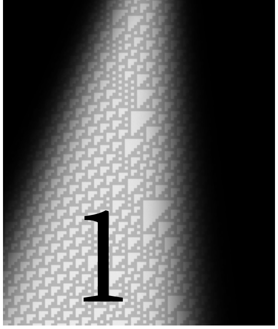
EXCERPTED FROM

STEPHEN WOLFRAM
A NEW KIND OF SCIENCE

SECTION 1.1

An Outline of Basic Ideas



The Foundations for a New Kind of Science

An Outline of Basic Ideas

Three centuries ago science was transformed by the dramatic new idea that rules based on mathematical equations could be used to describe the natural world. My purpose in this book is to initiate another such transformation, and to introduce a new kind of science that is based on the much more general types of rules that can be embodied in simple computer programs.

It has taken me the better part of twenty years to build the intellectual structure that is needed, but I have been amazed by its results. For what I have found is that with the new kind of science I have developed it suddenly becomes possible to make progress on a remarkable range of fundamental issues that have never successfully been addressed by any of the existing sciences before.

If theoretical science is to be possible at all, then at some level the systems it studies must follow definite rules. Yet in the past throughout the exact sciences it has usually been assumed that these rules must be ones based on traditional mathematics. But the crucial realization that led me to develop the new kind of science in this book is that there is in fact no reason to think that systems like those we see in nature should follow only such traditional mathematical rules.

Earlier in history it might have been difficult to imagine what more general types of rules could be like. But today we are surrounded by computers whose programs in effect implement a huge variety of rules. The programs we use in practice are mostly based on extremely complicated rules specifically designed to perform particular tasks. But a program can in principle follow essentially any definite set of rules. And at the core of the new kind of science that I describe in this book are discoveries I have made about programs with some of the very simplest rules that are possible.

One might have thought—as at first I certainly did—that if the rules for a program were simple then this would mean that its behavior must also be correspondingly simple. For our everyday experience in building things tends to give us the intuition that creating complexity is somehow difficult, and requires rules or plans that are themselves complex. But the pivotal discovery that I made some eighteen years ago is that in the world of programs such intuition is not even close to correct.

I did what is in a sense one of the most elementary imaginable computer experiments: I took a sequence of simple programs and then systematically ran them to see how they behaved. And what I found—to my great surprise—was that despite the simplicity of their rules, the behavior of the programs was often far from simple. Indeed, even some of the very simplest programs that I looked at had behavior that was as complex as anything I had ever seen.

It took me more than a decade to come to terms with this result, and to realize just how fundamental and far-reaching its consequences are. In retrospect there is no reason the result could not have been found centuries ago, but increasingly I have come to view it as one of the more important single discoveries in the whole history of theoretical science. For in addition to opening up vast new domains of exploration, it implies a radical rethinking of how processes in nature and elsewhere work.

Perhaps immediately most dramatic is that it yields a resolution to what has long been considered the single greatest mystery of the natural world: what secret it is that allows nature seemingly so effortlessly to produce so much that appears to us so complex.

It could have been, after all, that in the natural world we would mostly see forms like squares and circles that we consider simple. But in fact one of the most striking features of the natural world is that across a vast range of physical, biological and other systems we are continually confronted with what seems to be immense complexity. And indeed throughout most of history it has been taken almost for granted that such complexity—being so vastly greater than in the works of humans—could only be the work of a supernatural being.

But my discovery that many very simple programs produce great complexity immediately suggests a rather different explanation. For all it takes is that systems in nature operate like typical programs and then it follows that their behavior will often be complex. And the reason that such complexity is not usually seen in human artifacts is just that in building these we tend in effect to use programs that are specially chosen to give only behavior simple enough for us to be able to see that it will achieve the purposes we want.

One might have thought that with all their successes over the past few centuries the existing sciences would long ago have managed to address the issue of complexity. But in fact they have not. And indeed for the most part they have specifically defined their scope in order to avoid direct contact with it. For while their basic idea of describing behavior in terms of mathematical equations works well in cases like planetary motion where the behavior is fairly simple, it almost inevitably fails whenever the behavior is more complex. And more or less the same is true of descriptions based on ideas like natural selection in biology. But by thinking in terms of programs the new kind of science that I develop in this book is for the first time able to make meaningful statements about even immensely complex behavior.

In the existing sciences much of the emphasis over the past century or so has been on breaking systems down to find their underlying parts, then trying to analyze these parts in as much detail as possible. And particularly in physics this approach has been sufficiently successful that the basic components of everyday systems are by now completely known. But just how these components act together to produce even some of the most obvious features of the overall behavior we see has in the past remained an almost complete mystery. Within the framework of the new kind of science that I develop in this book, however, it is finally possible to address such a question.

From the tradition of the existing sciences one might expect that its answer would depend on all sorts of details, and be quite different for different types of physical, biological and other systems. But in the world of simple programs I have discovered that the same basic forms of behavior occur over and over again almost independent of underlying details. And what this suggests is that there are quite universal principles that determine overall behavior and that can be expected to apply not only to simple programs but also to systems throughout the natural world and elsewhere.

In the existing sciences whenever a phenomenon is encountered that seems complex it is taken almost for granted that the phenomenon must be the result of some underlying mechanism that is itself complex. But my discovery that simple programs can produce great complexity makes it clear that this is not in fact correct. And indeed in the later parts of this book I will show that even remarkably simple programs seem to capture the essential mechanisms responsible for all sorts of important phenomena that in the past have always seemed far too complex to allow any simple explanation.

It is not uncommon in the history of science that new ways of thinking are what finally allow longstanding issues to be addressed. But I have been amazed at just how many issues central to the foundations of the existing sciences I have been able to address by using the idea of thinking in terms of simple programs. For more than a century, for example, there has been confusion about how thermodynamic behavior arises in physics. Yet from my discoveries about simple programs I have developed a quite straightforward explanation. And in biology, my discoveries provide for the first time an explicit way to understand just how it is that so many organisms exhibit such great complexity. Indeed, I even have increasing evidence that thinking in terms of simple programs will make it possible to construct a single truly fundamental theory of physics, from which space, time, quantum mechanics and all the other known features of our universe will emerge.

When mathematics was introduced into science it provided for the first time an abstract framework in which scientific conclusions could be drawn without direct reference to physical reality. Yet despite all its development over the past few thousand years mathematics itself has continued to concentrate only on rather specific types of abstract systems—most often ones somehow derived from arithmetic or geometry. But the new kind of science that I describe in this book introduces what are in a sense much more general abstract systems, based on rules of essentially any type whatsoever.

One might have thought that such systems would be too diverse for meaningful general statements to be made about them. But the crucial idea that has allowed me to build a unified framework for the new kind of science that I describe in this book is that just as the rules for any system can be viewed as corresponding to a program, so also its behavior can be viewed as corresponding to a computation.

Traditional intuition might suggest that to do more sophisticated computations would always require more sophisticated underlying rules. But what launched the whole computer revolution is the remarkable fact that universal systems with fixed underlying rules can be built that can in effect perform any possible computation.

The threshold for such universality has however generally been assumed to be high, and to be reached only by elaborate and special systems like typical electronic computers. But one of the surprising discoveries in this book is that in fact there are systems whose rules are simple enough to describe in just one sentence that are nevertheless universal. And this immediately suggests that the phenomenon of universality is vastly more common and important—in both abstract systems and nature—than has ever been imagined before.

But on the basis of many discoveries I have been led to a still more sweeping conclusion, summarized in what I call the Principle of Computational Equivalence: that whenever one sees behavior that is not obviously simple—in essentially any system—it can be thought of as corresponding to a computation of equivalent sophistication. And this one very basic principle has a quite unprecedented array of implications for science and scientific thinking.

For a start, it immediately gives a fundamental explanation for why simple programs can show behavior that seems to us complex. For like other processes our own processes of perception and analysis can be thought of as computations. But though we might have imagined that such computations would always be vastly more sophisticated than those performed by simple programs, the Principle of Computational Equivalence implies that they are not. And it is this equivalence between us as observers and the systems that we observe that makes the behavior of such systems seem to us complex.

One can always in principle find out how a particular system will behave just by running an experiment and watching what happens. But the great historical successes of theoretical science have typically revolved around finding mathematical formulas that instead directly allow one to predict the outcome. Yet in effect this relies on being able to shortcut the computational work that the system itself performs.

And the Principle of Computational Equivalence now implies that this will normally be possible only for rather special systems with simple behavior. For other systems will tend to perform computations that are just as sophisticated as those we can do, even with all our mathematics and computers. And this means that such systems are computationally irreducible—so that in effect the only way to find their behavior is to trace each of their steps, spending about as much computational effort as the systems themselves.

So this implies that there is in a sense a fundamental limitation to theoretical science. But it also shows that there is something irreducible that can be achieved by the passage of time. And it leads to an explanation of how we as humans—even though we may follow definite underlying rules—can still in a meaningful way show free will.

One feature of many of the most important advances in science throughout history is that they show new ways in which we as humans are not special. And at some level the Principle of Computational Equivalence does this as well. For it implies that when it comes to computation—or intelligence—we are in the end no more sophisticated than all sorts of simple programs, and all sorts of systems in nature.

But from the Principle of Computational Equivalence there also emerges a new kind of unity: for across a vast range of systems, from simple programs to brains to our whole universe, the principle implies that there is a basic equivalence that makes the same fundamental phenomena occur, and allows the same basic scientific ideas and methods to be used. And it is this that is ultimately responsible for the great power of the new kind of science that I describe in this book.

Relations to Other Areas

Mathematics. It is usually assumed that mathematics concerns itself with the study of arbitrarily general abstract systems. But this book shows that there are actually a vast range of abstract systems based on simple programs that traditional mathematics has never considered. And because these systems are in many ways simpler in construction than most traditional systems in mathematics it is possible with appropriate methods in effect to go further in investigating them.

Some of what one finds are then just unprecedentedly clear examples of phenomena already known in modern mathematics. But one also finds some dramatic new phenomena. Most immediately obvious is a very high level of complexity in the behavior of many systems whose underlying rules are much simpler than those of most systems in standard mathematics textbooks.

And one of the consequences of this complexity is that it leads to fundamental limitations on the idea of proof that has been central to traditional mathematics. Already in the 1930s Gödel's Theorem gave some indications of such limitations. But in the past they have always seemed irrelevant to most of mathematics as it is actually practiced.

Yet what the discoveries in this book show is that this is largely just a reflection of how small the scope is of what is now considered mathematics. And indeed the core of this book can be viewed as introducing a major generalization of mathematics—with new ideas and methods, and vast new areas to be explored.

The framework I develop in this book also shows that by viewing the process of doing mathematics in fundamentally computational terms it becomes possible to address important issues about the foundations even of existing mathematics.