened. Wood then put his hand into and out of the path, and the group kept calling out, "Dimmer!" and "Brighter!" At no time, however, did the calls coincide correctly with the position of Wood's hand.

The Blondlot group then showed experiments in which the light was photographed with and without a piece of wet cardboard blocking the N rays. Since water was opaque to N rays, the photographs should be dimmer when the wet cardboard was in the way. When the wet cardboard half-blocked the light, one side of the photograph should be brighter than the other.

Wood remained skeptical, however, for it seemed to him that, in a number of different ways, there was room for error and that the results were far from conclusive.

Then Blondlot performed a particularly complicated

* experiment. He had the N rays fall on an aluminum

, prism that spread them out so that they fell on a strip of > phosphorescent paint in four different places—as evidenced by the fact that, according to Blondlot, four spots on the strip were particularly bright. The conclusion was that the N rays had been divided into four

^separate wavelengths.

^- Wood, however, could not for the life of him make out any sign of brighter areas on the phosphorescent

- strip.

Jf,   So Wood decided to do something drastic. The exper-iment had to be conducted in a darkened room so that £the phosphorescence would stand out better. In the ^darkness, then, Wood abstracted and pocketed the alu-'ium prism. He then asked that the experiment be •eated. Since it was the aluminum prism that refracted and

157

separated the N rays into four different wavelengths, the absence of the prism ought to destroy the results of the experiment completely. Nevertheless, when the experiment was repeated without the prism, the Blondlot group reported the same four areas of brightness.

In another experiment, a large steel file was used as a source of N rays. Wood managed to abstract that and substitute a similar piece of wood which was not supposed to be a source of N rays. Nevertheless, the experiment was reported to have worked perfectly.

Wood might have suspected fakery, but the obvious willingness of the Blondlot group to cooperate, their almost naive enthusiasm, and the very borderline nature of the observations, made it seem clear to him that it was all a matter of self-delusion.

Wood reported everything he had observed and done and the whole business of N rays was dropped at once— except in France.

Blondlot clung to N rays. He tried to reply to Wood's criticisms. He devised new and better automatic procedures for measuring the level of light. He called on the support of other (French) scientists. For a while, the nationalistic tone became ugly. It came down, according to some French enthusiasts, to a matter of sensitivity. Anglo-Saxon and German eyesight was simply not as delicate and refined as French eyesight was.

But finally even French scientists turned against the hard core of N-ray enthusiasts at Nancy. In 1906, a French team of scientists devised an experiment. They prepared two wooden boxes of equal size, weight, shape, and appearance. One contained a piece of tempered steel that was supposed to be an N-ray source that would pass through the wood, and the other a piece of lead that was not an N-ray source. The boxes were sealed and secretly identified. Blondlot was challenged to test, publicly, the two

158

boxes for N rays in any way he chose and to tell which one had the steel in it. Blondlot hesitated, and then refused to subject himself to the test. With that, the whole matter of N rays died. It had been alive for three years.

Blondlot's scientific career was at an end. He lived out the remaining quarter century of his life in obscurity. Perhaps he gained some satisfaction in living to witness the end of World War I and the return of Alsace-Lorraine to France. Since he died in 1930, at the age of eighty-one, he was spared the disaster of 1940, when France was totally defeated by a resurgent Germany and lost Alsace-Lorraine a second time (but for only five years).

What are the lessons from all this? First, scientists are human and can be driven by hopes and desires into error and folly.

Second, science is and should be international. The intrusion of patriotism and ideology can only be mis-'chievous. Just as French patriotism powered the N-ray < affair, at least in part, so did English patriotism power , the Piltdown hoax. Again, Soviet ideology made Lysenko possible, while Anglo-Saxon ideology made yril Burt possible.

Third and most important of all, we see that science a strong tendency to be self-correcting. Confirma-Ition of all findings are required and is not easy to come y. Without confirmation, findings are thrown out. What's more, if there is the faintest ground for sus-:ting a hoax, or

incompetence, or even mere folly, ientific reputations and careers can be punctured or istroyed. There is no forgiveness for deliberate falsity, id very little forgiveness for foolishness. Compare this with almost any other realm of human

159

endeavor. We have all seen, in recent years, how figures in government, in industry, in finance, even in religion can commit stupidities and even outright crimes, and admit to them, and be made heroes as a result.

This does not happen in science. In science "(and, I believe, in science alone) one cannot make up for stupidity and incompetence by cultivating a charming smile and a carefree wave of the hand.

Part IV Magnetism

160

] 1

Iron, Cold Iron

JL

A couple of weeks ago, I was standing in the hall at 1 Doubleday, waiting for an elevator. I had an advance copy of my new novel, Fantastic Voyage II: Destination f Brain, in my briefcase.

A young man, new at Doubleday, came rushing into the hall and said, "Pardon me, are you Isaac Asimov?"

I said, "That was who I was this morning. I guess I 'still am."

He said, "I knew you were a Doubleday author, but I lidn't think I'd ever get to see you."

I said, "I hope you're not disappointed now that you jfhave. My books are better than I am."

He said (almost inevitably, for few can resist), "How lany books have you published now?"

I thought of the fresh-minted novel in my briefcase

id said, with considerable satisfaction, "Three hun-

dred sixty-five." At this, I paused, and during the pause a gentleman entered the hall who, as it quickly turned out, did not recognize me at sight, or, possibly, had never even heard of me. I paid no attention to him, but, having paused, I then added something to my earlier remark to the young man.

I said, "I've published one book for every day in the year."

At this, the gentleman who had just stepped into the hall smiled in a most friendly fashion at me, patted my shoulder consolingly, and said, "I'm sure there must be days every once in a while when it seems like that," and went his way cheerfully.

The young man said softly, "What does he mean, 'seems'?"

But I just laughed and said, "It's all right. Three hundred sixty-five doesn't sound believable even to me."

In fact, this essay is the three hundred and fifty-fourth I have written for F & SFt which means that in eleven months (always assuming no catastrophe intervenes) I will have reached the mark of one-for-each-day-in-the-year for this series, and that, too, won't sound particularly believable—even for me.

But I intend to shoot for it (and beyond) just the same, so here goes . . .

Iron was one of the metals known to the ancients, but in some ways, it doesn't measure up. Gold, silver, and copper are, each in its way, beautiful, but iron is a gray and ugly metal.

Gold does not rust, but retains its beauty indefinitely. Silver is almost as good, and copper isn't entirely bad. Besides, even if silver and copper tarnish and discolor, they are easily polished back to the original shine. Iron, however, rusts much more readily than the other metals

do, and the rust is not only an ugly, brick red in color but it crumbles as it forms. Iron would seem to have no esthetic qualities at all.

Yet surface beauty isn't all there is. As long as iron can be kept from rusting, it is, or can be made, harder and tougher than any other

metal known to the ancients. It can hold a sharper edge and it is much more difficult to blunt.

Gold, silver, and copper are far too soft to use for long-lasting tools, for tough weapons of war, for protective armor. Copper can be hardened by alloying it with tin to form bronze, and, in the early days of warfare, soldiers fought with bronze swords, bronze-tipped spears, bronze-layered shields, and so on. Homer's Iliad is the great literary production that describes warfare in the Bronze Age.

An iron sword, however, can hew through a bronze shield, and an iron shield will blunt and bend a bronze sword. A properly iron-equipped arm can easily destroy one that is merely bronzed.

Or, as Rudyard Kipling said, in a poem he wrote in 1910:

Gold is for the mistress—silver for the maid— Copper for the craftsman cunning at his trade. "Good!" said the Baron, sitting in his hall,

"But Iron—Cold Iron—is master of them all."

Of course, metals were rare and hard to find (the very word metal is from a Greek word meaning "to search for").

Yet, perhaps as long ago as 5000 B.C., it was discovered that when certain blue rocks were heated in a wood fire, beads of copper appeared. The discovery was made accidentally at first, I'm sure, but it eventually led to the deliberate search for metal ores and for the de-

165

velopment of metallurgical techniques by about 3500

B.C.

The metallurgical techniques first developed were insufficient to squeeze iron out of its ores so that the only iron available in the first two thousand years of metallurgy was that which was to be found already in metallic form.

Earth's supply of iron never appears in metallic form, but, fortunately, there is iron in the sky. At intervals an iron meteorite would strike the Earth, and the iron so brought down was actually a nine-to-one mixture of iron and nickel, and this alloy was harder, tougher, and more rust-resistant than iron itself. Such meteorites were searched for avidly, so that no iron meteorite from the past is

ever found in places where the earliest civilizations flourished. The ancients had scavenged them all. Isolated cases of iron smelting may have taken place as early as 3000 B.C., but the technique was not developed in a systematic way until 1500 B.C., when the Hit-tites in Asia Minor learned how to make use of charcoal fires to get the temperature high enough for the purpose.

The Hittites undoubtedly kept their secret for some centuries, for much the same reason that we tried to keep the nuclear bomb a secret. It was easier to keep secrets in ancient times, and the Hittites retained a monopoly on iron until 1200 B.C., when their empire was finally destroyed. Even the Hittites formed iron in only small quantities and could not field a completely iron-equipped army. Eventually the pressure of outside enemies became too much for them.

The Hittite iron workers spread out and practiced

their skill elsewhere, teaching it to others, and iron

weapons became more common and widespread—but

still not universal.

When the Israelite tribes entered Canaan about 1200

166

B.C., they were uncivilized nomads who lacked the ability to form their own iron. They were amazed and daunted by the fact that the more civilized, town-dwelling Ca-naanites did have some iron. For instance: ". . . Og king of Bashan remained of the remnants of giants; behold, his bedstead was a bedstead of iron . . ." (Deuteronomy 3:11).

It was because of this that when the Israelites first entered Canaan, they spoke of the inhabitants as "giants." Later generations accepted the term literally, but it makes much more sense to suppose that the Israelites were awed by the Canaanite's iron technology. The Ca-naanites were giants in that sense.

Thus, the Israelites complained to Joshua that "all the Canaanites that dwell in the land of the valley have chariots of iron" (Joshua 17:16). And when the Israelites fought Sisera in northern Canaan, "Sisera gathered together all his chariots, even nine hundred chariots of iron . . ." (Judges 4:13).

Of course, the Israelites patriotically describe themselves, under the

leadership of Joshua, as victorious over the Canaanites, but this can be doubted. For at least two centuries after their appearance in Canaan, the Israelites were often under the domination of non-Israelitic groups according to the Bible itself. As late as 1000 B.C., they "served" the Philistines.

The Philistines had cold iron, you see. ". . . There was no smith found throughout all the land of Israel: for the Philistines said, Lest the Hebrews make them swords or spears: But all the Israelites went down to the Philistines, to sharpen every man his [plow] share, and his coulter, and his axe, and his mattock" (1 Samuel 13:19-20).

It was only under King David, soon after 1000 B.C., when presumably the Israelites managed to iron-equip

167

their army, that the Philistines were defeated and the Israelites became, for a time, a dominating force.

Again, by 1100 B.C., the Bronze Age Greeks who were the descendants of the warriors at Troy were overthrown by another tribe of Greeks from the north—the Dorians —who had iron weapons.

At that same time, the Assyrians were making use of iron weapons, too, and began to establish a large and powerful empire in what is now Iraq. Indeed, by 800 B.C. the Assyrians were the first to iron-equip an army thoroughly, so that for a while they were unbeatable.

Eventually, iron metallurgy was developed to the point where iron and its alloys became the cheapest of metals, so that to this day we use iron and steel when we need strength and affordability. But now I will move on to another type of property which, when it was first discovered, seemed to belong to iron and iron ore exclusively.

The property might well have been noted in very early times, but it was not till about 585 B.C. that observations were recorded and the phenomenon systematically studied.

According to the story, as detailed in the writings of the Roman encyclopedist Pliny (A.D. 23-79), who recorded everything he read or heard, a Greek shepherd who had iron nails in his shoes and an iron ferrule at the bottom of his staff, found that shoes and staff seemed to cling to a certain rock he encountered. It was not a universal stickiness, for nothing that was not iron stuck to the rock. The shepherd is supposed to have lived near the Greek city of Magnesia, located in what is now the Aegean coast of Turkey.

Samples of this sticky rock found their way to the most noted Greek scholar of the time, Thales (624-546

168

B.C.), who lived in Miletus, which was about ninety miles south of Magnesia.

Thales studied the properties of what he is supposed to have called ho magnetes lithos ("the Magnesian stone") and he found that it did indeed attract iron, but no other material available to him. Ever since, we call such an iron-attracting material, in English, a magnet from Thales' phrase, and the phenomenon is referred to as magnetism.

The particular rock which displayed magnetic properties is a relatively uncommon oxide of iron, which is now called magnetite. In earlier times, it was called loadstone, or fadestone, for reasons I will describe a bit later.

The ancients were fascinated by this mysterious and highly specific attraction. Thales thought it indicated the presence of some kind of life within the magnet, and that the attraction of iron was the result of a sort of affection between the two.

Some noticed that if a bit of iron was attracted to a magnet, that bit of iron, while in contact with the magnet, would attract a second bit of iron, which attracted a third bit, and so on. Plato (427-347 B.C.) has Socrates (470-399 B.C.) refer to this and make an analogy to the way an accomplished teacher can inspire a pupil and imbue him with the enthusiasm that will enable him to inspire a pupil of his own, and so on.

There were also those who noted that, under some circumstances, magnetism exerted a repelling effect.

By far the most important early discovery concerning magnetism, however, was made in China. No one knows how it came about, but here is how I imagine it might have happened . . .

If someone has a sliver of loadstone, it is bound to be fun to play with. One game would be to place it on a piece of wood and float it in a tub of water. It is then free to turn in any direction and, if you have a piece of

169

iron, you can "tease" it and make it turn toward the iron in this direction and that. I dare say that children, particularly, would consider this fun.

Eventually, one would get tired of the game and, perhaps leave the loadstone floating; then later come back and play the game again.

Eventually, some observant person was likely to notice that when the sliver of loadstone was left to itself, it always ended up aligned in a north-south direction. The magnet not only seeks iron, it would seem, but also seeks the north (or the south). There is a reference to this sort of thing in Chinese books dating as far back as

A.D. 121.

The Chinese, however, as far as we know, made no practical use of this property of a magnet. It may have been used in magic shows. There is also the suggestion that when Chinese traders or soldiers made their way across vast stretches of central Asia, they made use of the magnet to give them a notion of direction, but I find that a little hard to believe.

They seem definitely not to have used it at sea. For the most part, the Chinese were not great sea travelers. Self-satisfied to a fault, they felt that they had the only part of Earth worth anything and tended to stay at home.

They did make reference to the use of magnets for finding direction at sea as early as 1086, but the reference, then, was to foreign sailors, presumably from what is now Indonesia.

At just about that time, an English scholar, Alexander Neckham (1157-1217), made the first European reference to the use of magnetism to find direction at sea. How the news spread from China to western Europe, we don't know. It is conceivable, I suppose, that the discovery was made in Europe independently, but we don't know about that, either.

170

Prior to 1200 or so, sea voyagers got their best ideas of direction by observing the Sun at midday, when it was always in a due southerly direction. At night, they observed the North Star, which was always due north. Then, too, the Sun rose in the east and set in the west and that was useful, too. Once you know one direction, you know all the others, too.

The trouble with all this, though, is that many days and nights are cloudy. The Sun and stars are then not seen and direction-finding breaks down. As a result, sailors rarely dared get far out of sight of land, lest they be unable to find their way back and so perish.

But suppose you pivoted a loadstone sliver on a horizontal card so that it was free to turn in any direction around the card. It should eventually come to a halt in the north-south direction, with one end, distinguished by a touch of paint, perhaps, pointing north.

The word "load" is an archaic term for way, or route, or direction of journey. Therefore, anything that revealed the proper direction could be given that word. The North Star was sometimes called the loadstar, and that is why the magnetic rock came to be known as loadstone.

The word compass comes from a Latin term meaning "to measure around a circle." That is why the device geometers use to mark off a circle is called a compass. In the same way, the card with the pivoting needle able to go around in any direction is also called a compass. Tb distinguish the two compasses, the one using the loadstone is a magnetic compass.

The magnetic compass, as used in its early centuries, was crude, but it worked. It made it possible to move away from the coast and venture into the open sea for now one could determine directions, be it ever so cloudy, and there was a sharply reduced fear of getting lost and being unable to return to land.

171

To be sure, a magnetic compass is not an absolute necessity for sea travel. About the time the compass came into use in European vessels, the Polynesians were moving all over the vast Pacific Ocean in open, primitive vessels, with nothing but Sun, stars, currents, and bird flights to help them make their way between the tiny dots of land that were scattered widely over the sea.

Nevertheless, the Polynesian feat was a difficult one that they just barely managed, and that was sure to leave them stranded on particular islands for long periods of time. The west Europeans, with the compass, began, soon after 1400, to move across the seas and to begin an Age of Exploration that, for a period of time, allowed a few small nations—Portugal, Spam, England, France, the Netherlands—to dominate the world. All because of the compass—and gunpowder.

The first person to study magnetism with something like modern thoroughness was a French scholar whom we know only as Petrus

Peregrinus ("Peter, the Pilgrim")- He was born about 1240 and we don't know when he died. He was an engineer hi the army of the French king, Louis IX, and, in 1269, while he was engaged in the dull and long-drawn-out siege of an Italian city, he wrote a letter to a friend in which he described his experiments with magnets.

Peregrinus showed that the magnetic properties of a magnet were concentrated at the ends, or poles. He was the first to call them this, and we still speak of them as such, sometimes specifying them as magnetic poles to differentiate them from geographic poles that come at the two ends of an axis of rotation.

He showed further that it was always the same pole that pointed toward the north, so that one could speak of a north magnetic pole and a south magnetic pole.

172

(He apparently failed to notice, however, that the north magnetic pole of one magnet attracted the south magnetic pole of another, but that two north magnetic poles or two south magnetic poles repelled each other.)

Peregrinus also showed that it was impossible to isolate one of the poles from the other. Both always existed on a given magnet. If a magnetic sliver was broken in two, the half with a north magnetic pole developed a south magnetic pole at its broken end; the half with a south magnetic pole developed a north magnetic pole at its broken end.

He was also the first to study the behavior of iron filings when shaken on a card underneath which a magnet existed. From this he deduced the presence of what we now call a magnetic field.

In addition, he was the first to suggest that a ship's compass not be pivoted on a mere unmarked card but on one on which the various directions were marked. (He also had the erroneous notion that the needle would slowly work its way around the card in twenty-four hours, matching the rotation of the Earth, so that the compass could be used as a clock.)

It is possible to get the exact direction of north without a compass. When the Sun moves about the sky from east to west, it crosses the north-south line when it is at its highest point. It is then (at least when viewed from the northern hemisphere) due south.

This can be followed more easily by observing the shadow of a stick

hammered vertically into the ground. i As the shadow swings about from west to east, it grows from long to short to long again. When the shadow is at its shortest, it is pointing directly north. One can also mark the line of the shadow at sunrise and again at sunset. The angular bisector of the angle thus formed

173

wil! point due north. Then, too, the position of the North Star, if averaged over different times of the night and the year, also gives you the true north.

It is possible, then, to note that the position of north indicated by a magnetic compass often deviates somewhat from the true north. Chinese observers made note of this now and then, even as early as the 700s. However, these were isolated observations and nothing came of them. In Europe, too, there might have been isolated notices of this deviation of the compass from the true north—something called magnetic declination.

Magnetic declination was first studied systematically by Christopher Columbus (1451-1506) on the occasion of his famous voyage of discovery in 1492. Not only did Columbus discover America, but his was the first voyage we know of that yielded important scientific information beyond the mere fact of geographic discoveries. After all, Columbus was more than a dreamer and a brave man; he was a skilled navigator and he had the kind of credentials that must allow him, for his time, the status of "scientist."

Columbus noted that the direction of magnetic north not only deviated from the true north, but that the extent and even direction of the deviation varied as he traveled. The compass slowly turned from pointing a bit west of north to pointing a bit east of north and somewhere in midocean he passed a line where the magnet did, for a time, indicate the true north.

He made careful observations of this but kept it secret. He had a hard job keeping his sailors to the task of sailing ever westward, and if they had found out that the compass wasn't telling them the truth, they would undoubtedly have panicked, mutinied, killed Columbus, and headed back east in a desperate desire to regain land before they were lost forever. And without Colum-

174

bus's firm hand on the controls, they were not likely ever to have made it.

If that had happened, Columbus would have set out from Spain and would simply have disappeared. Who knows, then, when another explorer might have been mad enough to try the same voyage, especially as five years after Columbus's discovery, the Portuguese really reached India by going around Africa.

Compasses are always so pivoted that they can swing only clockwise and counterclockwise in a plane parallel to the surface of the Earth. What if they are pivoted in such a way that they are fixed in the horizontal and can't move right or left, but can move up and down? In that case, the north magnetic pole dips downward to some degree toward the Earth's surface. This is called magnetic dip.

It may be that the first person to note this was a German vicar named Georg Hartmann. In 1544, he observed magnetic dip and wrote a letter on the subject, but it aroused no interest.

In 1576, an English navigator, Robert Norman (born about 1560, with the date of his death unknown) also made note of magnetic dip and this time the discovery made its mark.

Meanwhile, while all this was going on, it was only natural that people would wonder why a compass always insisted on pointing north. How did the compass know which direction north was?

Since it was known that a compass would point in the direction of a lump of iron because of the attraction between itself and the iron, why not suppose, then, that somewhere far in the north there was a really huge lump

>of loadstone, a whole vast mountain of it, and that that

;was what the compass was pointing to?

175

The first to suggest that such a mountain existed was Pliny, who told the story of the discovery of magnetism. He not only suggested a mountain but two such mountains, one of which attracted iron, while the other repelled it. He placed these mountains off the coast of India which, at that time, was considered the home of all marvels.

Pliny thought that anyone with iron nails in his shoes could not place foot on the repelling mountain, and could not lift his shoes off again if he once stepped on the attracting mountain.

A century later, the Greek astronomer Claudius Ptolemy (A.D. 100-70) reduced matters to an attracting mountain only, and placed it farther off, on the southern coast of China. However, he imagined the magnetic pull to be so strong that ships with iron nails were pulled forcibly to the mountain if they approached too closely and were held there forever. In the Middle Ages, the story was that the mountain pulled the iron nails out of the ship, reducing it to isolated planks. Everyone on the ship was then plunged into the sea and drowned. In The Thousand and One Nights, the ship of Sinbad the Sailor, in one of his voyages, does venture too near the magnetic mountain and is shipwrecked as a result.

Of course, once the Europeans began to explore the seas it was clear that no such mountain existed in any part of the known world. It would have to be far up north amid the polar ice, in any case, if the compass pointed to it.

That would acccount for magnetic dip because the compass would point straight at the mountain through the bulge of the spherical Earth. It would also account for the existence of magnetic declination if the magnetic mountain were not precisely at the north pole.

However, as the 1500s progressed, arctic exploration showed no signs of the nearness of a magnetic moun-

176

tain, and as the fact that magnetic declination changed with time became better understood, that gave rise to the puzzle that the magnetic mountain would have to be drifting.

The time was ripe for new insights, and that's for the next essay.

177

12 From Pole to Pole

Occasionally, I have problems I really don't expect.

These days I am writing a weekly science column for the Los Angeles Times Syndicate. (I can hear you say: "My goodness, Asimov, don't you have enough to do without that?" . . . The answer is, "No. I wish I had the strength and ability, as I have the desire, to write all day long, every day.")

Some months ago, in one of these columns, I referred to the new supernova in the Large Magellanic Cloud and said that it was 150,000

light-years away.

I realized that my newspaper audience might not have a quick grasp of what 150,000 light-years was, so I did a little calculation. One light-year is roughly 5.88 trillion (5.88 xlO12) miles, so 150,000 light-years is about 8.8 x 1017 miles, or nearly 1018 miles.

I hesitated to say 1018 miles, because I didn't think

178

that that would be easily grasped; nor would a 1 followed by 18 zeroes.

It seemed to me that the largest number that could be easily grasped by today's reasonably literate people was a billion. After all, we know what a billionaire is and we also know that the national debt is now two thousand billion. I decided, then, that the easiest way of presenting the distance of the supernova was to say it was nearly a billion billion miles away.

What I wrote, then, was this: "The light, as it reached us across a gap of 150,000 light-years (nearly a billion billion miles). . . ." There, I felt I had done everything as neatly as could be expected.

The syndicate sent out the article as I had written it, with the "billion billion miles." I know they did, because I questioned them and they sent me a copy of their version, and there it was.

However, in one of the newspapers that printed the column, there must have been some editorial soul-searching. After all, every once in a while a writer somehow manages to repeat a word in a sentence and says, "John gave Mary the the book," or "John gave gave Mary the book." Such a thing is just a careless oversight, so the rewrite man simply omits the extra "the" or "gave" and everyone is happy.

Faced, then, with "nearly a billion billion miles," |some rewrite man smiled in a fatherly fashion and did ic the favor of leaving out one of the billions, making it Ifead that the supernova, being 150,000 light-years away, [was also "nearly a billion miles" away.

The planet Saturn is nearly a billion miles away! The ipernova is a billion times as distant. Conceive, then, my embarrassment when I received a rtter from a very intelligent little girl named June Tay-|tor, who explained that she was in the third grade. (And was, for the letter was written on ruled paper in

179

what was clearly a nine-year-old's printing.) In her letter, she carefully went through the calculation of multiplying 150,000 by the number of miles in a light-year and got the right answer.

She then said: "As I have selected your article as a current event for my school work, I would appreciate your clarification."

I was appalled. It's been a long time since I was caught dead to rights by a nine-year-old. Naturally, I wrote a letter at once, explaining the matter. It took me a considerable time to recover.

Fortunately, in the case of my F & SF series, this can't happen because the Noble Editor always sends me galleys. This doesn't prevent me from making foolish errors, because I'm the world's worst proofreader, but at least the errors are my own and I always find it easier to forgive myself than to forgive others.

Anyway, we're on the subject of magnetism and I'll now continue.

I ended the previous essay with the problem of why the compass needle pointed north, and why there was such a thing as magnetic dip. The answer was provided by an English physician and physicist, William Gilbert (1544-1603), who spent the last two years of his life as physician to Queen Elizabeth I.

In 1600, he wrote a book entitled Concerning Magnets, which was full of careful observation and experimentation, so that Gilbert shares with Galileo the popularization of the notion of modern experimental science.

He tested some notions about magnetism by direct experiment. There were people who maintained, for instance, that garlic destroyed the magnetism of a compass needle. In those days, it was enough merely to quote some "authority" to that effect. Gilbert got him-

180

self a mess of garlic and rubbed it all over a magnet and was able to show that it had no effect on the magnetism whatever.

Others maintained that iron rubbed by diamonds would be magnetized just as though it were rubbed by a loadstone. (Why not? Diamonds are so valuable!) Gilbert went to the expense of obtaining seventy-five diamonds and, in front of plenty of witnesses, used them in various ways to attempt to magnetize iron. It didn't work.

The most important thing he did, however, was to take a large piece of loadstone and fashion a globe out of it. He located the magnetic poles on it, and showed that a compass needle would point "north" if placed near the surface of this spherical magnet.

What's more, if he arranged for the compass needle to swivel vertically, it showed magnetic dip for it pointed straight at the magnetic pole through the body of the object. In fact, if the compass needle was held above the magnetic pole, it pointed straight down.

Gilbert concluded, then, that compass needles acted

ic way they do, not because there is a magnetic iron Imountain in the north, but because the Earth is itself a ?huge magnet.

He was the first to maintain this and he was correct the whole, though wrong in details. For instance, he lought that the Earth was literally a large loadstone,

it that the surface, through long weathering by wind

id water, had lost its immediate magnetic properties ccept for occasional pieces of unaltered loadstone.

He also committed the very common scientific fault >f forcing facts to fit a theory. He assumed that the

ignetic poles would coincide with the geographic

)les of rotation so that the compass should point to the

ic north everywhere on Earth, which it clearly didn't.

" course, in Gilbert's time, not much was known about

181

the Earth's arctic regions and still less about its antarctic regions.

Inevidentally, the north pole of a magnet is defined as that end of the magnet which turns north. It was afterward discovered that north magnetic poles attract south magnetic poles but repel other north magnetic poles. The fact, then, that the north pole of a compass needle points north means that the Earth's magnetic pole in the north is a south magnetic pole.

However, no one is going to speak of a south magnetic pole in the far

north and no one is going to switch all the north poles in magnets into south poles and vice versa. We end up, therefore, with the paradox of having the north magnetic pole of a compass needle attracted to the north magnetic pole of the Earth.

Incidentally, Gilbert's book was not very popular in England, partly because he was a strong proponent of Copernicanism and he used his magnetic findings to argue that the Earth went about the Sun. This was considered preposterous by many scholars, who dismissed the book in consequence.

Gilbert's insistence that the magnetic declination (the direction in which the compass needle points) was unchanging was finally disproved by an English astronomer, Henry Gellibrand (1597-1636). He carefully recorded the direction in which compass needles pointed and, in 1635, published his findings. He showed that in the past half century, magnetic declination in London had shifted by seven degrees. The angle of the magnetic dip also changed. (Even the intensity of the Earth's magnetic field changes with time, we now know.)

The reason for the change in magnetic declination and magnetic dip was a mystery. There was even some speculation that there might be four magnetic poles on Earth. In 1698, the English astronomer Edmund Halley (1656-1742), of later Comet Halley fame, was sent off on

182

•tf

| an ocean voyage to discover, if he could, east and west f magnetic poles. It was the first ocean voyage designed "> for a specifically scientific purpose, and not for explora-' tion. However, Halley did not find the additional poles ?• since they were not there to find.

It would help, of course, if we could find out where | the magnetic poles of Earth are actually located.

About 1830, the German mathematician Carl Frie-

drich Gauss (1777-1855), making use of observed com-

ipass measurements, calculated that Earth behaved as

" though there were a very powerful bar magnet buried in

its center.

He showed that this bar magnet was not set parallel to

-the axis of Earth's rotation. This was the first indication

^

"lat Earth's magnetic poles were not located at Earth's ^geographic poles. Instead, the so-called geomagnetic axis K passed through the center of the Earth, making an angle 12 degrees to the rotational axis. Gauss calculated that the north geomagnetic pole was •cated at 78.5 degrees North Latitude and 70 degrees /est Longitude. This is located in Hayes Peninsula in northwestern Greenland, just 35 kilometers north of the

* icrican base at Thule, and about 750 kilometers from le north pole.

The south geomagnetic pole would be at precisely the Opposite side of the globe, 78.5 degrees South Latitude id 110 degrees East Longitude. This is deep in Antarc-ica, very nearly at the maximum distance from the ^ocean and, therefore, at the region of greatest cold where the Soviet base, Vostok, is established. (This is pure coincidence, of course. It is 750 kilometers from the south pole.)

Once Gauss had established where the geomagnetic poles ought to be in theory, one of the goals of polar

183

expeditions came to be the confirmation of this. Explorers wanted to find that spot in the Arctic where the north pole of the compass needle pointed straight down and that spot in the Antarctic where it pointed straight up, and see how close Gauss had gotten with his theoretical calculations.

It was quickly discovered that, in the Arctic at least, Gauss's calculation was way off. It came to be clear that the north magnetic pole was not at the north geomagnetic pole. In fact, the north magnetic pole wasn't even in Greenland.

The Scottish explorer James Clark Ross (1800-62) discovered the north magnetic pole on June 1, 1831. It was on the western shore of Boothia peninsula at 70.85 degrees North Latitude and 96.77 degrees West Longitude. This point is located in the northernmost extension of the North American continent and is 1,100 kilometers southwest of Gauss's north geomagnetic pole, and fully 2,100 kilometers from the

north pole itself. (The distance between the north magnetic pole and the north pole is equivalent to that from New York City to Dallas, Texas.)

The south magnetic pole seemed a much harder nut to crack. It seemed certain to be somewhere in the body of Antarctica and no one had yet managed to penetrate the continent. They were merely nosing about the icy coastlines.

In 1840, however, a French explorer, J. S. C. Dumont d'Urville (1790-1842), was sailing along the Antarctic shore and found a section where the compass needle pointed nearly straight upward. He knew he was fairly close to the south magnetic pole, though not right on it.

By 1909, exploration parties were beginning to penetrate Antarctica and one of them, under an Australian explorer, Edgeworth David (1858-1934), located the south magnetic pole, 250 kilometers inland from the

184

jstern shore of the Ross Sea. It was at 72.42 degrees >uth Latitude and 155.27 degrees East Longitude. It about 1,400 kilometers northeast of Gauss's south >magnetic pole and 1,600 kilometers from the south >le itself.

Tb make matters more complicated, both magnetic les move. Since its discovery, the north magnetic pole moved 500 kilometers closer to the north pole, the south magnetic pole is moving away from the 'outh pole and is now almost exactly on the Antarctic ore where Dumont d'Urville would have discovered it he had come at the right time. What's more, the magnetic axis, that is, the line pass-through the Earth from pole to pole, from the north ictic pole to the south magnetic pole, does not pass >ugh the center of the earth. It misses the center by less than 1,100 kilometers.

| So you see, there are a number of questions about the ictic poles. Why are "they so far from the geo-iphic poles? Why do they move? Why aren't they on ictly opposite sides of the globe? Most of all, why is the intensity of the field changing? Since 1800, the intensity of Earth's magnetic field has :lined by about 10 percent. If this goes on, then in a)uple of thousand years, it will decline to zero and then rerse, so that the north magnetic pole of the compass jdle will begin to point southward. This has happened fnumber of times in Earth's history, with such magnetic mis taking place at very irregular intervals.... No knows why.

Let's think about magnetism. Why does iron behave so differently from other materials? In the course of the nineteenth century, it became clear that electricity and magnetism were closely allied; that electric currents

185

could show magnetic properties and that magnets could produce electric currents. The laws governing the properties of the electromagnetic field were worked out between 1864 and 1873 by the Scottish mathematician James Clerk Maxwell (1831-79). He showed that it was impossible to have electricity without magnetism, and vice versa.

Every substance is made up of atoms. Every normal atom has, in its outskirts, electrons. Every electron has an electric charge and is, therefore, a small magnet.

Electrons have something called spin, and the spin can be in either one direction or the opposite direction. If an atom has an even number of electrons, half spinning in one direction and half in the other, the magnetic effects tend to cancel out.

If there is an odd number of electrons, at least one is not going to be in balance. Sometimes, even with an even number, two may be unbalanced. In those cases, atoms and the material made up of them will show weak magnetic effects.

In some cases, there is a tendency for the imbalance in the electrons of many atoms to line up in the same direction, so that the magnetic effects pile up. Such a substance can exhibit strong magnetic properties. Loadstone, a naturally occurring oxide of iron, is an example. |j

Ordinary iron has domains within which large numbers of atoms with unbalanced electrons line up. The ,| individual domains, however, point in every possible direction and cancel each other out. If iron is stroked with loadstone, the domains are all pulled into the same direction and the iron becomes strongly magnetized. Once the loadstone is removed there is a tendency for the domains to move into different directions again and the magnetism is lost. Thus, iron tends to be a temporary magnet.

The iron atoms in steel are held more tightly and,

186

>nce they are stroked into the same direction, cannot iily move out of alignment again. Steel, therefore, ;nds to form a permanent magnet. This ability to line up vast numbers of unbalanced

slectrons to produce a strong magnetic effect is called

zrromagnetism, the prefix coming from the Latin word

for "iron."

Although iron is by far the most common substance mt displays ferromagnetism, it is not the only one. lere are two metals that are very like iron, chemically, id these are nickel and cobalt. Both of them are ferro-lagnetic and will be attracted by a magnet. All this seems to strengthen the view that the Earth's lagnetic field is based on a central core of iron. After 1, some 7 percent of the meteorites that fall are a mix-re of iron, nickel, and cobalt in a ratio of 90, 9, and 1, id they may be remnants of central cores of asteroids, icre is no doubt that iron is the most common of the lore complex elements of the Universe, and the density * the Earth, as a whole, fits the thought that there is a rge nickel-iron core at the center. Of course, that doesn't explain why the magnetic >les are off center, move about, and so on, but those little details that can be taken care of later—were it

|ot that the iron-core theory falls apart altogether for le following reason:

A ferromagnetic substance retains its strong magnetic >perties as long as its atoms are held firmly in place in ich a way that the unbalanced electrons are all lined >. At any temperature, though, the atoms are vibrat-l, and the higher the temperature, the more vigor-isly they vibrate. Eventually, if the temperature is suf-aently high, the atoms are vibrating energetically

Enough to slip their moorings and, with their electrons, jgin to take on all sorts of random positions. This was demonstrated to be so in 1895 by the French

187

chemist Pierre Curie (1859-1906), who, in that same year, married Marie Sklodowska, who was to become the famous Madame Curie. The temperature at which ferromagnetic substances lost their ability to be ferromagnetic (the Curie temperature) varies.

The Curie temperature of nickel is 358 C, that of iron is 770 C, and that of cobalt is 1131 C

Oddly enough, there is a fourth ferromagnetic element, one that is not chemically related to these so-called iron elements. The fourth ferromagnetic element is gadolinium, one of the rare earth metals. There are thirteen other very similar rare earth metals, but only gadolinium seems to be ferromagnetic. (Please don't ask me why.) Its Curie temperature is only 16 C (60 F), so that on a chilly day, gadolinium will be attracted by a magnet but as the day turns fairly mild, the metal will drop off.

This business of the Curie temperature seems to knock the iron-core theory of Earth's magnetism for a loop. The latest determinations show that the metallic core of the Earth is at a temperature of 3500 C at its outermost rim. That temperature goes up steadily to one at the very center of 6600 C. All of it, then, every bit of it, is far above the Curie temperature of any known substance, which means that the center of the Earth is simply not a magnet in the ordinary sense of the word.

Why, then, does the Earth have a magnetic field?

The German-American physicist Walter Maurice El-sasser (b. 1904) feels that the answer may lie in electro-magnetism. In 1939, he suggested that the Earth's rotation sets up slow eddies in the iron core, which is hot enough to be liquid (except, perhaps, at the very center, where high pressure keeps it solid). A moving electric conductor sets up a magnetic field and it is that which we experience.

Of course, we should expect the eddies to be parallel

188

to the direction of rotation, so that the magnetic axis

t-will be lined up with the axis of rotation. This isn't so.

i The magnetic poles are far from the geographic poles,

the magnetic poles move, and the line from pole to pole

does not pass through Earth's center. No doubt these

asymmetries can eventually be explained, but what that

explanation may be, I don't know.

Then, too, we might suppose the magnetic intensity

decreases or increases according to whether the speed

of swirl decreases or increases. Right now the liquid

; iron core is swirling more and more slowly. Eventually,

I it will come to rest and the magnetic field will disappear.

Then, it will begin swirling in the opposite direction and

the magnetic field will reverse itself.

But why does the swirl decrease and increase? If the

|iarth always turns from west to east, why would the

liquid core swirl with the Earth's rotation at some times

and against it at other times? We don't know—or, at any

rate, / don't know.

We can test other heavenly bodies, however. If Els' sasser is correct, there are two things necessary for a Iplanetary magnetic field. First, there has to be a liquid ire capable of carrying an electric current. Second, there must be a period of rotation fast enough to set up ^swirls in that liquid.

Earth meets both requirements. The Moon, on the other hand, meets neither. From its low density, we Jknow that it is rock all the way through and it is simply ,not hot enough at the center to melt that rock (rock has a higher melting point than iron has). Even if the rock were molten, it would not carry an electric current. On top of that, the Moon rotates on its axis, relative to the Universe generally, in 27V3 days, rather than in 1 day, as Earth does. The result of all this is that you wouldn't expect the Moon to have a magnetic field—and it doesn't.

189

How about Mars? Like the Earth, it rotates on its axis pretty quickly— 241/2 hours. To be sure, it's a distinctly smaller body than Earth is so that its speed of rotation is not much more than half that of Earth, but it is fast enough to set up swirls.

Or it would be fast enough to set up swirls if there were something to swirl. Mars's density is low enough so that we can conclude it has

little or no liquid metal core, and, therefore, despite its rapid rotation, it should have no magnetic field—and it doesn't.

Venus is almost as large as Earth and almost as dense as Earth. It undoubtedly has a liquid metal core and the liquid metal is surely iron. So far, so good—but Venus has a period of rotation of 243 days, the slowest period of rotation in the Solar system. That's not enough to produce a set of swirls despite its liquid iron core. It should, therefore, have no magnetic field to speak of— and it doesn't.

Jupiter is made up almost entirely of hydrogen, with a little helium thrown in. There may be a solid ball of rock and metal at the very core—quite small in comparison to Jupiter itself, but, for all we know, as large as Earth. We just don't know enough about Jupiter's interior to be able to tell.

However, suppose that Jupiter is largely hydrogen. Under the huge pressure at the center, the hydrogen is in metallic form. That means the single electron of the hydrogen atom is very loosely held and the hydrogen can carry an electric current easily.

In addition, Jupiter has a very rapid rotation rate of just under ten hours despite the fact that it must turn through a circumference eleven times that of Earth. There is thus a liquid material at its center capable of carrying a current and a rate of rotation that should make it swirl like crazy. It should not only have a magnetic field; it should have an extremely intense one.

190

And it does. Its magnetic field, measured by probes that skimmed past the planet, is some nineteen thousand times as intense as Earth's.

Probes have also measured the magnetic fields of Saturn and Uranus, the properties of which are like those of Jupiter, though less extreme. Uranus's magnetic field is only fifty times as strong as Earth's. Its magnetic axis is tipped no less than 60 degrees to the rotational axis, and the center of the magnetic axis misses the geographical center of the planet by a full 8,000 kilometers.

Voyager 2 has just observed Neptune, and it, too, has a magnetic field, as astronomers were certain it would have.

The Sun, like the gas giants, apparently has a conducting interior and though it rotates only once in 26 days, its huge size makes the rotation rate fast enough for swirls. Hence, there is a strong magnetic field, as is evidenced by the sunspots, if nothing else.

That leaves Mercury. It is a small planet, smaller than Mars, though larger than the Moon. It is, however, just about as dense as Earth. Considering that it is smaller and that its central regions must therefore be less compressed than Earth's are, we can assume not only that it must have a metallic core, but that that core is probably a bit larger in proportion to its overall size than the Earth's is.

However, Mercury rotates slowly, only once in 59 days. It is not as slow a rotator as Venus, but it is slower than the Moon and it should not turn quickly enough to swirl the metallic core. So it should not have a magnetic field. . . . But it does. Just a weak one, but a more intense one than you would expect.

My own feeling is that there is just a chance that small Mercury has a central temperature that is cool enough

191

to allow a little ferromagnetism. It doesn't seem likely, but perhaps there's just a chance.

In any case, there's a great deal about astronomical magnetic fields that people do not understand.

192

13 The Fire of Life

t When my parents first arrived in the United States, with

| my three-year-old self in tow, they moved into a very

primitive apartment, for that was ai! they could afford.

It had no electricity, but only gas jets for illumination. It had a wood-burning stove, and an icebox rather than a refrigerator.

The stove was my special delight. My mother would |. light it with old newspapers and then put in sticks, and in cold weather, she would leave the door open for a while to help warm the kitchen. I would watch eagerly, for the fire seemed alive, consuming the paper and then seizing hold of the wood and creeping along it, turning blue and yellow and liberating a delightful odor along with its warmth.

Years later I realized that the place was a horrible slum apartment but, of course, I didn't know that at the

195

time, and when, after two years, we moved out, I wept bitterly.

Nor did the new apartment comfort me. It had no gas jets, but electric lights instead, which burned with a dead, unchanging glare one couldn't look at (it took a while for us to learn about frosted bulbs). And the new gas stove was a terrible disappointment. I never saw living flame in a stove again and I didn't see how a gas stove could possibly cook. It didn't even have stovelids, which one could remove with a special holder and look inside at the fire. Of course, in mature life, I have occasionally watched someone's fireplace, but the magic was never quite the same—one must be a child.

Since I have never been part of a conflagration (and I never want to be, you understand), I have no experience of the horror and deadly danger of fire. I remember only the delight and beauty of it when I think of those old, old days. Nowadays, when I know somewhat more than I knew when I was a little boy, I think of the intimate relationship of fire and life—and particularly of fire and human life—so that's what I'd like to talk about now.

The usual state for any planet is that of being dead. By that I mean that all the changes that can take place on it have just about taken place, and nothing much will or can happen further. The Earth, in the first few hundred million years of its existence, was nearly dead. It had developed an ocean and an atmosphere. The ocean was mostly water; the atmosphere was essentially a mixture of carbon dioxide and nitrogen.

The Earth, however, was not completely dead. Any time there is a difference in temperature between one part of a body and another part, that body is not dead, for heat will flow from the point of high temperature to

196

the point of low temperature and that will produce change.

Earth has two types of temperature difference. First, the planetary interior is much hotter than the planetary surface. This produces the cracking and shifting of the crust, together with earthquakes, volcanoes, mountain ranges, ocean deeps, and so on.

Second, the Sun is much hotter than the Earth, so that heat flows from the Sun to the Earth's surface during the day, and from the

Earth's surface to outer space during the night.

If a planet is completely dead, it might be viewed as having rolled downhill and to be resting motionlessly in the deepest part of the valley. The flow of heat, whether from the Sun or from the planetary interior, tends to drive the planet uphill slightly. The balance between the flow of heat and the natural tendency to move downhill balances and the planet stays slightly uphill at all times.

From the chemical standpoint, the flow of heat forces the simple molecules of the ocean and atmosphere to combine into more complex molecules, which have a larger energy content. The formation of these more complex molecules represents an uphill movement. With time, more and more complex molecules are formed until eventually some are so complex that they have the properties we associate with life.

The chief property of life is its ability to maintain itself in an uphill position by pushing parts of the environment downhill and making use of the energy liberated.

We then have the situation of the radiation of the Sun pushing certain molecules uphill and of life maintaining itself uphill by ruthlessly pushing some of those uphill molecules downhill again. This sort of thing is sufficient to maintain a population of bacterial cells in Earth's

197

oceans and, for over two and half billion years, that's all there was on Earth.

Life managed to improve on the situation, however. Certain bacterial cells developed photosynthesis, evolving substances that made it possible to use the energy of visible light to form complex molecules in much more massive quantities.

Photosynthesis made it possible to drive Earth much farther uphill and this, in turn, made available much more energy when the compounds were allowed to drop downhill again. This gave the cells a much larger food supply so that they had the wherewithal to grow more complex and to associate with each other to form mul-ticellular organisms.

What's more, in driving chemicals uphill, molecules were produced that retained the carbon and hydrogen atoms (and some other atoms, too), but retained only a few of the oxygen atoms. The oxygen not retained was discharged into the atmosphere so that slowly the

carbon dioxide content declined and the oxygen content rose.

It was the removal of oxygen atoms that increased the energy content of the molecules. In moving downhill, the molecules that were rich in carbon and hydrogen and poor in oxygen combined with atmospheric oxygen, giving up part of its energy content, which could be made use of by the various life forms.

Therfree oxygen content of Earth's atmosphere is maintained by the photosynthetic action of life. If photosynthesis (found in the green plants of the world) were to disappear, the complex molecules that now exist on Earth would slide downhill, combining with the oxygen and producing carbon dioxide. The oxygen would disappear and would not be replaced and the Earth's atmosphere would become a mixture of carbon dioxide and nitrogen as it was in the early period before life had

198

come into existence. And life could not exist at the stage past the bacterial.

It is possible, under certain conditions—say, a temporary rise in temperature to a high point—for the downhill slide to reach catastrophic speed. In that case, there is a release of a great deal of energy in a short time, energy that makes itself felt as heat and seen as light. In short, there is a fire,

This can only happen if there is a certain amount of free oxygen in the atmosphere, so that much of it is available for runaway combination with carbon/hydrogen compounds. What's more, it can only happen if water is present in only limited quantities, for water, present in excess, prevents temperature from rising high enough. And since water does not itself combine with oxygen, it tends to dilute and damp out the combination of other materials with oxygen.

This means that even if ample supplies of free oxygen are present in the atmosphere, fire is impossible as long as life is confined to the waters of the world.

This is not to say that there wouldn't be heat and light in the world. A volcanic eruption may send a stream of glowing lava across Earth's land surface. Lightning may send flashes of heat and light through the atmosphere. There is, however, nothing for the lava or the lightning to set on fire.

It was not until about 450 million years ago that plant life began to

invade the dry land and only by 410 million years ago were there the first forests.

Land plants are, to a large extent, dry, so that if a stream of lava flowed into a forest, or a lightning bolt hit a tree, the catastrophic downhill movement of the molecules would be initiated and there would be a fire. The fires would then continue till the denser parts of the forest were burned out, or until a rain fell.

(Recently, there was a report that analysis of trapped

199

air bubbles in amber showed that, in the time of the dinosaurs, the atmosphere was 32 percent oxygen, instead of 21 percent as it is today. I could not believe that. With an atmosphere that was one third oxygen, forest fires, it seemed to me, would never go out, and land life would become sparse indeed,)

It is odd to think that fire has only been possible during the final tenth of Earth's existence so far, but land animals have only existed during that final tenth, so fire has been part of their total experience.

There is no telling when or where a fire could start. We can't tell when a volcano will blow its top, or which one will do so next. Even if a forest is well away from any volcano, it could still be subject to the blind blow of a lightning bolt. Once a fire does start, plant life, which is immobile, can do nothing but burn, and animal life which is too slow to outrace the fire can do nothing but burn, too. Those animals that can run fleetly, however, do so, and there can be no panic like that of trying to stay ahead of a deadly, devouring monster whose hot breath you feel behind you.

It is with respect to fire that humanity has marked itself out clearly from all other life forms.

There are other animals that are tailless; other animals that walk on two feet; other animals that communicate rather subtly; other animals that use tools and even make them; other animals that can, after a fashion, reason or create. In almost all respects human beings differ from other animals in degree rather than in kind. With respect to fire, however, the difference is absolute. All human societies, without exception, make use of fire. No species of life that is not human makes use of fire or has ever made use of fire. How did this come about?

200

We don't know, of course. We can only speculate. There comes a time when a forest fire dies down. It .} has run out of easily available fuel, or it has been '- drenched in a rain, but a few twigs, shrubs, or patches of grass are still smoldering, or are burning in final, feeble gasps.

Human beings have run from the fire along with all

;   other forms of life capable of doing so, but now only

| human beings, with their overpowering curiosity, will

f linger to watch. My own feeling is that it was children

who watched the fire, when it seemed safe, with the

same absorption that I watched the fires in the stove

when I was a child.

It is inevitable that, as the fire died down, some child would feed it another twig or a handful of brush. It is also inevitable that the mother would come and snatch the child away and stamp—stamp—stamp the fire. . . . And maybe hand the child a juicy one on the ear, for his or her own good.

It is also inevitable that, eventually, some adult would say to himself (or herself), "Hey, if we drag that thing inside and we're very, very careful, it will light the place, and keep us warm."

In any case, a cave near Peking was discovered about 1927 in which bones were found that indicated occupancy by very early human beings—say, 500,000 years ago. Along with these bones were signs of campfires.

Consequently, the use of fire dates back at least 500,000 years, which means it was not discovered by Homo sapiens but by our hominid predecessor, Homo

erectus.

Fire was an enormous boon to hominids. By giving light and warmth, it made it possible for hominids to move out of the tropics. It was also useful as a way of inflicting a salutory fear on other animals, even the fiercest.

201

Only human beings learned not to be unreasonably afraid of fire. A fire in a cave, or within a circle of stones, would keep the predators away. They might snarl and slink about the outskirts, but that would be all. In fact, I imagine people would carry burning branches to scare game and set them to stampeding into traps,.

Then, too, fire made it possible to cook food. Meat was made softer and tastier if roasted. What's more, the roasting killed worms and bacteria so that the meat was safer to eat. Eventually, fire made plant food, otherwise inedible, most palatable. Try eating rice or corn on the cob before heating them and you'll see what I mean.

Fire also made possible various chemical changes in inanimate matter (soft clay into hard pottery, sand into glass, ores into metals, and so on). In short, fire introduced humanity's first age of comparative "high tech."

To begin with, of course, fire could be obtained only after it had been started by natural means. Once one had a fire, it had to be kept burning continuously, for if it ever died out the search for another fire would have to be instituted at once.

The time came, however, when techniques were developed for starting a fire where none had been before. This could be done by friction, by turning a pointed stick in a depression in another stick, a depression that contained very dry shreds of wood, leaves, or fungus (tinder). The heat of friction might eventually ignite the tinder.

We don't know when such methods were first developed, but the technique of starting a controlled fire where none existed before would represent another enormous step forward.

The original fuel for fire was wood in one form or another, whether a huge log, or a bundle of twigs and

202

grass, or anything in between. The material was all around and it was easy to burn.

From a chemical standpoint, wood, or plant tissue generally, is extraordinarily complex, but the chief component is cellulose, which consists of giant molecules that are, in turn, made up of a rather simple building block. The building block is made up of six carbon atoms, ten hydrogen atoms, and five oxygen atoms and that combination can be used, more or less, to represent wood as a fuel.

Notice that wood is partly oxidized. Oxygen is already present in combination with carbon and hydrogen, but that is only a partial oxidation. If the carbon and hydrogen were completely oxidized, they would become carbon dioxide (with molecules consisting of one carbon atom and two oxygen atoms) and water (with molecules consisting of two hydrogen atoms and one oxygen atom). Every carbon atom would require two oxygen atoms and every two hydrogen atoms would require one oxygen atom.

This means that the six carbon atoms of the cellulose building block would require twelve oxygen atoms and the ten hydrogen atoms would require five oxygen atoms for a total of seventeen. Only five oxygen atoms exist in the molecule, so twelve more must be obtained from somewhere, and that somewhere is the atmosphere, where it exists as oxygen molecules made up of a pair of oxygen atoms each.

So we combine the building block formula with six oxygen molecules to get six carbon dioxide molecules and five water molecules.

In order for wood to burn completely, in accordance with the equation, oxygen has to reach all parts of the wood. This happens generally as far as the wood in a campfire or in a fireplace is concerned. The wood is piled together loosely and the heat of the fire causes the

203

air above it to rise, producing a draft that brings fresh air into the neighborhood of the wood.

However, on occasion people might want a large fire —to roast a whole antelope at some festival, for instance—and, in that case, oxygen doesn't get to the bottom portions of the pile of wood in any great quantity.

The heat of the fire makes the complex molecules in wood break down, causing water to steam off and also producing small molecules of carbon-containing vapors. These vapors are inflammable, mix with air, and combine with the oxygen content to give off light and heat over a sizable volume of the mixture. The actual flame of a fire is the mixing and combining of inflammable vapors and oxygen. As the wood breaks down, releasing water and inflammable vapors, there is a residue left behind that is richer and richer in carbon atoms until finally what is left over is almost entirely carbon.

The carbon left behind can be made to burn, but the burning is

difficult to get started. Once the burning does start, it does so without flame, since carbon does not vaporize until extremely high temperatures are reached. It therefore burns only at the surface, glowing quietly and persistently, and with a higher temperature than that of ordinary burning wood. This carbon residue is called charcoal, which may come from old words meaning "turning to ember." (An "ember" is a lump of matter that burns without actual flame.)

The conversion of wood into charcoal may have been the first chemical process developed by human beings for the production of a useful substance. We don't know when it first happened, but it must have taken place deep in prehistoric times.

Charcoal may have had only limited uses, however, until 1500 B.C. By that time, metallurgy had existed for a couple of thousand years. Ores had been heated to obtain silver, copper, and bronze, for instance. Iron would

204

have been a particularly valuable metal but it didn't seem to exist in the ores. (Iron was known, however, because it could be found in the form of meteorites.)

Someone must have started a charcoal fire on rocks that happened to be ir6n ore and found drops of iron in the residue because by about 1500 B.C., the Hittites in eastern Asia Minor had developed a technique for smelting iron ore with charcoal. The higher temperature of burning charcoal was needed to force the oxygen atoms that were in combination with the iron atoms to combine with carbon atoms instead, leaving the iron atoms free. (Properly done, as was learned some centuries later, some of the carbon mixed with the iron to produce steel, a particularly hard and tough alloy of carbon.)

Since iron soon turned out to be absolutely necessary for tools, weapons, and armor, the demand for charcoal grew rapidly and charcoal production became a vital industry.

Now a vicious circle set in. As fire was used for more and more purposes to increase the food supply and add to human security, the population naturally increased so that still more fire had to be used to continue to produce the good things and make further population increase inevitable.

It must have seemed to early man that the supply of wood was infinite, or virtually so, since new wood grew as fast as old wood was used. And yet, as the population grew and the uses of fire multiplied, a

deforestation began to take place. This was hastened as human beings turned to the production of charcoal in quantity. Charcoal production is very wasteful of wood since so much wood must be burned away in order that some of it is left behind as charcoal residue.

People had to go farther and farther afield to find

205

wood and, eventually, there was pressure to find an alternate fuel.

The answer to the problem came about because the process of charcoal formation had already taken place in nature on an extremely large scale (and extremely slowly—but what's time to a planet?).

Beginning about 345 million years ago, and continuing for over 100 million years, huge forests of primitive trees grew in large areas of low, flat, swampy land. These trees eventually died and fell into shallow water where they were slowly covered by mud and sediment not particularly rich in oxygen. This made total decay difficult. Some decay did take place, but the residue grew richer and richer in carbon. There developed therefore a kind of charcoal.

Ordinary charcoal, made by human beings, is rather light and crumbly. The charcoal made out of the decaying trees that were covered by mud and sediment was compressed under the weight of overlying layers and became dense and hard. It still burns and smolders, but it does not resemble ordinary charcoal. It is called simply coal, therefore.

Even today, coal is forming. There are swampy, boggy areas where decaying plant material can be dug up and dried out to be used as fuel. This is peat. Some of the hydrogen and oxygen has already been lost as vapors so whereas fresh wood is about 50 percent carbon, peat is 60 percent carbon. The next stage is lignite, which, when it is dry, is nearly 70 percent carbon.

Beyond that is a kind of coal that is about 85 percent carbon. If this coal is heated in the absence of air so that it doesn't burn, the 15 percent that is not carbon is driven off, along with some of the carbon. This type of coal is called bituminous coal because the material

206

driven off is a black tar, or pitch, and in ancient times pitch was called bitumen.

Finally, there is a kind of coal that is at least 95 percent carbon. This burns with a red-hot glow, forming an ember, as charcoal does. The Greek word for "ember" is anthrax, so this kind of coal is called anthracite coal.

Coal is forming much more slowly these days than in past ages, when those large primitive forests in the swamps existed. Peat and lignite therefore make up only a small percentage of all the coal in the world. Anthracite coal forms in only a few areas where there was a great deal of pressure. It too makes up only a small percentage of all the coal in the world.

Most coal is bituminous coal, and there is a great deal of that under the ground. There may be as much as 8,000 billion tons of it here and there in the Earth.

Though almost all of this coal is underground, some few coal seams may have been heaved upward and uncovered by geological processes so that occasionally lumps of coal might have been found lying on the ground. They can scarcely have attracted much attention. Yet, once in a while a piece of coal might happen to burn. Perhaps a lump of coal was accidentally kicked into a campfire; or perhaps a lump just happened to be on the ground in the place where a campfire was built. Then it might be noticed that, after the fire was out, this odd piece of black rock was still smoldering. Eventually, people must have started looking for such black rocks in order to use them as fuel.

The Chinese did so first, and when Marco Polo visited China in 1275, he took note of this and, in writing his travel book in 1295, he spoke of the Chinese using black stones as fuel.

So Europeans started looking for lumps of coal and some people, when they found them, must have wondered if they could find more if they dug underground.

207

Such digging was done first in the Netherlands, and underground coal was found.

The English learned of this and took particular note because, by 1600, most of their native forest was gone, and what was left was earmarked for the English navy on which the nation's security depended.

The English therefore started looking for coal with particular intensity and, by 1660, were producing 2 million tons of coal each year. This was more than 80 percent of all the coal that was being

produced in the world.

At first, coal was used only as fuel to cook food and to warm the houses in winter. It was bituminous coal that was used and it burned with a smoky, sooty, smelly flame. London became a dirty city indeed.

Despite the coming of coal, wood still had to be burned to produce charcoal for iron smelting. In 1603, however, an Englishman, Hugh Platt, discovered that if bituminous coal was heated in the absence of oxygen and the pitch was driven off, what was left behind was something very much like charcoal. It was called coke.

At first, the coke was of indifferent quality and did not work well as far as iron smelting was concerned. It wasn't till 1709 that an Englishman, Abraham Darby, was able to use coke on a large scale for iron smelting.

Then, as more and more coal was needed, some way had to be found to pump the water out of the coal mines quickly. In order to pump out the water, steam engines were invented, and they could be used in quantity largely because the steam could be formed by burning coal under the water containers. And because steam engines could be used in quantity, they could be used to power factory machinery, steamships, steam locomotives, and so on.

In short, it was coal that was the power behind the Industrial Revolution, and it was England's experience

208

with coal mining that made it certain the Industrial

Revolution would begin there and not somewhere else.

Yet coal was not destined to remain the king of fuels

forever, either. I'll continue the story in the next essay.

209

14

The Slave of the Lamp

My periodontist is a wise guy. No doubt he thinks I am one, too, but he has an advantage over me. Four times a year, he pokes around my gums with sharp instruments of torture and makes comments about

their condition, comments which verge on personal insult.

Naturally, I try to hand it back, but since he usually arranges to have my mouth full of blood, my natural ebullience is dampened.

Last week, though, I got him.

He said to me, "Your gums are in pretty good condition. What have you done? Changed your life-style?"

I said gravely, "I attribute it to excellent periodontal care, Joel."

Whereupon Joel, smiling fatuously, said, "All right FU accept that." -

210

To which I replied, "And then, of course, I sometimes come here."

All his muttering, jabbing, and general butchery couldn't keep me from grinning for the rest of the session.

So while I'm still in good humor, I'll continue the discussion of fuels that I began in the previous essay.

In the previous essay, I talked about solid fuels: wood, charcoal, and coal, where charcoal and coal are ultimately derived from wood.

Wood, however, though the most easily available fuel in very ancient times, was by no means the only one. There was another fuel and it must have been discovered in Neanderthal times. I imagine the discovery must have been an accidental one.

After all, if meat is roasting over an open fire, fat upon it will sizzle and burn. Or it will melt, drip down, and burn in the fire beneath. Eventually, people watching this will get the idea that animal fat (or plant fat, like olive oil, for that matter) will burn.

So, at some dim time in the past, torches were invented. Perhaps the idea arose when resinous wood was burned. Such wood burned with a brighter light and for a longer time than dry wood did, but once the resin was burned, the advantage was gone.

Some prehistoric genius, therefore, thought of making wood artificially resinous by dipping a piece of porous wood, or a bundle of reeds, into oil or melted fat. The torch would then burn brightly and, when the flame started to fade, it could always be extinguished and dipped in liquid fuel again. (Or a new piece of wood might be used—

wood was cheap.)

But then someone was bound to think of the fact that the wood was unnecessary.

211

Suppose you hollow out a depression in a rock and fill it with absorbent material, such as tinder or moss. You then soak the material with oil or melted fat and set it on fire. It will burn for a long time and, when the fire burns low, you need only add a little more liquid fuel carefully. This is a lamp (from a Greek word for "torch") and it came into use from 20,000 to 70,000 years ago.

Of course, you can carry a torch and hold it high for better or wider illumination, whereas a primitive lamp is too easily tipped and spilled to be portable.

Naturally, lamps would be improved. Instead of making them out of rock, you could make them out of clay or, later, metal, giving them a more convenient shape and making them lighter.

Furthermore, the wick must have been invented early on. In a sense, it was merely a tiny, artificial torch. One only needed something porous, some twisted moss, a pithy reed, or, later, a strip of textile material, which would absorb oil. One end is placed in the oil, which soaks up into the wick, and the other end is set on fire. As the oil burns, more oil invades the wick from below.

The lamp can be covered, to minimize the danger of spillage, though of course, some opening had to remain for the wick to emerge. More than one wick could be used to give more than one flame and produce more light. As many as twenty wicks in one lamp have been found in archeological digs. However, the more wicks there are and the more light one gets, the faster the oil is used up. (This may have been one of the earliest hints to humanity that there is no such thing as a free lunch.)

By ancient Greek times, lamps looked something like teapots, with a handle at one end, so they could easily be carried about, and the wick in the spout. This is the familiar Aladdin's lamp shape.

We can wonder why it was that rubbing a lamp—

212

rather than a vase or a chair—should produce that wonderful genie. It

strikes me that a lamp already puts a slave at one's service. The lamp makes it possible to carry light wherever one goes, and you can't overestimate the importance of light in primitive times (or, for that matter, now). It seems to me that the slave of the lamp (light) is so important that getting a genie out of it to shower you with palaces, wealth, and women is something you would expect of a lamp.

Of course, it is possible to have a wick without a lamp. If you impregnate the wick with solid fat of one sort or another, and pile the fat about it, you can set fire to the wick on top and it will slowly burn downward, as melted fat soaks up the wick. This is a candle (from a Latin word meaning "to glisten").

Candles go back to at least 3000 B.C.

What were the fuels used in lamps? In northern climates, where fire was more needed and more used and where lamps and candles may have been invented, the blubber from sea animals was the logical choice. Even as late as the nineteenth century, whale oil was a common lamp fuel.

In more southerly climates, it would be plant oils that were used— olive oil, linseed oil, and so on.

For candles, what was mostly used was tallow, the solid fat of cattle and sheep. Wax could also be used; in particular, beeswax, which was hard and which burned cleanly and odorlessly. However, beeswax was expensive and was used mainly in churches and in aristocratic homes. Spermaceti, a wax from sperm whales, was used in more recent times.

Advances were made both in candles and in lamps in the nineteenth century. For instance, as candles burned down, the charred wick (or snuff) would gradually stick

213

up above the flame, looking ugly and producing smoke. Therefore, anyone who used candles had to keep "snuffing" it—that is, cutting off the spent wick judiciously—and that was a bother.

In 1824, however, a Frenchman, Jean Jacques de Cambacere (1753-1824), invented a braided wick that bent as it charred so that its end moved into the hot part of the flame and gradually burned away. There was no need of snuffing with such candles—a small matter, but something that must have been a delight to candle users. As a result, "snuffing" is no longer used in the old sense, but is now used to mean

"extinguishing."

Then there was a French chemist, Michel Eugene Chevreul (1786-1889—yes, he lived to be 103), who worked with fats, found they were glyceryl esters of fatty acids, and isolated the fatty acids. In 1825, he took out a patent on the manufacture of candles made out of these fatty acids. They were harder than tallow candles, less greasy, gave a brighter light, needed less care, and didn't smell bad.

It is because candles were so improved that we can still use them today for show (not for light). You'll find candles at every banquet and at almost every restaurant table, doing nothing but lending "atmosphere." I keep thinking that if Cambacere and Chevreul had minded their own business and if candles still needed snuffing and still stank, they wouldn't be there, and there'd be one fire hazard the less.

Lamps were also improved in modern times. A Swiss physicist, Aime Argand (1755-1803), about 1783, invented a lamp with a glass chimney (the familiar lamp of nineteenth-century rural America) and a device for introducing a current of air through the lamp that resulted in a brighter light and less smoke.

Then there was the Austrian chemist Karl Auer von Welsbach (1858-1929), who thought that a lamp light

214

might be even brighter if the flame would heat some chemical that would then glow with a brilliant white light. He tried many substances that might glow at high heat without melting and then finally found what he wanted.

If he impregnated a cylindrical fabric with thorium nitrate, to which was added a small percentage of cerium nitrate, he got a brilliant white glow. This Welsbach mantle was patented in 1885 and produced the best oil lamps yet seen.

Now let us backtrack a little.

Plant life, as I explained in the previous essay, could slowly, under pressure, and in the relative absence of oxygen lose what oxygen and hydrogen it had and turn into coal, which is mostly carbon.

Animal life, too, can undergo changes. The fat droplets from innumerable one-celled organisms can lose what little oxygen they have and become a complex mixture of hydrocarbons, compounds

whose molecules are made up of carbon and hydrogen atoms only. This mixture is called petroleum, from Greek words meaning "rock oil," because it is a liquid fuel that comes from the rocky ground, rather than from plants or animals, (Of course, it came from animals originally, but those who named it didn't know that.)

Generally, petroleum deposits exist underground, where they are slowly formed, but the viscissitudes of geological change sometimes bring them fairly close to the surface or even right up to it. In that case, the smaller molecules, which evaporate easily, do evaporate and vanish, leaving behind a tarry residue made up of larger molecules.

This residue is most commonly found in those places which are the richest in underground reservoirs of pe-

215

troleum—in the Middle East. What we now call Iraq and Iran (but which ancient Greeks called Mesopotamia and Persia) were the richest.

The residue has received a variety of names. It might be called asphalt, for instance (a word of uncertain origin). There is enough asphalt about the Dead Sea for the Jewish historian Flavius Josephus (A.D. 37-100) to call the sea, in Latin, Lacus Asphaltites ("Lake Asphalt").

Asphalt might also be called bitumen, or slime, or, most commonly, pitch.

The ancients who lived in the Middle East found uses for pitch. It was sticky; it wouldn't mix with water; and it wouldn't allow water to soak through. If pitch were smeared on wooden objects, and if it filled the cracks between them, it would make them waterproof. Hence, it was of great use in shipbuilding.

Thus, when God directs Noah to build the ark, he says, "Make thee an ark of gopher wood; rooms shalt thou make in the ark, and shalt pitch it within and without with pitch" (Genesis 6:14).

Bitumen could also be used as a mortar to hold bricks together. Thus, when the builders of the tower of Babel got to work "they had bricks for stone, and slime had they for mortar" (Genesis 11:3).

When there was a battle in the vale of Siddim near the Dead Sea, the Bible remarks that "the vale of Siddim was full of slimepits" (Genesis 14:10).

What's more, there was no question that the ancients knew that bitumen would burn, for Isaiah, when he wanted to describe how miserable the world situation would be if God got slightly annoyed with humanity, said: "And the streams thereof shall be turned into pitch, and the dust thereof into brimstone and the land thereof shall become burning pitch" (Isaiah 34:10). The most interesting mention, however, is in connec-

216

tion with Moses' babyhood. As the Hebrew boy babies were being killed, Moses' mother, to save him, "took for him an ark of bulrushes, and daubed it with slime and with pitch, and put the child therein" (Exodus 2:3).

This makes sense, for an "ark of bulrushes" would be a little boat made out of papyrus reeds, which is just the sort of thing an Egyptian would make. The pitch would be added to make it watertight.

The catch is that there was no pitch in Egypt to speak of. The Egyptians only started using it in later days when they imported it from Mesopotamia. Whatever they used to make their boats waterproof, it wasn't pitch.

Why, then, does the tale of the ark of bulrushes talk about pitch? . . . Because it is a borrowing from another story.

Sargon of Agade, a Mesopotamian conquerer who lived perhaps twelve centuries before the time of Moses, was the kind of hero concerning whom later storytellers invented legends, and a favorite legend for any hero would deal with how the hero escaped death as a baby. The Greeks told such stories of baby-escape about Perseus, Oedipus, and Hercules. The Romans told it of Romulus and Remus. The Israelites told it of Abraham as well as of Moses. The Christians told it of Jesus. But Sargon of Agade, as far as we know, was the first.

In order to save him from death, Sargon was placed in a little boat in the Euphrates river and he was saved by a gardener. Undoubtedly, Sargon's boat was well-coated with pitch.

The story was borrowed by the biblical legend-makers and used for Moses, and they borrowed the pitch, too.

There would, of course, be places where petroleum seepage produced small-molecule fractions that were in the process of evaporating, but were seeping upward; more or less as rapidly as they evaporated. In

that case,

217

there would always be vapors present that, if present in enough concentration, would be inflammable.

I imagine that, every once in a while, someone would start a campfire near one of those places and, if conditions were right, there might be a flash of light and then a flame flickering along the ground in some particular spot.

Anyone involved in such a thing would want to hurry away, I suppose. If sufficient curiosity were aroused, though, he might watch from what he thought was a safe distance and, if so, he might note that the flame seemed to have no intention of going out, and didn't seem to consume fuel in the ordinary way.

Moses is supposed to have seen something like that. "And the angel of the Lord appeared unto him out of the midst of a bush: and he looked, and, behold, the bush burned with fire, and the bush was not consumed" (Exodus 3:2).

Such unconsuming "eternal fires" may have stirred the religious feelings of some. Even ordinary fires were, hi a way, mysterious things that clearly brought great good to humanity and offered dangers, too. It would not be unusual for some primitive people to attribute divine qualities and powers to fire. The Zoroastrian Persians did so and are sometimes referred to as fire-worshipers as a result.

On the other hand, some may have been frightened by these fires from the ground and thought them the work of demons. Such fires, and the experience of volcanoes, may have helped convince people of an underground of eternal fire, thus giving rise to the legendary existence of a Hell in which the spirits of the dead were tormented.

In the places where petroleum seeped upward, a liquid might be obtained which burned. As a fuel, it would seem very much like ordinary oil from plants and ani-

218

mals. The Persians called this burning liquid neft, which have meant "liquid." The Greeks picked up the and called it naphtha. '• Naphtha is mentioned in two places in the Apocrypha. The Book of Daniel tells of three young men, Sha-drach, Meshach, and Abednego, who were thrown into a fiery furnace for defying Nebuchadnezzar's religious

views, but who were saved by a divine miracle. In the apocryphal book "The Song of the Three Young Men," it says in verse 23: "Now the king's servants who threw them in did not cease feeding the furnace fires with naphtha, pitch, tow, and brush."

In the book of 2 Maccabees, written some time in the first century B.C., the tale is told of the building of the second Temple, five centuries earlier, after the Persians permitted some Jews to return to Jerusalem. There would naturally be a search for some relic of the First Temple that would represent a continuation of sanctity for the Second. In particular (says the story), they were looking for some holy fire that might have been preserved by pious men, or by a divine miracle, for the seventy years or so that had elapsed since the destruction of the First Temple. However, "they had not found fire but thick liquid" (2 Maccabees 1:20). They sprinkled this liquid on the wood on which materials for a sacrifice had been laid and "a great fire blazed up, so that all marveled" (2 Maccabees 1:22).

This mysterious liquid, according to 2 Maccabees 1:36, was called nephthar, the meaning of which was given as "purification," but the verse goes on to say "by most people it is called 'naphtha.' "

Pitch was not found in the Middle East only. There were petroleum seepages reported in various parts of

219

Europe, and once Europeans discovered the Americas, seepages were found there, too.

In March 1595, Walter Raleigh (1552-1618) visited the island of Trinidad, where he was the first European to see Pitch Lake, which is a lake consisting of about 10 million tons of asphalt.

People valued such pitch, for new uses were found for them. Asphalt was used for paving roads, the softer portions of pitch were used as a liniment. Clear oil obtained from it (mineral oil) was used as a laxative. The thicker portions, when they burned, produced a foul-smelling smoke that was used to fumigate houses.

In the nineteenth century, inflammable liquids were sought for use in lamps, liquids that might be cheaper and hi more dependable supply than whale oil. Coal was heated to yield coal oil, for instance. It was also possible to heat and obtain oil out of asphalt from Trinidad, or out of certain kinds of rocks called shale that seemed to be impregnated with oily material (hence it was called oil shale).

In 1853, a British physician, Abraham Gesner (1797-1864), developed a process that would yield an inflammable liquid from asphalt. Because it was driven out of a waxy mixture of solid hydrocarbons, Gesner called the liquid kerosene from a Greek word for "wax." The British call it paraffin these days, but in the United States it is still called kerosene. Kerosene was ideal for lamps (and nowadays when we think of oil lamps, we think of them as kerosene lamps, as though that were a single word). The trouble was, though, that even with Gesner's process there wasn't enough kerosene to meet the great demands the lamps of Europe and America represented.

The short supply was bound to continue as long as people dealt with petroleum that had reached the surface, been exposed to open air, and had dried out. The

220

Ikerosene fraction was vaporized and gone and only lall amounts could be squeezed out of the pitch that

^remained.

But what if one could dig down and come across the

I petroleum before any of it had evaporated, when it

; might be rich in the small-molecule fractions that would include kerosene? In that case, the liquid petroleum might be heated and made to give off kerosene in enormous quantities.

This notion of digging for liquid was a very old one. After all, one digs down to the water table and has a well, which will yield cold, fresh water at all times.

As long as two thousand years ago, people in China and Burma were digging not for fresh water, but for brine. This they would heat to obtain salt for use in preserving food and for other purposes. Apparently, every once in a while they brought up petroleum, too. They had no direct use for this, but they didn't throw it away, either. They would collect it and use it as a fuel for a flame that would drive the water away from the brine, leaving salt behind.

We now switch to a railway conductor named Edwin Laurenline Drake (1819-80). He had been born in New York State, and he worked in New Haven, Connecticut.

As a matter of investment, he had bought some stock in the

Pennsylvania Rock Oil Company. (Remember that "rock oil" is English for petroleum.) The company made its money by collecting petroleum that had seeped up to the surface near Titusville, Pennsylvania, and selling it for medicinal purposes. Titusville is in the northwestern part of the state, about ninety miles north of Pittsburgh.

There was enough petroleum seepage for medical use, but not enough to satisfy the lamps of the nation.

221

Drake, in view of his investment, would have liked a lot of petroleum and a lot of sales to lamp owners.

As it happened, he knew about the Chinese drilling for brine and their habit of occasionally bringing up petroleum, so he studied the methods for such drilling. Then, in 1858, he persuaded the company to lease him some land on which he might start drilling operations.

He started drilling and, on August 18, 1859, having drilled down for 691/2 feet, he struck oil. It was the first oil well to have been drilled into the surface of the Earth.

Once Drake succeeded, others flocked to the spot and began drilling for oil on their own. Northwestern Pennsylvania became the first oil field in the world, and boom towns sprang up. Drake hadn't patented his methods, however, and he wasn't a clever businessman, so he didn't become rich. In fact, he died a poor man.

However, people continued to drill, and not just in Pennsylvania, either. Before the 1800s were finished, there were oil wells in fourteen states, from New York in the East to California in the West; from Wyoming in the North, to Texas in the South. Oil wells were dug overseas, too, in Baku in the Caucasus, for instance.

The petroleum was refined and used as a source for kerosene chiefly, and the fifty years between 1860 and 1910 were the golden age of the kerosene lamp. With the glass chimneys, and the wicks, and the air currents, and soon the Welsbach mantles in addition, the lamps lit up homes as they had never been lit up before.

Kerosene put whale oil out of business and removed that reason, at least, for killing the magnificent cetaceans. (Unfortunately, other reasons cropped up.) What's more, there seemed enough petroleum in the ground to supply kerosene for lighting for many centuries.

However, something happened, and it was called the

222

^ electric light (see "To The Top," F & SF, September ,1976).* In 1879, Thomas Alva Edison (1847-1931) invented a practical electric light, and designed the kind ,of generating station that could keep lights burning steadily even as some were turned on and others were turned off. It was the greatest invention of the greatest inventor we know by name.

The electric light did not sweep the world instantly. Generating stations had to be built, cables and wires had to be laid, electric light fixtures had to be installed. What's more, the first lightbulbs didn't last long and were unpleasant to look at, with their bare filaments. The bulbs had to be improved by filling them with nitrogen, rather than with vacuum; by frosting the glass rather than leaving it clear; by substituting tungsten filaments for carbon ones, and so on.

Still, as early as October 10, 1881, the Gilbert and Sullivan comic opera Patience moved to a new theater, the Savoy, the first theater to be equipped with electric lighting. When the next play, lolanthe, opened on November 25, 1882, the chorus of fairies had their wands tipped with electric lights, which made a great sensation.

It was not till the time of World War I, however, that electric lights had won their victory, leaving the kerosene lamps to become a charming antique and nothing more. (Unlike candles, they are not even used for ceremonial reasons.)

Nevertheless, as I mentioned at the start of the previous essay, even as late as 1925,1 was living in a Brooklyn apartment that did not have electric lights.

You would have thought that, with the passing of the kerosene lamp and the steady dwindling of the need for kerosene, the petroleum industry, having had its short-

* See my book Quasar, Quasar, Burning Bright (Doubleday, 1978).

223

lived boom, would now dwindle and pine and become as antique as the lamps themselves.

Not a bit of it. The industry continued to grow, and became an enormous giant.

We'll continue with the subject, therefore, in the next essay.

15

The Horse Under the Hood

224

On September 18,1957,1 received a letter from the late Robert P. Mills, who was then editor of The Magazine of Fantasy and Science Fiction and its sister magazine, Venture Science Fiction. He wanted to know whether I would be willing to write a short science column for Venture.

Yes, I would. Of course!

For several years, I had been writing occasional science articles for Astounding Science Fiction and I had enjoyed it, but I hated having to get approval for each article first and then having to risk a rejection. (I'm funny that way.) Bob offered to let me have a free hand as long as I didn't miss a deadline.

Good! I promptly wrote an article and it appeared in the January 1958 issue of Venture. That was the seventh issue of the magazine. I then wrote three more articles

225

which appeared in the eighth, ninth, and tenth issues of the magazine. But with the tenth, the July 1958 issue, the magazine ceased publication.

My days as a science columnist had ended so quickly, and just as I was getting into the swing of it, too. I was chagrined.

Then, on August 12, 1958, I had lunch with Bob in New York, and he asked me if I would continue the column after all, but for F & SF.

I agreed at once without bothering to conceal my glee. I was back in business on the same terms as before —an absolutely free hand provided I did not miss a deadline.

The magazine and I both held to the agreement. My first column appeared in the November 1958 F & SF, and I went on, issue after issue since then. The magazine never objected, not even once, to my choice of subject, nor offered to change a word, and I never missed a deadline regardless of family crises, bouts of ill health, or whatever.

In any case, the point of all this is that this essay that you are now reading is my three hundred and sixtieth. With the October 1988 issue, I complete thirty years of my column. Next month is the column's thirtieth anniversary and with it I will start my thirty-first year.

I just thought I'd mention it. ... And to tell you that some of you out there may be tired of the column, but I'm not. I'm shooting for another thirty years, so here goes . . .

The horse has been a servant to man ever since about 2000 B.C., when the nomads of central Asia tamed it.

In some ways, it is an ideal animal. For a combination of speed and strength, there is nothing like a horse. Anything bigger and stronger, such as a rhinoceros, is

226

slower; anything faster, such as an antelope, is smaller and weaker.

On the other hand, anatomically it leaves something to be desired, if it is compared with the ox, which is the most useful prehorse animal when it came to work. The ox is stupid, placid, uncomplaining, strong, and has huge hulking shoulders with which to push. ... It is also terribly slow.

A horse is built otherwise. Its shoulders are narrow and if it must pull, there has to be a broad strip of hide crossing the horse's chest. Under those circumstances, when a horse pulls hard, it succeeds in closing its windpipe. So a horse doesn't pull hard—neither would you in its place.

But the central Asian nomads reduced the job of pulling to a minimum. They devised a chariot that was little more than a platform on an axle between two large spoked wheels. Two men stood on the platform, one to control the horse, and one to handle the weapons.

In the centuries after 2000 B.C., the charioteers swept down upon the settled civilizations of the time, and defeated them all, from China to Egypt. There was no standing up to the charging chariots until the conquered peoples learned to use the horse themselves.

In the next three thousand years and more, the horse remained an indispensable adjunct of the aristocracy. And there were many improvements, too. The chariot went out of fashion once horses were

bred that were sufficiently large and strong to bear the weight of a heavy man, and still be able to run fleetly.

Stirrups were invented so that a rider could stick his boots into them and sit firmly. That meant he could thrust with his spear without pushing himself off the back of the horse.

Horseshoes were invented, which protected the deli-

227

cate hooves of a horse and kept him from turning up lame every other day.

About 1000 A.D., the horse collar was invented, which gave the horse a pair of artificial shoulders to push with, so that for the first time he could pull with full strength. That made him into a superior work animal about the farm. He was the ideal animal to pull improved plows so that the food supply in northwestern Europe was increased manyfold.

Eventually, coaches were devised that made it possible for people without horses to travel at horse speed according to set schedules, but the coaches were pulled by horses, of course. There were carriages-for-hire, reapers, omnibuses—all pulled by horses.

All the way down to the closing decades of the nineteenth century, a galloping horse was as fast as a man could go overland, and life without horses seemed unthinkable.

That didn't mean that people didn't dream of impossible improvements. The winged horse Pegasus is the most charming creature in Greek mythology. In theAra-bian Nights, there is an object that flew by turning a peg, but it is in the shape of a wooden horse. Of course, the Greek myths have Daedalus frying on artificial wings, and the Arabian Nights has flying carpets, too.

But in 1769, Watt's steam engine came into being and, for the first time, human beings didn't need fantasies. They had a reasonably efficient way of drawing upon the inanimate energy of burning fuel. By 1781, Watt had improved his device to make it possible for it to bring about rotary motion, and, by turning wheels, it could power mill machinery—and transportation devices.

In 1807, the first commercially viable steamship came

228-

into being, and, in 1825, the first commercially viable steam locomotive.

The steamships were fine, but steam locomotives were clearly lacking in versatility as land transportation. The locomotive required rails and could only travel on those rails. Furthermore, it was only economic as a large device carrying many goods or many people.

Was there no way to personalize the locomotive? Could it not be made small to accommodate an individual, or a small family? Could it not be independent of rails, so that it could go anywhere an ordinary road would take it?

In short, what was wanted, was a private carriage, which a wealthy man could own, or a commercial carriage for hire, which a man of moderate means could use—but without a horse. A horseless carriage, in other words.

Even before Watt's steam engine, people had dreamed of horseless carriages. They thought of them as powered by sails (but you would then have to depend on the fickle wind). There was also thought of clockwork devices (which you would then have to wind up with considerable effort).

Steam did away with all that. Once Watt's steam engine came into being, people thought only of steam carnages. Steam carriages were indeed built, and some of them indeed worked, but there were enormous problems.

They tended to be heavy. No matter how you skimped on the carriage itself, a steam carriage had to carry a big, strong boiler. What's more, the boiler had to be fed fuel, so usually a steam carriage had to have a platform behind for the stoker, who would keep feeding the fire. The water would boil away steadily and you would have to stop frequently to fill up on additional water. What's more, you couldn't start till you had heated the water to

229

boiling and worked up a head of steam, and if you've ever waited for a trivial kettle of water to boil so that you can have a cup of tea, you know that waiting for water to boil can be tedious. And once you did start, the steam carriage was likely to lumber along like a laden

ox.

Nor did the various horse-related industries, the coach owners, for instance, sit idly by. They claimed that the vehicles would scare the horses and that was a powerful argument. No one wanted to ride on a panicky runaway horse, or be in a coach pulled by some.

Even the public was hostile. The steam carriages tended to tear up the roads, and fill the air with noise and steam. In Great Britain, so hostile was Parliament that, in 1865, they passed a red-flag law that kept all steam carriages to a top speed of four miles an hour in rural areas (the speed of a brisk walk) and two miles an hour in towns. What's more, someone with a red flag had to walk along in front of the steam carriage so that people would be warned of its approach. The law wasn't repealed till 1896.

Even so, inventors worked doggedly to make steam carriages more efficient and commercial. By 1900, there were flash boilers that allowed one to build a head of steam quickly. The machines were made lighter, simpler, and faster.

Two brothers, Francis Edgar Stanley (1849-1918) and Freelan O. Stanley (1849-1940), began to manufacture steam carriages in 1897, producing the famous Stanley Steamer. In 1906, they produced a steam carriage that broke the world record for speed. It went a mile in 18.2 seconds, which is equivalent to a speed of 127 miles per hour. However, the steam carriage was overtaken by events. Something better had come along.

*   *   * 230

The steam engine is an external-combustion engine. That is, the fuel is burned outside the engine to produce steam and the steam then enters the engine where its pressure moves the piston.

Naturally, it occurred to some people that matters would be improved if the fuel were somehow burned inside the cylinder housing the piston, so that the energy of the chemical combination could move the piston directly. That would be an internal-combustion engine.

With an internal-combustion engine, there wouldn't have to be a large water boiler within which to make steam. There would be no heat lost in bringing the boiler and water to the steaming point. Furthermore, a vehicle would start instantly when fuel was burning in the cylinder; there would be no need to wait for the water to boil and a head of steam to build up.

But what would the fuel be? Obviously, you can't stick slivers of wood or bits of coal into the cylinder. You need inflammable vapors that

will mix with air, and explode readily when detonated with, say, an electric spark. That means gases—or, possibly, liquids that are easily evaporated and give off gases at ordinary temperatures.

As early as 1820, someone built an engine intended to work on exploding mixtures of nydrogen and oxygen, but it wasn't commercially useful.

The first internal-combustion engine that could be viewed as even remotely practical was built in 1859 by a Belgian-French inventor, Jean Joseph Etienne Lenoir (1822-1900). He used illuminating gas as a fuel, the kind of gas that in those days was obtained by heating coal in the absence of air, forcing inflammable vapors to be given off. In 1860, Lenoir inserted his engine in a small conveyance and putt-putted it around the streets. This was the first motor carriage (as distinguished from a steam carriage) or, in briefer form, the first motorcar.

231

The engine was very primitive and inefficient, however, making use of only about 4 percent of the burning fuel. Still, in the course of five years, Lenoir sold three hundred of his engines.

The piston in the Lenoir engine was a two-stroke device, in and out, but, in 1862, a French engineer, Al-phonse Eugene Beau de Rochas (1815-1893), pointed out that much greater efficiency would result from the use of a four-stroke device.

1. The piston would push outward, creating a partial vacuum and sucking in a mixture of inflammable vapor

and air.

2. The piston would move inward, compressing the mixture.

3. Ignition at maximum compression would explode the mixture and drive the piston outward. That would be the power stroke because it is the one that delivers the impulse that turns the wheels.

4. The piston would move inward again, expelling the products of combustion.

After that, the piston moves outward again, sucking in a new mixture of vapors and air, and the cycle proceeds over again—and again—indefinitely.

Describing this in a theoretical way is one thing. Actually building a device that incorporates these ideas and making it ail work in a useful and practical way is quite another. Beau de Rochas didn't try to put his ideas into practice. Nor did anyone else over the next fourteen years.

In 1876, however, a German inventor, Nikolaus August Otto (1832-91), built one that actually worked. As a result, the four-stroke cycle is sometimes called the Otto cycle, and internal-combustion engines making use of the Otto cycle are sometimes called Otto engines. Otto patented his engine in 1877, and formed a company that sold thirty-five thousand such engines in a few

232

years. It was clearly the best internal-combustion engine that had been designed and, by 1890, it was the only one.

Next came the matter of building a vehicle with an Otto engine that would run more efficiently than Le-noir's vehicle.

The first to do this was a German mechanical engineer, Carl Friedrich Benz (1844-1929). He mounted the engine in the back of something that looked very much like a buggy. It had three bicycle wheels, a small one front center, and two large ones on either side in the back.

Not only did Benz make use of an Otto engine, but his fuel was gasoline and that is worth a small digression.

Kerosene and gasoline are both obtained from petroleum. Kerosene is made up of hydrocarbons with ten to twelve carbon atoms per molecule. Gasoline is made up of smaller molecules containing only four to eight carbon atoms.

This means that gasoline has a lower boiling point than kerosene does, and vaporizes much more easily. In fact, it is because it gives off vapors of inflammable gas so readily that it is called gasoline. Of course, that means it is inevitably abbreviated as gas, which it isn't. It is a volatile liquid. The word came into use in the 1870s.

The French call gasoline essence de petrol, meaning "extract of petroleum," which it is, but then they abbreviate it to essence, which seems foolish. The British, just to be contrary, borrowed the French expression and abbreviated it to petrol.

Whatever you call it, though, gasoline is too vaporous and too ready to explode to be used in a lamp. You need the more decorous and quiet

kerosene for the purpose.

233

On the other hand, kerosene wouldn't work in an Otto engine, for it doesn't give off enough vapors. There we want gasoline.

And so it came about that just as the electric light was killing the kerosene lamp, and it looked as though petroleum would become a drug on the market, the coming of gasoline-powered motorcars gave petroleum a new lease on life. A new and better lease on life, for more gasoline was gobbled up, by far, in the new cars, than lamps could consume kerosene. The entire process of petroleum refinement switched from converting as much of it as possible into kerosene, to converting as much of it as possible into gasoline.

That answers the question with which I concluded the previous essay, as to how the petroleum industry could survive the decline of the kerosene lamp. However, as long as we're on the subject of motorcars, let's continue . . ,

Benz built his first three-wheeler in early 1885. It was a gas buggy (an American slang term for the motorcar) almost literally. Also, since it had an Otto engine and used gasoline as fuel, it was the first representative of what we today call an automobile.

The word came into use just before the time when Benz's device was built, and it is an uncomfortable one. Auto means "self" and mobile means "moving," so automobile means "self-moving" (no horse, that is), and that surely sounds like a great description. The trouble is, however, that auto is from Greek and mobile is from Latin, and mixing the two languages in this fashion is a no-no for linguistic purists.

In proper Greek, the device should be an autokinesis, and in proper Latin, it would be an ipsemobile. The chance, however, of doing anything about this is pre-

234

cisely zero. Automobile it is, and automobile it will stay, and Benz was its inventor,

Benz ran his first automobile around a cinder track right next to his factory. He made four laps before something broke, and he only stalled twice. His wife and his workmen ran around the track with the automobile in wild excitement.

Benz made his first public run in the autumn of 1885, and either forgot how to steer, or had trouble doing so, for he ran into a wall. He made his first sale in 1887, flourished, and, in 1890, began to manufacture four-wheelers.

Second in the field was another German inventor, Gottlieb Wilhelm Daimler (1834-1900). Daimler had! worked with Otto at first, but left him in 1883 because he found Otto too conservative in his outlook.

Daimler constructed a high-speed engine, making id lighter and more efficient, and he also used gasoline as a fuel. He fitted such an engine to a boat in 1883 and had the first motorboat. In 1885, he fitted an engine to a bicycle and had the first motorcycle.

He built his first automobile in 1887, which put him two years behind Benz, but his automobile was a four wheeler at the start, which put him three years ahead ol Benz in that respect. What's more, his automobile wal the first to have the engine in front, and the horse, so tJ speak, was finally under the hood.

In the United States, an inventor, George Baldwii Selden (1846-1922), claimed priority because he ha 'obtained a patent for an automobile design as early a 1879. However, all he had was the design. He didiJ build an automobile. The first American gasoline-pofl ered automobile actually built was devised by Chark Edgar Duryea (1861-1938). Duryea drove his car on tfi streets of Springfield, Illinois, on September 22, 1893|

235

To begin with, the automobile was an expensive toy that might easily have developed into something meant exclusively for the amusement of the world's rich (like yachts). Some, however, made efforts to produce automobiles cheaply.

One step in that direction was the establishment of part-interchangeability, of making every part so exactly to specification that any part could be used in any automobile. This was a practice used in other industries, but the American engineer Henry Martyn Leland (1843-1932) was the first to apply it successfully to automobiles.

He built the first Cadillac in 1903; in 1908, he put on a show in which three Cadillacs were disassembled, the parts mixed up, and some replaced from dealers' stocks, and then three Cadillacs were assembled out of the mess and were driven five hundred miles

without trouble.

A revolution, however, came with the American engineer Henry Ford (1863-1947). He built his first automobile in 1893, and founded a company for the manufacture of automobiles of his own design. He was intent on making cars cheaply and he tried eight designs which he labeled by various letters: Model A, Model B, and so on. The eighth he called Model S. Those models which were cheaper sold better.

In 1908, then, Ford got the idea of the assembly line. The parts moved along a belt and went to the workmen, rather than vice versa. Each man in line did one job and the product was then passed on to the next man who did another job, and so on. At the end of the assembly line, a finished car rolled off onto the floor.

Ford used the assembly line to manufacture his ninth

236

model, which he labeled the Model T, and that was, by all odds, the most famous automobile in history.

It cost only $950 to start with, but the prices steadily dropped until, in 1926, it was only $290. (Of course, that was in 1926 dollars, which had much more buying power than our own feeble item of 1988, but it was still cheap.)

The Model T was the first car available to the middle classes, but there was still one thing about it that kept it from being truly a vehicle for everyone. In order to start it, it had to be cranked. The engine had to be given a good hard turn in order for it to catch and, therefore, keep going on indefinitely.

I've never cranked a car myself (well, I'm not that old) but, in my imagination, I can see exactly how it went. You got the crank, went out to the front of the car, stuck it into the little hole under the radiator, and felt it grip a projection which it would turn and which would, in turn, turn the engine.

You then spat on your hands, got a firm grip on the crank, and pushed it down with all your might and as sharply as you could. The engine would cough once or twice and die. Your lips would set more grimly and you would repeat the process and get another double cough. You might have to do it a half dozen more times, getting sweatier and angrier and cursing more and more freely —and then, finally, it would catch and you raced quickly into the car and put it in gear so you could get going before it died again.

To make it a little worse, the time might come when the engine caught unexpectedly—when you weren't set for it and your grip wasn't quite firm enough, or you were off balance. When that happened, the crank would manage to yank itself out of your hand, turn with the engine, coming around to the other side of your forearm, giving it a sharp blow and possibly breaking one or

237

both bones. I imagine that would be less fun than almost anything else about a car.

In any case, as long as a car had to be cranked, the job fell to the strongest person in the family—usually to the lord and master—and women and half-grown youngsters were out of it.

It didn't last, of course. The American engineer Charles Francis Kettering (1876-1958) invented the electric self-starter, a device in which an electrically powered clamp gripped a projection of the engine and turned it like a relentless arm that could twist harder and longer than a human arm possibly could. . . . The motor would catch while all you did personally was to twist a key in the dashboard into position to make contact and close a circuit.

The self-starter appeared first in the 1912 Cadillac, but it spread gradually to all the cars, even down to the cheapest, and in the course of the 1920s the crank disappeared and is now scarcely remembered.

With the self-starter, the automobile was driven as easily by women as by men, and by adolescents as by adults. The day of the virtual universality of the automobile had finally come.

The automobile changed American society from top to bottom. It gave rise to a nation on wheels. It offered the possibility of a home in the suburbs, for one was no longer necessarily enslaved to a house near the office or factory. It made it possible to take a vacation someplace farther away than the vacationer's backyard. It helped disintegrate the family, for it was easier for children, when grown up, to find jobs at a distance (but to return for a reunion, too). It meant greater freedom for teenagers, who could escape parental supervision by car and use it for sexual experimentation.

It produced a network of paved roads and gas stations and garages and built an industry on which the nation's

238

prosperity depended and was maintained. And it also introduced us to traffic congestion, to air pollution, and to the killing and maiming of Americans by the tens and hundreds of thousands per year. , . . , t But whatever the difficulties, we can t give it up. Just to point up the universality of the automobile with a specific case, even /, a person who doesn t know which end of a hammer you saw a plank with, learned ho* to drive in 1950, and of all the modes of travel, I find that a journey by car, with my own hands on the wheel, is by far the least unpleasant.

239

16

The Unforgiving Minute

When I was young, I encountered, as most avidly reading youngsters did, inspirational writings of many kinds. I did not fail, for instance, to come across "If—," written in 1910 by Rudyard Kipling (1865-1936).

I read it with cynicism, I'm afraid. Young as I was on the day I stumbled across it, I knew that I couldn't live by its precepts. I doubted that anyone could.

There were the lines that went: "If you can meet with Triumph and Disaster / And treat those two imposters just the same . . ."

I knew I wouldn't. I knew I would jump up and down and wave my arms with glee in case of Triumph. I knew even more firmly that I would skulk in a corner and be very sorry for myself in case of Disaster.

What's more, I thought, even as a child, that anyone

243

who would make "one heap of all your winnings / And risk it on one turn of pitch-and-toss" was a jackass.

There was one bit that got me, however; that I kept repeating to myself over and over. It was the following:

If you can fill the unforgiving minute

With sixty seconds' worth of distance run, Yours is the Earth and everything that's in it, And—which is more—you'll be a Man, my son!

I won't say that those lines centrally guided my life because there were a number of other factors that made me keep the old nose to the grindstone day after day and year after year, but if, at any time, I thought, "Well, why not take it easy?" it was Kipling's "unforgiving minute" that popped into my mind, the minute that would never forgive being wasted and would never return, and it was that which turned me back and forced me to give it my sixty seconds' worth of distance run.

So it came about that once an interviewer asked me if I had a fixed routine before starting work.

"What do you mean, a fixed routine?" I asked, puzzled.

"Well, do you start out by sharpening pencils, or by looking out the window, or by doing deep knee bends, or anything else that would serve to get you into the mood of writing?"

"Oh, that," I said, "Sure! I have something I never fail to do before I start working." "Good! Tell me what it is!"

"The first thing I do," I said, "is get close enough to the typewriter for my fingers to reach the keys." So let's talk about the unforgiving minute.

*   *   * 244

The time units that forced themselves on human beings to begin with were the three that depended on astronomical facts: the day, the month, and the year. I have devoted essays to each of these three natural units of time, the most recent being "Time Is Out of Joint" (F & SF, February 1986), which dealt with the day.

Even the day, which is the shortest of the three, is quite long and it was unavoidable* that human beings would divide it into smaller portions: dawn, sunrise, morning, noon, afternoon, sunset, twilight, and, of course, night.

These are not precise divisions, but they are sufficient for many purposes.

There are occasions, however, when you might want something more precise. You might want to make sure you finished a job before the heat of the afternoon made you stop, or that you began a journey to the next town with the secure knowledge you would not be overtaken by nightfall. For those reasons, you might want a pretty close idea as

to just what time of day it was.

We don't know who first thought of following the shadow of a stick as it crept along the ground in response to the fact that the Sun was making its way across the sky. Such sundials, however, came into use in early civilized time in Egypt and the path of the shadow was divided into twelve equal periods.

Why twelve? Probably because the astronomic periods suggested the number. After all, there are about 12 months to the year, about 60 days (12x5) in two months, and about 360 days (12 x 30) in a year.

Why those numbers? Early civilized humanity had plenty of trouble handling fractions and it so happened that 12 could be divided evenly by 2, 3, 4, and 6—no fractions. No other number close to that size could be divided evenly by as many as four different factors.

As for 60, that could be divided evenly by 2, 3, 4,5, 6,

245

10, 12, 15, 20, and 30; while 360 could be divided evenly by 2, 3, 4, 5, 6, 8,10,12,15,18, 20, 24, 30, 36, 45, 60, 72, 90, 120, and 180. These were unique numbers made to be easily handled—as anyone could plainly see—by the all-wise gods.

So the Sumerians divided the circle into 360 equal parts (which we call degrees from the Latin word meaning "to step down"). Each degree was divided into 60 equal parts, and each of those parts into 60 still smaller equal parts. The first set was called, in Latin, pars minuta prima ("first small part") and the next set was called para minuta secunda ("second small part"). These phrases were shortened to minute and second, respectively.

Once the day was divided into twelve hours of daytime and twelve hours of night, it seemed natural to divide each hour into sixty minutes and each minute into sixty seconds. That is how the unforgiving minute got its start and why each one had to have its sixty seconds' worth of distance run.

Naturally, minutes and seconds of time were just mathematicians' devices at first; you couldn't actually measure them. Sundials only sufficed to give you an estimate of rather sizable fractions of hours. Furthermore, sundials only worked during the daytime, and only when the sky was not clouded over.

Could there be some way of measuring time on cloudy days or by

night making use of some device that could be checked against the sundial when that was possible?

What was needed for the purpose was some natural process that took place at a fixed speed over an extended period of time, and to standardize just how much of the process took place in exactly one hour by the sundial. You would then have a dock. (The word is from the word for "bell" in most European languages,

246

including medieval Latin, since the passage of each hour would be announced by the tolling of a bell.)

Thus, you could keep time by the burning of a candle made in a fixed size of fixed material, or by having dry sand drift from an upper chamber into a lower one through a narrow orifice. Such devices could work day and night, cloudy or clear, and they would be portable

besides.

You could continue the timekeeping by substituting a new candle as the old one burned out, or by turning the sand clock over when all the sand had drifted out of the upper chamber. Still, these devices weren't very good. Different candles were oound to burn at different rates and even the same candle burned more rapidly or less rapidly depending on such things as air currents. As fon sandglasses, the sand drifted through the Orifice mord rapidly when there was a weight of much sand above ij than when there was little sand there.

Perhaps the best clock the ancients had was the clep\ sydra, in which it was water that dropped from an uppe chamber to a lower one. The word clepsydra is from thJ Greek, meaning "to steal water," because the watd seemed to be stolen slowly out of the upper chamber ra the lower. It is just as useful, however, to call it a watd

clock.

The earliest water clocks have been traced back d 1400 B.C. in ancient Egypt, but it was not until about id B.C. that a Greek engineer, Ctesibius, devised one wifl the obvious sources of error removed. He arranged fd a continuous flow of water into the upper chamber, WH an overflow. In this way the upper chamber always ha the same head of water and the rate of drip did n| change with time.

Eventually, water clocks were fitted with little floa that supported pointers that rose with the water level

247

the lower chamber. The pointer thus automatically indicated the number of each hour as it passed.

However good a water clock might be, the use of water was an inconvenience. There had to be a continuous water supply; the clock was not easily portable; and however careful one was, leaks or spills ensured that there would always be wetness about.

Yet clocks were needed to a slowly increasing extent. In the Middle Ages, monks and others in the religious life had to engage in prayers at set times for the sake of discipline. It is easy to see that those who had to say their prayers under such conditions might grow to feel their souls were in danger not only if they neglected to say them, but even if they were merely to say them at

the wrong time.

People in houses of worship therefore had to have clocks, and they got rid of water and its inconveniences by making use of gravity instead. They wrapped a cord around a drive shaft and suspended a heavy weight from it. The weight, as it was pulled downward by gravity, forced the drive shaft to turn, and a pointer attached to it marked off the hours on a dial. The trick was to so arrange the workings so that the pointer turned at a constant, slow speed that took it around the dial in twelve hours, or two complete turns in a day.

About 1300, something called an escapement was invented. This was a device with teeth that engaged the turning drive shaft and allowed it to move only so far. Then it disengaged and another tooth caught it. This helped the drive shaft turn slowly enough and constantly enough for the purpose,

Until medieval gravity clocks were invented, attempts were made to take into account the varying length of daytime as the seasons progressed, making the daytime

248

hours longer in summer and shorter in winter. With the gravity clocks, however, this was abandoned. The hours were made a fixed length all year long and it was agreed to let the Sun rise and set at different times by the clock through the year.

All clocks of ancient and medieval times, by the way, even at their very

best, could be counted on to end the day at least a quarter of an hour fast or slow. They would have to be adjusted manually at frequent intervals by checking them against sundials.

This is not intended as a sneer, of course. A loss or gain of a quarter of an hour a day represents an error of just about 1 percent. Considering the level of technology then available, I think this small error speaks highly indeed for the ingenuity and for the pains taken by the early timekeepers.

What's more, prior to about 1600 there was little need for greater accuracy where ordinary people, even clerics, were involved. There were certain specialized activities, however, that did require better timekeeping, and it is to these we now must turn.

Until 1581, the regular motions human beings used for their clocks were progressive. Candles always burned downward; sand, water, and weights always moved downward.

In 1581, however, the Italian scientist Galileo Galilei (1564-1642), who was only seventeen at the time, discovered a regular motion that could be under the control of human beings, and that went back and forth— that was periodic.

He was attending services at the Cathedral of Pisa, and he found himself watching a swinging chandelier that was shifting with air currents, now in a wide arc, now in a small one. It seemed to Galileo that whatever

249

SiSS^i^^4""1*OI

"e pendulum (fro* * tat uffl

^gmg"K e it was at once «^rf a?to* but ta pnnaple, rt^   t the mot.ons
    the

Galileo coul   "°\   ^enta sc   se

SSTrf nature-a -—^^toply ^ """

^^rltSa^Ff^S^

.^^&^^±^^^rfi

worked

transfer enough energy to the pendulum to keep it swinging indefinitely and for making the swinging pendulum control the escapement so that that became much more precise than before. (Of course, the descending weights had to be wound back to the top periodically just as they would have to be in the absence of a pendulum.)

Huygens'spend«/«m clock was the first timepiece that was accurate not to the hour but to the minute. For the first time, a timepiece could be profitably given another hand—a minute hand, making a complete circle while the hour hand advanced one hour.

The big disadvantage of the pendulum clock was its size. The pendulum had to be a yard long to beat out seconds and, in general, the pendulum clock was nonportable.

The English physicist Robert Hooke (1635-1703) had, however, begun to study springs in 1658 and had shown that they could oscillate with constant periods, even as pendulums did, and took up less room in so doing.

In 1675, then, Huygens worked out a miniature clock. In this a stiff mainspring gradually uncoiled, supplying a steady force that kept a much thinner hairspring oscillating steadily. The hairspring kept the escapement going and the clock thus produced was small enough to keep in a pocket.

Such a small clock was useful to sentries or other people who had to watch (that is, stay awake) during the night hours. The length of time they had to do so before being relieved was therefore a watch; and the instrument that told them when their watch was over and when relief should be coming was also a watch.

Watches, too, had to be rewound periodically to recoil the mainspring. (Nature simply won't give you something for nothing.)

251

Navigation represented another timekeeping problem. On the open seas, there were no roads, no landmarks, no one to ask the way. One had to determine latitude and longitude. For latitude, it was only necessary to measure the maximum height of the Sun in the course of the day. Longitude, however, depended on knowing the time difference between the moment of highest Sun at the home port and

the moment of highest Sun at the position at sea.

Prior to 1400, longitude didn't matter, for ships only made short voyages, hopping from shore to shore. Even if they missed their goal, they would be sure to reach some piece of land, and could make a new try.

During the 1400s, however, Europeans began to make long ocean voyages that kept them out of touch of known land for weeks, or even months. The absence of timekeeping equipment forced them to guess at then-longitude and they could easily lose themselves in the trackless sea. Nations like England and the Netherlands began to depend, more and more, on world-wide commerce and could not afford to fool around with lost ships.

Pendulum clocks wouldn't do on board ship, since the swaying would put the pendulum out of action. Ordinary watches wouldn't do either, because they weren't accurate enough. What was needed was a chronometer (Greek for "time measurer") that was small enough to be portable, unaffected by the swaying of a ship, and very accurate over long periods of time.

In 1713, therefore, the British government offered a prize of 20,000 pounds (an enormous fortune in those days) to anyone inventing such a timepiece. A British mechanic, John Harrison (1695-1776), managed to do

252

the needful, constructing a chronometer that kept time to within one minute after five months at sea.

The "gentlemen" of Parliament, however, objected strongly to paying a fortune to a mere mechanic and it took poor Harrison forty years to collect his prize money. King George III actually had to interfere on Harrison's side to make Parliament disgorge.

Clocks and watches continued to improve and to take into account, for instance, changes in temperature. They proved essential for the workings of an industrial society. Train travel, air travel, radio, and television all have to work on the minute, or even the second, if they are to work at all.

It came about, therefore, that almost everyone came to carry a timepiece in his pocket or on his wrist, and was constantly checking the time (at least, if he is as time-bound as I am). What this costs us all in endlessly being driven by each unforgiving minute—what it costs in

terms of ulcers and heart attacks—I can't say, but there's nothing to be done about it.

Individuals might deliberately step back into a timeless "simple life," but science, industry, and society in general simply cannot.

By 1950, the best mechanical clocks could keep such accurate time that they would gain or lose no more than a second in nineteen months, or less than a minute in an entire lifetime. It might seem silly to look for still more accurate methods of timekeeping, but greater accuracy was sought and found.

In 1880, French chemist Pierre Curie (1859-1906) and his brother, Jacques, discovered the phenomenon of piezoelectricity (where piezo comes from the Greek word meaning "to press").

They discovered that if certain crystals were placed

253

under pressure they would develop an electric potential. That seemed mysterious, but we now know that crystals are built up of particles, some of which carry positive electric charges and some negative. Under pressure, these charges are separated slightly, producing the potential.

The reverse is also true. If a crystal is placed under an electric potential, it compresses.

If a crystal is placed under an oscillating electric potential, it compresses and relaxes in rapid alternation, producing soundwaves equal in frequency to that of the potential oscillation. This means that a beam of ultrasonic vibrations (far too rapid to be heard) is formed, and can be used for what we now call sonar.

The tiny vibrations of the crystal are far more rapid, and far more regular, than the mechanical vibrations of pendulums and springs. What is needed, then, is a watch containing a small electric battery to supply the power, a crystal to undergo the vibrations, and a coupling that will enable the vibrations to turn the hands of a watch. Here, at least, there is no frequent need for rewinding. So little electricity is required that even a small battery can deliver the necessary power for a year or two before having to be replaced.

The best crystals for the purpose are crystals of quartz, that are hard, uniform, durable, and have vibrations that are almost independent of temperature. The first clock driven by a quartz crystal was built in

1928, and now crystal watches, with tiny quartz crystals cut into the shape of tuning forks, are built so cheaply and in such numbers that watches that have stems and require winding have come to seem archaic and quaint.

The best crystal clocks are so accurate that they could go a hundred thousand years or so (if they could be made to last so long) without gaining or losing more than a second.

254

But we can do better still. Atoms themselves have natural oscillations. An atomic nucleus has a magnetic field that interacts with the field of the electrons. As a result, the nucleus behaves as though it has an axis of rotation that precesses, that is, moves so that its ends mark out circles, billions of times a second.

The basic understanding of this nuclear precession came with the work of the Austrian-American physicist Isidor Isaac Rabi (1898-1988), beginning in 1937. By 1945, Rabi could see that the precession was sufficiently regular to be potentially useful for time measurements, and suggested the construction of atomic docks.

Eventually, such atomic clocks were indeed built and were shown to be more accurate than even the best crystal clocks.

Atomic clocks have already served to time Earth's rotation accurately enough to show that our planet is a comparatively lousy clock. Its period of rotation jogs slightly up and down as earthquakes, snowfalls, and storms alter its distribution of mass. It also slows progressively because of tidal action. Atomic clocks can tell us when to add leap seconds to the year to keep Earth in step with true time.

The result is that it is no longer necessary to base the length of the standard second on an astronomic motion —on a certain fraction of the year, for instance. Instead, in 1967, the international definition of the second was set as equal to 9,192,631,770 periods of the oscillation of the cesium atom.

And this is not the ultimate either. It is possible to make use of oscillations of hydrogen atoms under specialized conditions that would yield a clock that would, if it could only be maintained indefinitely, gain or lose not more than a second in a hundred million years. In the entire lifetime of the Universe such a hydrogen clock (if it could have been kept going for that entire

255

period) would have gained or lost not more than two and a half minutes.

And there is room for still further improvement by making use of lasers and strong cooling. There are also astronomical objects known as millisecond pulsars that rotate nearly a thousand times a second, shooting out radio pulses with each rotation, without the gradual slowing effect exhibited by ordinary pulsars. The periodicity of the pulses makes them no better than our best atomic clocks, perhaps, but millisecond pulsars require no maintenance, would be upset by nothing short of astronomic catastrophe, and can endure indefinitely.

But why bother? Is there any point in keeping time so accurately?

Yes, there is.

Einstein's special theory of relativity indicates that time slows down with velocity. At extremely high velocities (those of energetic subatomic particles) this slowing effect is noticeable, has been measured, and has been shown to check the theory virtually on the nose.

There is, however, a tiny slowing effect even at ordinary velocities and this is usually described as "immeasurably small." Well, with the best modern clocks, it isn't immeasurably small, and it has been measured and shown to check the theory, where one clock is, for instance, kept stationary relative to Earth's surface while the other is carried around the Earth on planes.

By Einstein's general theory of relativity, time also slows in the presence of gravitational fields. This could be measured where the gravitational field is enormously intense, as with pulsars, but the slowing effect is present (though extremely tiny) even in connection with a gravitational field as weak as that of our Sun. That, too, can be measured by the use of atomic clocks.

256

The general theory also predicts that radio waves will take very slightly longer to reach us if they skim past the Sun in the course of their passage since they then follow a slightly curved path rather than a straight line. This, too, has been checked.

Furthermore, it is becoming more and more necessary, in science, to synchronize instruments. Thus, radio telescopes are observing the Universe by way of radio waves that are a million times or more

longer than light waves. To see as clearly with radio waves as with light waves would require radio telescopes a million times wider than light telescopes.

This is impractical, but we can build two or more radio telescopes a goodly distance apart and by concentrating on the same object at the same time, it would be as though we had one telescope as wide as the several are separated in distance.

This, however, would mean that the various radio telescopes be in exact synchronization—that a particular radio wave enter all the instruments at the same time. 'This requires the use of the best atomic clocks we have in order to get the synchronization sufficiently exact. The result of our clocks is that we see much more clearly, and in much greater detail, by radio than by light. (We are beginning, however, to make use of multiple light telescopes as well.)

Then, too, with extremely good atomic clocks we can measure the rate of rotations of pulsars and check sudden "glitches" in those rates. We can check all sorts of things thought to be constant that might not be quite constant.

In short, the better we can make our timekeeping, the more profoundly we can study the Universe in finer and finer detail and the more we can "fill the unforgiving second, with a quintillion equal splits of distance run."

And there you are, Rudyard, old man.

257

Something

17 A Sacred Poet

I heard, once, that the oratory of William Jenning* Bryan, the populist leader of the Democratic party hi the first decade of this century, was likened to the North Platte river of his home state, Nebraska. His oratory they said, like the river, "was two miles wide and a foo deep."

Well, last night I met a very amiable and likable gen tleman who had spent decades in researching a particu lar subject and the result was that his knowledge was, in my opinion, two miles deep, but only a foot wide.

He gave a talk and, in the question-and-answer session that followed,

I had a little set-to with him. Twice tried to make my point, and twice he drowned me i irrelevant chatter. When I tried a third time, with a ring ing "Nevertheless—," the moderator stopped me foi fear I would forget my manners and offend the man.

261

In the course of the few things I did have a chance to say, however, I quoted the Latin poet Horace. No, I didn't quote him in Latin because I am not that kind of scholar, but I quoted him in English, which was good enough. The quotation goes as follows: "Many brave men lived before Agamemnon, but all are overwhelmed in eternal night, unwept, unknown, because they lack a sacred poet."

By this (which, by the way, was quite a propos of the point I was trying to make) Horace meant that not all Agamemnon's deeds and heroisms and high rank would have helped him live in memory had it not been that Homer wrote the Iliad. It was the poet's work and not the hero's that lived in memory.

Though I didn't get to make my point as I wished, the quote remained in my mind and it led me to the following essay, which will be quite unlike any I have offered you for, lo, these many years. Be patient with me, for I am going to discuss poetry.

Let me make a few things plain. First, I am no expert on poetry. I have a certain facility with parodies and limericks, but there it stops.

Nor do I pretend to any ability at judging the worth of poetry. I can't tell a good poem from a bad one, and I have never had the impulse to be a "critic."

So what am I going to talk about when I discuss poetry? Why, something that doesn't require judgment or poetic understanding, or even critical ability (if there should be such a thing).

I want to talk about the effect of poetry. Some poems have an effect on the world and some poems don't. It has nothing to do with being good or bad. That is a subjective decision and I imagine there will always and eternally be disagreements on such a matter. But there can't be any disagreement about a poem's effectiveness.

Let me give you an example:

262

In 1797, the infant United States built its first warships. One of them,

built in Boston, was the Constitution. The ship had a brief workout when France and the United States had a minor unofficial naval war in 1798. The real test came in 1812, when the United States went to war with Great Britain for a second time. This war started with a humiliation on land. General William Hull, an utter incompetent, surrendered Detroit to the British virtually without a blow. (Hull was court-martialed and condemned to death for this, but was granted a reprieve because of his services in the Revolutionary War.)

What saved the American morale in those difficult first months was the feats of our small navy, which took on the proud British men-of-war and knocked them for a loop. The Constitution was under the command of William Hull's younger brother, Isaac Hull. On August 19, 1812, the Constitution met up with the British Guer-riere ("Warrior"), and in two and a half hours riddled it into a Swiss cheese so that it had to be sunk.

On December 19, the Constitution, under a new captain, destroyed another British warship off the coast of Brazil. In this second battle, the British cannonballs bounced off the seasoned timbers of the Constitution's hull, doing no damage, and the crew cheered the sight. One cried out that the ship's sides were made of iron. The ship was at once named Old Ironsides, and it has been known by that name ever since, to the point where I imagine few people remember its real name.

Well, ships grow old and by 1830, Old Ironsides was obsolete. The Navy was ready to scrap it, for it had far better ships now. Congress wasn't anxious to spend any more money on it, so the scrapping looked good. There were some sentimentalists who thought the ship ought to be preserved as a national treasure, but who cares

263

about a few soft-headed jerks. Besides, you can't fight City Hall, the saying goes.

In Boston, however, there lived a twenty-one-year-old youngster named Oliver Wendell Holmes. He had just graduated from Harvard, he was planning to study medicine, and he had dashed off reams of poetry. In fact, his fellow students had named him class poet.

So Holmes wrote a poem entitled "Old Ironsides." Perhaps you know it. Here's the way it goes:

Ay, tear her tattered ensign down!

Long has it waved on high, And many an eye has danced to see

That banner in the sky; Beneath it rung the battle shout,

And burst the cannon's roar— The meteor of the ocean air

Shall sweep the clouds no more.

Her deck, once red with heroes' blood,

Where knelt the vanquished foe, When winds were hurrying o'er the flood,

And waves were white below, No more shall feel the victor's tread,

Or know the conquered knee— The harpies of the shore shall pluck

The eagle of the sea!

Oh, better that her shattered hulk

Should sink beneath the wave; Her thunders shook the mighty deep,

And there should be her grave; Nail to the mast her holy flag,

Set every threadbare sail, And give her to the god of storms,

The lightning and the gale!

264

The poem was published on September 14, 1830, and was quickly reprinted everywhere.

Is the poem a good one? I don't know. For all I know, critics will say it is mawkish and overblown, and that its images are melodramatic. Perhaps. I only know that I have never been able to read it aloud with a steady voice, particularly when I get to the parts about the harpies and about the threadbare sails. I can't even read it to myself, as I did just now, without gulping and finding it difficult to see the paper.

To critics, that may make me an object of scorn and derision, but the fact is that I'm not, and wasn't, the only one. Wherever that poem appeared, a sudden roar of protest arose from the public. Everyone began contributing money to help save Old Ironsides. The schoolchil- dren brought their pennies to school. There was no stopping it. The Navy and the Congress found themselves facing an aroused public and

discovered that it wasn't Old Ironsides that was battling the god of storms; they were.

They gave in at once. Old Ironsides was not scrapped. It was never scrapped. It was rebuilt in 1833 and still exists, resting in Boston Harbor, where it will continue to exist indefinitely.

It was not Old Ironsides'5 feats of war that saved it; it was that it had a sacred poet. Good or bad, the poem was effective.

The War of 1812 gave us a poem called "The Defense of Fort McHenry," which was published on September 14, 1814, and was quickly renamed "The Star-Spangled Banner."

It's our national anthem now. It's difficult to sing (even professional singers have trouble sometimes) and the words don't flow freely. Most Americans, however patriotic, know only the first line. (I am rather proud of

265

the fact that I know, and can sing without hesitation, all

four stanzas.)

All four stanzas? Every Fourth of July, the New York Times prints the music and all the words of the anthem, and try as you might, you will count only three stanzas. Why? Because during World War II, the government abolished the third stanza as too bloodthirsty.

Remember that the poem was written in the aftermath of the British bombardment of Fort McHenry in Baltimore Harbor. If the fort's guns had been silenced, then the British ships could have disembarked the soldiers they carried. Those soldiers would surely have taken Baltimore and split the nation (which was still hugging the seacoast) in two. Those soldiers had already sacked Washington, which was a small hick town of no importance. Baltimore was an important port.

In the course of the night, the ships' guns fell silent and to Francis Scott Key, on board one of the British ships (trying to get the release of a friend), the whole question was whether the American guns had been silenced and knocked out, or whether the British ships had given up the bombardment. Once the dawn came, the answer would be plain—it would depend on whether the American flag or the British flag was flying

over the fort.

The first stanza asks, then, whether the American flag is still flying. The second stanza tells us it is still flying. The third stanza is a paean of unashamed triumph and here it is: And where is that band who so vauntingly swore

That the havoc of war and the battle's confusion, A home and a country should leave us no more? Their blood has washed out their foul footsteps'

pollution.

No refuge could save the hireling and slave

266

From the terro'r of flight, or the gloom of the grave;

And the star-spangled banner in triumph doth

wave

O'er the land of the free and the home of the brave.

Good poetry? Who knows? Who cares? If you know the tune, sing it. Get the proper scorn into "foul footsteps," the proper hatred into "hireling and slave," the proper sadistic glee into the "terror of flight, or the gloom of the grave," and you'll realize that it rouses passions a little too strongly. . . . But who knows, there are times when you may want those passions.

I should point out that music plays its part, too. Sing a poem and the effect is multiplied manifold.

Consider the American Civil War. For over two years, the Union suffered one disaster after another in Virginia. The muttonheads who ran the Union army were, one after the other, no match for Robert E. Lee and Thomas J. "Stonewall" Jackson. Those were the best soldiers the United States ever produced and, as Fate would have it, they fought their greatest battles against the United States.

Why did the North continue to fight? The South was ready to stop at any time. The North needed only to agree to leave the South alone and the war would be over. But the North continued to fight through one bloody debacle after another. One of the reasons for that was the character of President Abraham Lincoln, who would not quit under

any circumstances—but another was that the North was moved by a religious fervor.

Consider "The Battle Hymn of the Republic." It's a march, yes, but not a war march. It is God, not man, who is marching. The key word in the title is "hymn,"

267

not "battle," and it is (or should be) always sung slowly and with the deepest emotion.

Julia Ward Howe, who wrote the words (to the well-known tune of "John Brown's Body"), had just visited the camps of the Army of the Potomac in 1862, and was very moved. It must surely have expressed a great deal of what many Northerners felt. There are five stanzas to the poem and most Americans, today, barely know the first, but during the Civil War it was all five that were known. Here is the fifth:

In the beauty of the lilies Christ was born across the sea,

With a glory in His bosom that transfigures you and me;

As he died to make men holy let us die to make men free, While God is marching on.

"Let us die to make men free!" I don't say that everyone in the North had that fervor, but some did, and the words may have swayed those on the borderline. After all, something kept the Northern armies fighting through disaster after disaster, and the "Battle Hymn" was surely one of the factors.

And if slavery moved some Northerners as an evil that must be fought and destroyed at any cost, there were other Northerners to whom the Union was a benefit that must be supported and preserved at any cost, and there was a song for that, too.

The worst defeat suffered by the Union was in December 1862, when the unspeakable General Ambrose Burnside, perhaps the most incompetent general ever to lead an American army into battle, sent his soldiers against an impregnable redoubt manned by the Confederate army. Wave after wave of the Union army surged forward, and wave after wave was cut down.

268

It was after that battle that Lincoln said, "If there is a worse place

than Hell, then I am in it." He also remarked of Burnside on a later occasion that he "could snatch defeat from the very jaws of victory."

But, the story goes, as the Northern army lay in camp that night trying to recover, someone struck up a new song that had been written by George Frederick Root, who had already written "Tramp! Tramp! Tramp! The Boys Are Marching." This time he had come up with something called "The Battle-Cry of Freedom." Here is one of the stanzas:

Yes, we'll rally round the flag, boys, we'll rally once again,

Shouting the battle-cry of Freedom. We will rally from the hillside, we'll gather from the plain,

Snouting the battle-cry of Freedom. The Union forever! Hurrah, boys, hurrah! Down with the traitor, and up with the star! And we'll rally round the flag, boys, we'll raUy once again,

Shouting the battle-cry of Freedom.

Even I, with my tin ear, have a sneaking suspicion that this is not great poetry, or even adequate poetry, but (the story continues) a Confederate officer, hearing those distant strains from the defeated army, gave up hope at that moment. He felt that a defeated army that could still sing that song about "the Union forever" would never be finally defeated, but would keep coming back to the assault again and again, and would never give up till the Confederacy was worn out and could fight no more. . . . And he was right.

There's something about words and music together that have an amazing effect.

There's an ancient Greek story, for instance, that may

269

conceivably be true (the Greeks never spoiled a story by worrying over the facts of the case). According to it, the Athenians, fearful of a loss in a forthcoming battle, sent to the oracle at Delphi for advice. The oracle advised them to ask the Spartans to lend them one soldier.

The Spartans did not like to defy the oracle so they gave the Athenians one soldier, but, not particularly anxious to help a rival city to victory, were careful not to give Athens a general or a renowned fighter. They handed the Athenians a lame regimental musician. And at the battle, the Spartan musician played and sang such stirring music that the

Athenians, cheering, advanced on the enemy at a run and swept the field.

Then there's the story (probably also apocryphal) of an event that took place in the Soviet Union during the Nazi invasion, when a group of German soldiers, meticulously dressed in Soviet uniforms, marched into Soviet-held territory in order to carry out an important sabotage mission. A young boy, seeing them pass, hastened to the nearest Soviet army post and reported a group of German soldiers dressed in Soviet uniforms. The Nazis were rounded up and, I presume, given the treatment routinely accorded spies.

The boy was then asked, "How did you know those were German soldiers, and not Soviet soldiers?"

And the boy replied, "They weren't singing."

For that matter, did you ever see John Gilbert in The Big Parade, a silent movie about World War I? Gilbert has no intention of being caught up in war hysteria and joining the army, but his car is stopped by a parade passing by—men in uniform, the flag flying, instruments blowing and banging away.

It's a silent movie, so you don't hear any words, you don't hear any music (except for the usual piano accompanist), you don't hear any cheering. You see only Gilbert's face behind the wheel, cynically amused. But he

270

has to stay there till the parade is done, and after a moment one of Gilbert's feet is tapping out the time then both feet are, then he's beginning to look excited and eager and—of course—he gets out of the car to enlist.. Without hearing a thing, you find it completely convincing. That is how people get caught up and react.

I can give you a personal experience of my own. As you probably can guess from what I've already said, I am not only a Civil War buff, but with respect to that war I'm an ardent Northern patriot. "The Union forever," that's me.

But once when I was driving from New York to Boston, alone in my car, I was listening to a series of Civil War songs on the car radio. One song I had never heard before and I've never heard since. It was a Confederate song at a desperate time and it was pleading for the South to unite and, with all its strength, throw back the Yankee invaders. And, by the time the song was over, I was in utter distress,

knowing that the war had ended over a century before and that there was no Confederate recruiting station to which I could run and volunteer.

Insidious, the power these things have.

During the Crimean War, with Great Britain and France fighting Russia, the British commanding general, Baron Raglan, gave a command that was so ambiguous and was accompanied by a gesture that was so uncertain that nobody knew exactly what he meant. Because no one dared say, "That's crazy," the order ended by sending 607 men and horses of the Light Brigade charging pell-mell into the main Russian army. In twenty minutes, half the men and horses were casualties and, of course, nothing was accomplished.

The commander of the French contingent, Pierre Bosquet, stared in disbelief as the men rode their horses into the mouths of cannon and said, "C'est magnifique,

271

mais ce n'est pas la guerre." Freely translated, he was saying, "That's all very nice, but that's not the way you fight a war."

However, Alfred, Lord Tennyson wrote a poem about it, that starts with the familiar:

Half a league, half a league

Half a league onward, All in the valley of death

Rode the six hundred.

He wrote a total of fifty-five lines, hi a rhythm that mimicked perfectly the sound of galloping horses. Read it properly, and you'll think you're one of the horsemen careening forward on that stupid charge.

Tennyson didn't actually hide the fact that it was a mistake. He says:

"Forward the Light Brigade!" Was there a man dismayed? Not tho the soldier knew

Someone had blundered: Theirs not to make reply, Theirs not to reason why, Theirs but to do and die. Into the valley of death

Rode the six hundred.

The result is that, thanks to the poem, everyone thinks of the Charge in its heroic aspect, and no one thinks of it as an example of criminally inept generalship.

Sometimes a poem totally distorts history and keeps it distorted too.

In 1775, the British controlled Boston while dissident colonials were concentrated at Concord. General Gage, the British commander, sent a contingent of soldiers to confiscate the arms and powder that were being stored

272

at Concord and to arrest Samuel Adams and John Hancock, who were the ringleaders of dissent.

Secrets weren't kept well, and colonial sympathizers in Boston set out to ride through the night to warn Adams and Hancock to make themselves scarce, and to warn the people in Concord to hide the arms and powder. Two of the riders were Paul Revere and William Dawes. They took different routes but got to Lexington. Adams and Hancock were staying there and, on hearing the news, quickly rode out of town.

Revere and Dawes then went on to Concord, but were stopped by a British patrol and were arrested. That was it for both of them. Neither one of them ever got to Concord. Neither of them gave the vital warning to the men of Concord.

However, in Lexington, Revere and Dawes had been joined by a young doctor named Samuel Prescott, who was still awake because he had been with a woman, doing what I suppose a man and woman would naturally do when alone late at night.

He buttoned his pants and joined the two. He avoided the British patrol and managed to get to Concord. He got the Concord people roused and ready, with the arms parceled out for defense.

The next day, when the British stormed their way through Lexington and got to Concord, the Miriutemen were waiting for them behind their trees, guns in hand. The British just barely made it back to Boston and the War of the American Revolution had begun.

Lexington-and-Concord remained famous forever after, but the business about people riding to give warning was drowned out somehow. No one knew of it.

In 1863, however, the Civil War was at its height and the North was still looking for its great turning point victory (which was to come at Gettysburg in July of that year). Henry Wadsworth Longfellow felt the urge to

write a patriotic ballad to hearten the Union side, so he dug up the old tale no one remembered and wrote a poem about the warning ride in the night.

And he ended it with a mystical evocation of the ghost of that rider:

Through all our history, to the last, In the hour of darkness and peril and need, The people will waken and listen to hear The hurrying hoof-beats of that steed, And the midnight message of Paul Revere.

The poem proved immensely popular, and very heartening to its readers with its implication that the ghosts of the past were fighting on the side of the Union.

But there was an important flaw in it. Longfellow mentioned only Paul Revere, who, after all, had never completed the job. It was Prescott who warned Concord.

And did you ever hear of Prescott? Did anyone ever hear of Prescott? Of course not.

Prescott's role is no secret. Any reasonable history book, any decent encyclopedia, will tell you all about it. But what people know is not history, not encyclopedias, but:

Listen, my children, and you shall hear Of the midnight ride of Paul Revere . . .

That's the power of a poem, even (if you'll forgive my tin ear and let me make a judgment) of a rotten poem like "Paul Revere's Ride"!

ABOUT THE AUTHOR

Isaac Asimov is America's most prolific author, with over four hundred and thirty books published. He is known and loved the world over for his science fiction— I including the Foundation series— mystery stories, and nonfiction. The essays in this collection were first published in The Magazine of Fantasy and Science Fiction in 1987 and 1988.