The Cosmic Origins of Time's Arrow

One of the most basic facts of life is that the future looks different from the past. But on a grand cosmological scale, they may look the same

By Sean M. Carroll

The universe does not look right. That may seem like a strange thing to say, given that cosmologists have very little standard for comparison. How do we know what the universe is supposed to look like? Nevertheless, over the years we have developed a strong intuition for what counts as "natural"—and the universe we see does not qualify.

Make no mistake: cosmologists have put together an incredibly successful picture of what the universe is made of and how it has evolved. Some 14 billion years ago the cosmos was hotter and denser than the interior of a star, and since then it has been cooling off and thinning out as the fabric of space expands. This picture accounts for just about every observation we have made, but a number of unusual features, especially in the early universe, suggest that there is more to the story than we understand.

Among the unnatural aspects of the universe, one stands out: time asymmetry. The microscopic laws of physics that underlie the behavior of the universe do not distinguish between past and future, yet the early universe—hot, dense, homogeneous—is completely different from today's—cool, dilute, lumpy. The universe started off orderly and has been getting increasingly disorderly ever since. The asymmetry of time, the arrow that points from past to future, plays an unmistakable role in our everyday lives: it accounts for why we cannot turn an omelet into an egg, why ice cubes never spontaneously unmelt in a glass of water, and why we remember the past but not the future. And the origin of the asymmetry we experience can be traced all the way back to the orderliness of the universe near the big bang. Every time you break an egg, you are doing observational cosmology.

The arrow of time is arguably the most blatant feature of the universe that cosmologists are currently at an utter loss to explain. Increasingly, however, this puzzle about the universe we observe hints at the existence of a much larger spacetime we do not observe. It adds support to the notion that we are part of a multiverse whose dynamics help to explain the seemingly unnatural features of our local vicinity.

The Puzzle of Entropy

Physicists encapsulate the concept of time asymmetry in the celebrated second law of thermodynamics: entropy in a closed system never decreases. Roughly, entropy is a measure of the disorder of a system. In the 19th century, Austrian physicist Ludwig Boltzmann explained entropy in terms of the distinction between the microstate of an object and its macrostate. If you were asked to describe a cup of coffee, you would most likely refer to its macrostate—its temperature, pressure and other overall features. The microstate,

KEY CONCEPTS

- The basic laws of physics work equally well forward or backward in time, yet we perceive time to move in one direction only—toward the future. Why?
- To account for it, we have to delve into the prehistory of the universe, to a time before the big bang. Our universe may be part of a much larger multiverse, which as a whole is time-symmetric. Time may run backward in other universes.

—The Editors

KENN BROWN

[THE AUTHOR]

Sean M. Carroll is a senior research associate in physics at the California Institute of Technology. His research ranges over cosmology, particle physics and Einstein's general theory of relativity, with a particular expertise in dark energy. He has been awarded fellowships from the Sloan and Packard foundations, as well as the M.I.T. Graduate Student Council Teaching Award and the Villanova University Arts and Sciences Alumni Medallion. Outside of academia, Carroll is best known as a contributor to the blog Cosmic Variance, which is not only one of the most thoughtful science blogs but also the way he met his wife, science writer Jennifer Ouellette.

on the other hand, specifies the precise position and velocity of every single atom in the liquid. Many different microstates correspond to any one particular macrostate: we could move an atom here and there, and nobody looking at macroscopic scales would notice.

Entropy is the number of different microstates that correspond to the same macrostate. (Technically, it is the number of digits, or logarithm, of that number.) Thus, there are more ways to arrange a given number of atoms into a high-entropy configuration than into a low-entropy one. Imagine that you pour milk into your coffee. There are a great many ways to distribute the molecules so that the milk and coffee are completely mixed together but relatively few ways to arrange them so that the milk is segregated from the surrounding coffee. So the mixture has a higher entropy.

From this point of view, it is not surprising that entropy tends to increase with time. High-entropy states greatly outnumber low-entropy ones; almost any change to the system will land it in a higher-entropy state, simply by the luck of the draw. That is why milk mixes with coffee but never unmixes. Although it is physically possible for all the milk molecules to spontaneously conspire to arrange themselves next to one another, it is statistically very unlikely. If you waited for it to happen of its own accord as molecules randomly reshuffled, you would typically have to wait much longer than the current age of the observable universe. The arrow of time is simply the tendency of systems to evolve toward one of the numerous, natural, high-entropy states.

But explaining why low-entropy states evolve into high-entropy states is different from explaining why entropy is increasing in our universe. The question remains: Why was the entropy low to start with? It seems very unnatural, given that low-entropy states are so rare. Even granting that our universe today has medium entropy, that does not explain why the entropy used to be even lower. Of all the possible initial conditions that could have evolved into a universe like ours, the overwhelming majority have much higher entropy, not lower [see "The Arrow of Time," by David Layzer; Scientific American, December 1975].

In other words, the real challenge is not to explain why the entropy of the universe will be higher tomorrow than it is today but to explain why the entropy was lower yesterday and even lower the day before that. We can trace this logic all the way back to the beginning of time in our observable universe. Ultimately, time asymmetry is a question for cosmology to answer.

The Disorder of Emptiness

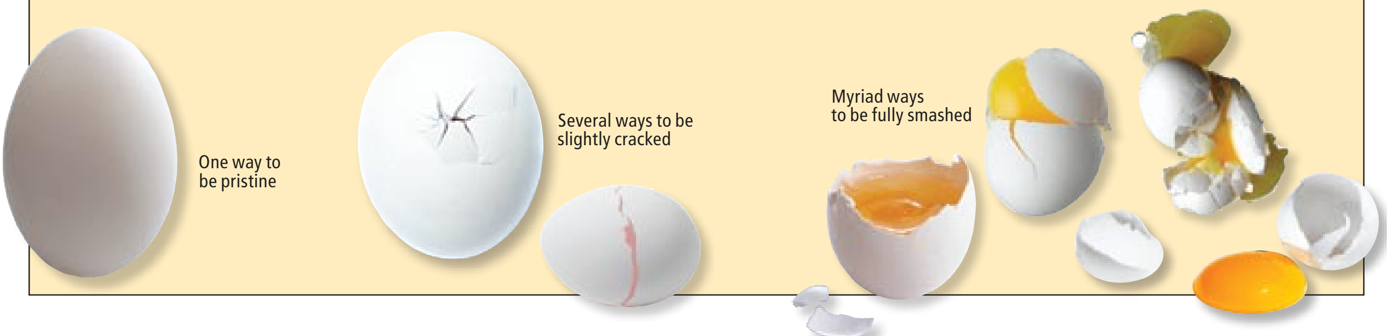
The early universe was a remarkable place. All the particles that make up the universe we currently observe were squeezed into an extraordinarily hot, dense volume. Most important, they were distributed nearly uniformly throughout that tiny volume. On average, the density differed from place to place by only about one part in 100,000. Gradually, as the universe expanded and cooled, the pull of gravity enhanced those differences. Regions with slightly more particles formed stars and galaxies, and regions with slightly fewer particles emptied out to form voids.

Clearly, gravity has been crucial to the evolution of the universe. Unfortunately, we do not fully understand entropy when gravity is involved. Gravity arises from the shape of spacetime, but we do not have a comprehensive theory of spacetime; that is the goal of a quantum theory of gravity. Whereas we can relate the entropy of a fluid to the behavior of the molecules that constitute it, we do not know what constitutes space,

[FROM ORDER TO DISORDER]

ENTROPY IN THE KITCHEN

A raw egg exemplifies the asymmetry of time: a fresh one breaks easily, but a broken one does not spontaneously put itself together again, for the simple reason that there are more ways to be broken than not. In physics jargon, the broken egg has a higher entropy.



KEN WEINGART (*Carroll*); ALL EGG PHOTOGRAPHY COURTESY OF GETTY IMAGES: RICHARD DRURY (*whole*); GRAEME MONTGOMERY (*cracked*); JAN STROMME (*cracked in half*); MICHAEL ROSENFELD (*half with yolk*); JONATHAN KANTOR (*seeping yolk and smashed*); DIAMOND SKY IMAGES (*over easy*) so we do not know what gravitational microstates correspond to any particular macrostate.

Nevertheless, we have a rough idea of how entropy evolves [*see box below*]. In situations where gravity is negligible, such as a cup of coffee, a uniform distribution of particles has a high entropy. This condition is a state of equilibrium. Even when particles reshuffle themselves, they are already so thoroughly mixed that nothing much seems to happen macroscopically. But if gravity is important and the volume is fixed, a smooth distribution has relatively low entropy. In this case, the system is very far from equilibrium. Gravity causes particles to clump into stars and galaxies, and entropy increases noticeably—consistent with the second law.

Indeed, if we want to maximize the entropy of a volume when gravity is active, we know what we will get: a black hole. In the 1970s Stephen Hawking of the University of Cambridge confirmed a provocative suggestion of Jacob Bekenstein, now at the Hebrew University of Jerusalem, that black holes fit neatly into the second law. Like the hot objects that the second law was originally formulated to describe, black holes emit radiation and have entropy—a lot of it. A single million-solar-mass black hole, such as the one that lives at the center of our galaxy, has 100 times the entropy of all the ordinary particles in the observable universe.

Eventually even black holes evaporate by emitting Hawking radiation. A black hole does not have the highest possible entropy—but just the highest entropy that can be packed into a certain volume. The volume of space in the universe, however, appears to be growing without limit. In

[ENTROPY OF A GAS]

WHAT GRAVITY DOES TO ENTROPY

What qualifies as low entropy or high entropy depends on the situation. Physicists identify the high-entropy state of a system based on how the system evolves over time. For example, if a diffuse and sufficiently cool gas feels the tug of gravity, it evolves to a clump. The law of entropy increase then implies that the clump has a high entropy, even though at first glance it might appear to be orderly (low entropy).

- **Gravity shut off**
- **Volume fixed**

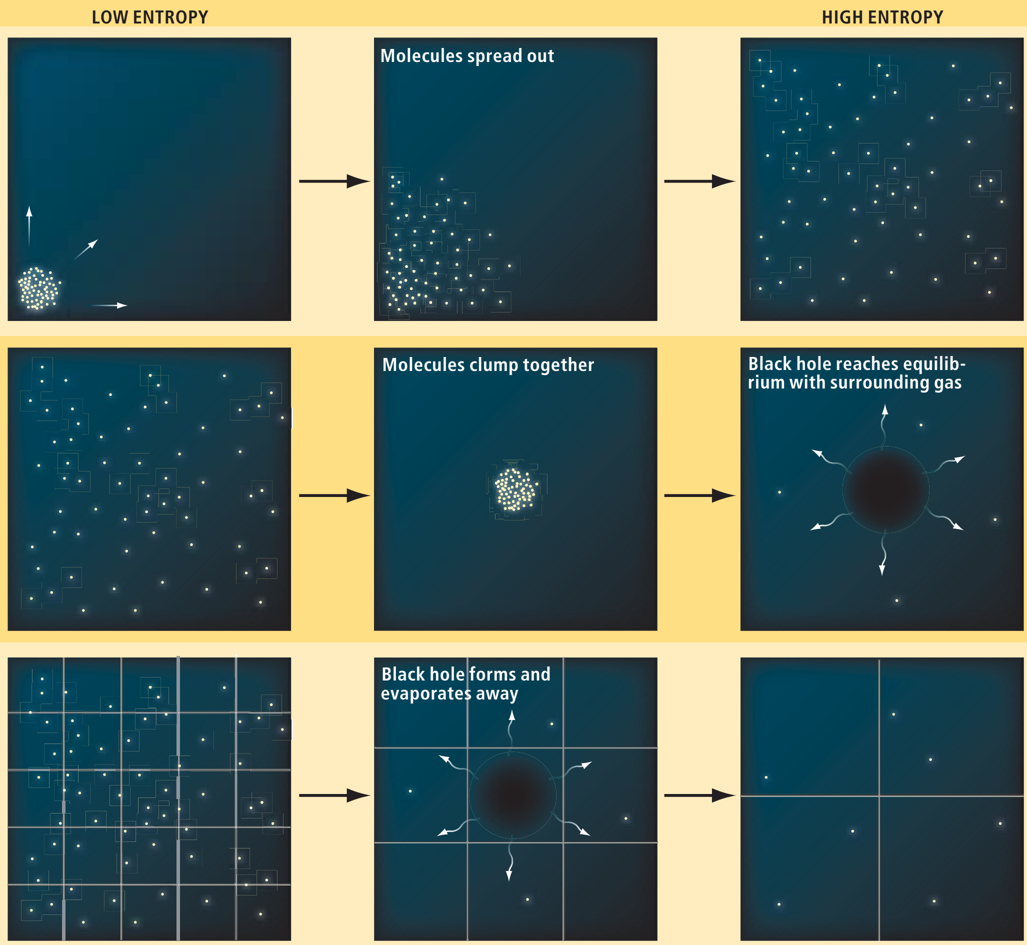
When gravity is negligible, a gas in a box has low entropy if it sits neatly in one corner and high entropy if it sprawls out. Thus, sprawl it does.

- **Gravity turned on**
- **Volume fixed**

Where gravity is significant, the opposite is true: the gas maxes out its entropy by collapsing to a black hole. Thus, a gravitating gas tends to clump rather than spread. The hole can survive forever in equilibrium with its surroundings.

- **Gravity turned on**
- **Volume expanding**

If the box is growing in size, the gas initially clumps and forms a black hole, but then the hole evaporates away. The gas it leaves behind continues to increase in entropy forever by spreading into an ever thinner gruel.



LUCY READING-IKKANDA

1998 astronomers discovered that cosmic expansion is accelerating. The most straightforward explanation is the existence of dark energy, a form of energy that exists even in empty space and does not appear to dilute away as the universe expands. It is not the only explanation for cosmic acceleration, but attempts to come up with a better idea have so far fallen short.

If dark energy does not dilute away, the universe will expand forever. Distant galaxies will disappear from view [see "The End of Cosmology?" by Lawrence M. Krauss and Robert J. Scherrer; SCIENTIFIC AMERICAN, March]. Those that do not will collapse into black holes, which in turn will evaporate into the surrounding gloom as surely as a puddle dries up on a hot day. What will be left is a universe that is, for all intents and purposes, empty. Then and only then will the universe truly have maxed out its entropy. The universe will be in equilibrium, and nothing much will ever happen.

It may seem strange that empty space has such a huge entropy. It sounds like saying that the most disorganized desk in the world is a completely empty desk. Entropy requires microstates, and at first glance empty space does not have any. In actuality, though, empty space has plenty of microstates—the quantum-gravitational microstates built into the fabric of space. We do not yet know what exactly these states are, any more than we know what microstates account for the entropy of a black hole, but we do know that in an accelerating universe the entropy within the observable volume approaches a constant value proportional to the area of its boundary. It is a truly enormous amount of entropy, far greater than that of the matter within that volume.

TIME'S ARROW FAQs, PART I

If entropy always increases, then how do low-entropy objects such as eggs form in the first place? The law of entropy applies to closed systems. It does not forbid decreases in entropy in open systems, including chickens. A hen takes in energy and goes through a great deal of effort to produce an egg.

Don't some particle processes have a built-in arrow of time? The decays of some elementary particles, such as neutral kaons, happen more frequently in one direction of time than the other. (Physicists do not need to travel backward in time to observe this asymmetry; they infer it from experiments on related particle properties.) But these processes are reversible, unlike the growth of entropy, so they do not explain the arrow of time. The Standard Model of particle physics does not seem to be of any help in explaining the low entropy of the early universe.

Past vs. Future

The striking feature of this story is the pronounced difference between the past and the future. The universe starts in a state of very low entropy: particles packed together smoothly. It evolves through a state of medium entropy: the lumpy distribution of stars and galaxies we see around us today. It ultimately reaches a state of high entropy: nearly empty space, featuring only the occasional stray low-energy particle.

Why are the past and future so different? It is not enough to simply posit a theory of initial conditions—a reason why the universe started with low entropy. As philosopher Huw Price of the University of Sydney has pointed out, any reasoning that applies to the initial conditions should also apply to the final conditions, or else we will be guilty of assuming the very thing we were trying to prove—that the past was special. Either we have to take the profound asymmetry of time as a blunt feature of the universe that escapes explanation, or we have to dig deeper into the workings of space and time.

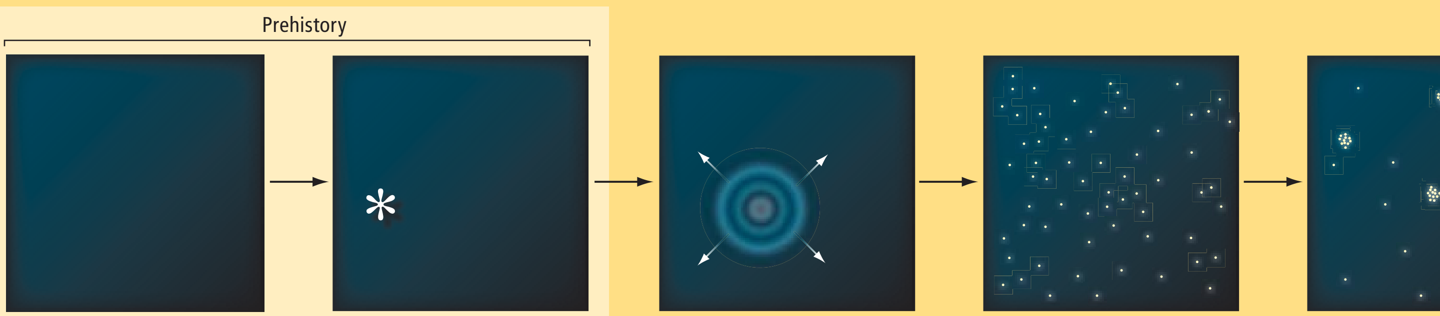
Many cosmologists have tried to attribute the time asymmetry to the process of cosmological inflation. Inflation is an attractive explanation for many basic features of the universe. According to this idea, the very early universe (or at least some part of it) was filled not with particles but rather with a temporary form of dark energy, whose density was enormously higher than the dark energy we observe today. This energy caused the expansion of the universe to accelerate at a fantastic rate, after which it decayed into matter and radiation, leaving behind a tiny wisp of dark energy that is becoming relevant again today. The rest of the story of the big bang, from

LUCY READING-IKKANDA (*timeline*)

[FROM "HEAT BIRTH" TO "HEAT DEATH"]

RESTORING SYMMETRY TO TIME

According to the standard model of cosmology, the universe began as a nearly uniform gas and will end up as nearly empty space—in short, it goes from low entropy to high entropy, a final condition that physicists call "heat death." But this model fails to explain what set up the initial



Space is nearly empty.

Quantum fields in one region fluctuate.

Space inflates.

Inflation ends and fills space with a nearly uniform primordial gas.

Clumps beget galaxies.

the smooth primordial gas to galaxies and beyond, simply follows.

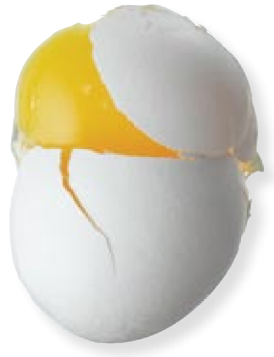
The original motivation for inflation was to provide a robust explanation for the finely tuned conditions in the early universe—in particular, the remarkably uniform density of matter in widely separated regions. The acceleration driven by the temporary dark energy smooths out the universe almost perfectly. The prior distribution of matter and energy is irrelevant; once inflation starts, it removes any traces of the preexisting conditions, leaving us with a hot, dense, smooth early universe.

The inflationary paradigm has been very successful in many ways. Its predictions of slight deviations from perfect uniformity agree with observations of density variations in the universe. As an explanation for time asymmetry, however, cosmologists increasingly consider it a bit of a cheat, for reasons that Roger Penrose of the University of Oxford and others have emphasized. For the process to work as desired, the ultradense dark energy had to begin in a very specific configuration. In fact, its entropy had to be fantastically *smaller* than the entropy of the hot, dense gas into which it decayed. That implies inflation has not really solved anything: it "explains" a state of unusually low entropy (a hot, dense, uniform gas) by invoking a prior state of even lower entropy (a smooth patch of space dominated by ultradense dark energy). It simply pushes the puzzle back a step: Why did inflation ever happen?

One of the reasons many cosmologists invoke inflation as an explanation of time asymmetry is that the initial configuration of dark energy does not *seem* all that unlikely. At the time of inflation, our observable universe was less than a centimeter across. Intuitively, such a tiny region does not have many microstates, so it is not so improbable for the universe to stumble by accident into the microstate corresponding to inflation.

Unfortunately, this intuition is misleading. The early universe, even if it is only a centimeter across, has exactly the same number of microstates as the entire observable universe does today. According the rules of quantum mechanics, the total number of microstates in a system never changes. (Entropy increases not because the number of microstates does but because the system naturally winds up in the most generic possible macrostate.) In fact, the early universe is the same physical system as the late universe. One evolves into the other, after all.

Among all the different ways the microstates of the universe can arrange themselves, only an incredibly tiny fraction correspond to a smooth configuration of ultradense dark energy packed into a tiny volume. The conditions necessary for inflation to begin are extremely specialized and therefore describe a very low entropy configuration. If you were to choose configurations of the universe randomly, you would be highly unlikely to hit on the right conditions to start inflation. Inflation does not, by itself, explain why the early universe has a low entropy; it simply assumes it from the start.



FAQs, PART II

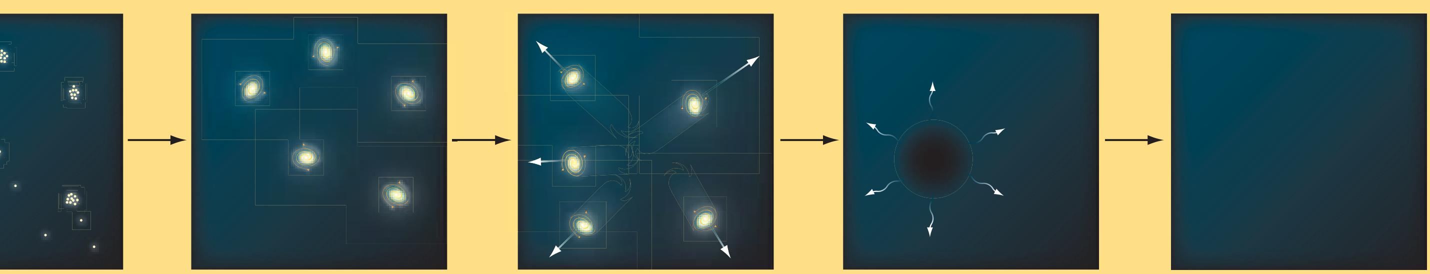
Doesn't quantum mechanics have an arrow of time? According to the standard interpretation of quantum mechanics, the measurement of a system causes its wave function to "collapse," a process that is asymmetric in time. But the reason wave functions collapse yet never uncollapse is the same reason that eggs break yet never unbreak—namely, because collapse increases the entropy of the universe. Quantum mechanics does not help explain why the entropy was low in the first place.

Why do we remember the past but not the future? To form a reliable memory requires that the past be orderly—that is, have a low entropy. If the entropy is high, almost all "memories" would be random fluctuations, completely unrelated to what actually happened in the past.

A Time-Symmetric Universe

Thus, inflation is of no help in explaining why the past is different from the future. One bold but simple strategy is just to say: perhaps the very far past is not different from the future after all. Perhaps the distant past, like the future, is actually a high-entropy state. If so, the hot, dense

low-entropy state. The author's model adds a period of prehistory. The universe began empty and will end up empty—the appearance of stars and galaxies is a temporary deviation from its usual equilibrium condition. (This figure is schematic; it does not show space expanding.)



The universe today.

Accelerating expansion pulls galaxies out of sight.

Each galaxy collapses to a black hole, which evaporates into a diffuse gas.

Space is nearly empty.

state we have been calling "the early universe" is actually not the true beginning of the universe but rather just a transitional state between stages of its history.

Some cosmologists imagine that the universe went through a "bounce." Before this event, space was contracting, but instead of simply crashing to a point of infinite density, new physical principles—quantum gravity, extra dimensions, string theory or other exotic phenomena—kicked in to save the day at the last minute, and the universe came out the other side into what we now perceive as the big bang. Though intriguing, bouncing cosmologies do not explain the arrow of time. Either entropy was increasing as the prior universe approached the crunch—in which case the arrow of time stretches infinitely far into the past—or the entropy was decreasing, in which case an unnatural low-entropy condition occurred in the middle of the universe's history (at the bounce). Either way, we have again passed the buck on the question of why the entropy near what we call the big bang was small.

Instead let us suppose that the universe started in a high-entropy state, which is its most natural state. A good candidate for such a state is empty space. Like any good high-entropy state, the tendency of empty space is to just sit there, unchanging. So the problem is: How do we get our current universe out of a desolate and quiescent spacetime? The secret might lie in the existence of dark energy.

In the presence of dark energy, empty space is not completely empty. Fluctuations of quantum fields give rise to a very low temperature—enormously lower than the temperature of today's universe but nonetheless not quite absolute zero. All quantum fields experience occasional thermal fluctuations in such a universe. That means it is not perfectly quiescent; if we wait long enough, individual particles and even substantial collections of particles will fluctuate into existence, only to once again disperse into the vacuum. (These are real particles, as opposed to the short-lived "virtual" particles that empty space contains even in the absence of dark energy.)

Among the things that can fluctuate into existence are small patches of ultradense dark energy. If conditions are just right, that patch can undergo inflation and pinch off to form a separate universe all its own—a baby universe. Our universe may be the offspring of some other universe.

Superficially, this scenario bears some resemblance to the standard account of inflation. There, too, we posit that a patch of ultradense dark energy arises by chance, igniting inflation. The difference is the nature of the starting conditions. In the standard account, the patch arose in a wildly fluctuating universe, in which the vast bulk of fluctuations produced nothing resembling inflation. It would seem to be much more likely for the universe to fluctuate straight into a hot big bang, bypassing the inflationary stage altogether. Indeed, as far as entropy is concerned, it would be even more likely for the universe to fluctuate straight into the configuration we see today, bypassing the past 14 billion years of cosmic evolution.

In our new scenario, the preexisting universe was never randomly fluctuating; it was in a very specific state: empty space. What this theory claims—and what remains to be proved—is that the most likely way to create universes like ours from such a preexisting state is to go through a period of inflation, rather than fluctuating there directly. Our universe, in other words, is a fluctuation but not a random one.

Emit fo Worra

This scenario, proposed in 2004 by Jennifer Chen of the University of Chicago and me, provides a provocative solution to the origin of time asymmetry in our observable universe: we see only a tiny patch of the big picture, and this larger arena is fully time-symmetric. Entropy can increase without limit through the creation of new baby universes.

Best of all, this story can be told backward and forward in time. Imagine that we start with empty space at some particular moment and watch it evolve into the future and into the past. (It goes both ways because we are not presuming a unidirectional arrow of time.) Baby universes fluctuate into existence in both directions of time, eventually emptying out and giving birth to babies of their own. On ultralarge scales, such a multiverse would look statistically symmetric with respect to time—both the past and the future would feature new universes fluctuating into life and proliferating without bound. Each of them would experience an arrow of time, but half would have an arrow that was reversed with respect to that in the others.

The idea of a universe with a backward arrow of time might seem alarming. If we met someone from such a universe, would they remember the future? Happily, there is no danger of such a rendezvous. In the scenario we are describing, the only places where time seems to run backward are enormously far back in our past—long before

FAQs, PART III

Is the multiverse theory testable?
The idea that the universe stretches far beyond what we can see is not really a theory—it is a prediction made by certain theories of quantum mechanics and of gravity. Admittedly, it is a prediction that is hard to test. But all theories of physics force us to go beyond what we can directly see. For instance, our current best model for the origin of cosmic structure, the inflationary universe scenario, requires us to understand the conditions even before inflation.

[RETROCHRONOLOGY]

The History of the Observable Universe

Here is a timeline of important events in the history of our observable universe, according to conventional cosmology:

- Space is empty, featuring nothing but a tiny amount of vacuum energy and an occasional long-wavelength particle formed via fluctuations of the quantum fields that suffuse space.
- High-intensity radiation suddenly sweeps in from across the universe, in a spherical pattern focused on a point in space. When the radiation collects at that point, a "white hole" is formed.
- The white hole gradually grows to billions of times the mass of the sun, through accretion of additional radiation of ever decreasing temperature.
- Other white holes begin to approach from billions of light-years away. They form a homogeneous distribution, all slowly moving toward one another.
- The white holes begin to lose mass by ejecting gas, dust and radiation into the surrounding environment.
- The gas and dust occasionally implode to form stars, which spread themselves into galaxies surrounding the white holes.
- Like the white holes before them, these stars receive inwardly directed radiation. They use the energy from this radiation to convert heavy elements into lighter ones.
- Stars disperse into gas, which gradually smooths itself out through space; matter as a whole continues to move together and grow more dense.
- The universe becomes ever hotter and denser, eventually contracting all the way to a big crunch.

Needless to say, this is not the usual way in which we describe the history of the universe—it is the conventional sequence of events told backward in time. But the laws of physics work equally well run forward or backward in time. Thus, this sequence is as legitimate as the usual one. It serves the purpose of driving home just how unlikely the entire history of our observable universe really is. *—S.M.C.*

our big bang. In between is a broad expanse of universe in which time does not seem to run at all; almost no matter exists, and entropy does not evolve. Any beings who lived in one of these time-reversed regions would not be born old and die young—or anything else out of the ordinary. To them, time would flow in a completely conventional fashion. It is only when comparing their universe to ours that anything seems out of the ordinary—our past is their future, and vice versa. But such a comparison is purely hypothetical, as we cannot get there and they cannot come here.

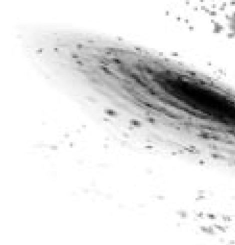
As of right now, the jury is out on our model. Cosmologists have contemplated the idea of baby universes for many years, but we do not understand the birthing process. If quantum fluctuations could create new universes, they could also create many other things—for example, an entire galaxy. For a scenario like ours to explain the universe we see, it has to predict that most galaxies arise in the aftermath of big bang–like events and not as lonely fluctuations in an otherwise empty universe. If not, our universe would seem highly unnatural.

But the take-home lesson is not any particular scenario for the structure of spacetime on ultralarge scales. It is the idea that a striking feature of our observable cosmos—the arrow of time, arising from very low entropy conditions in the early universe—can provide us with clues about the nature of the *unobservable* universe.

As mentioned at the beginning of this article, it is nice to have a picture that fits the data, but cosmologists want more than that: we seek an understanding of the laws of nature and of our particular universe in which everything makes sense to us. We do not want to be reduced to accepting the strange features of our universe as brute facts. The dramatic time asymmetry of our observable cosmos seems to be offering us a clue to something deeper—a hint to the ultimate workings of space and time. Our task as physicists is to use this and other clues to put together a compelling picture.

If the observable universe were all that existed, it would be nearly impossible to account for the arrow of time in a natural way. But if the universe around us is a tiny piece of a much larger picture, new possibilities present themselves. We can conceive of our bit of universe as just one piece of the puzzle, part of the tendency of the larger system to increase its entropy without limit in the very far past and the very far future. To paraphrase physicist Edward Tryon, the big bang is easier to understand if it is not the beginning of everything but just one of those things that happens from time to time.

Other researchers are working on related ideas, as more and more cosmologists are taking seriously the problem posed by the arrow of time. It is easy enough to observe the arrow—all you have to do is mix a little milk into your coffee. While sipping it, you can contemplate how that simple act can be traced all the way back to the beginning of our observable universe and perhaps beyond. ■



DON DIXON

MORE TO EXPLORE

Time's Arrow and Archimedes' Point: New Directions for the Physics of Time. Huw Price. Oxford University Press, 1996.

Spontaneous Inflation and the Origin of the Arrow of Time. Sean M. Carroll and Jennifer Chen. Submitted on October 27, 2004. **www.arxiv.org/abs/hep-th/0410270**

Dark Energy and the Preposterous Universe. Sean M. Carroll in *Sky & Telescope,* Vol. 109, No. 3, pages 32–39; March 2005. Available online at **www.preposterousuniverse.com/writings/skytel-mar05.pdf**