The Outer Limits of Reason

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

Library of Congress Cataloging-in-Publication Data

Yanofsky, Noson S., 1967–

The outer limits of reason : what science, mathematics, and logic cannot tell us / Noson S. Yanofsky.

pages cm

Includes bibliographical references and index.

ISBN 978-0-262-01935-4 (hardcover : alk. paper)

1. Knowledge, Theory of. 2. Science—Philosophy. 3. Mathematics—Philosophy. I. Title.

Q175.32.K45Y36 2013

001.01—dc23

2012050531

10 9 8 7 6 5 4 3 2 1

Please email all comments and criticisms to noson@sci.brooklyn.cuny.edu.

d_r0

Table of Contents

Preface

> The more we learn about the world, and the deeper our learning, the more conscious, specific, and articulate will be our knowledge of what we do not know, our knowledge of our ignorance.[1]
> —Karl Popper

> A man's got to know his limitations.
> —Harry Callahan, *Magnum Force* (1973)

> Everything should be made as simple as possible, but not simpler.
> —Attributed to Albert Einstein

With understanding comes ambivalence. Once we know something, we often find it boring and trite. On the other hand, the mysterious and unknown fascinates us and holds our attention. That which we do not know or understand is what interests us, and what we *cannot* know intrigues us even more. This book explores topics that reason tells us we cannot know because they are beyond reason.

Many books convey the amazing facts that science, mathematics, and reason have revealed to us. There are also books that cover topics that science, mathematics, and reason have not yet fully explained. This book is a little different. Here we study what science, mathematics, and reason tell us *cannot* be revealed. What cannot be predicted or known? What will never be understood? What are the limitations of computers, physics, logic, and our thought processes? What is beyond the bounds of reason? This book aims to answer some of these questions and is full of ideas that challenge our deep-seated beliefs about the universe, our rationality, and ourselves.

Along the way we will study simple computer problems that would take trillions of centuries to solve; consider perfectly formed English sentences that have no meaning; learn about different levels of infinity; leap into the bizarre and wonderful world of the quantum; discuss specific problems that computers can never solve; befriend butterflies that bring about blizzards; ponder particles that simultaneously dance at different parties; hear about paradoxes and self-referential paradoxes; see what relativity theory tells us about our naive notions of space, time, and causality; understand Gödel's famous theorems about the limitations of logic; discover certain problems in mathematics and physics that are impossible to solve; explore the very nature of science, mathematics, and reason; wonder why the universe seems perfect for human beings; and examine the complex relationship between our mind, reason, and the physical universe. We will also attempt to peek beyond the borders of reason and see what, if anything, is out there. These and many other fascinating topics will be presented in a way that is clear and comprehensible.

While exploring these various limitations in diverse areas, we will see that many of the limitations have a similar pattern. These patterns will be investigated in order to better understand the structure of reason and its limits.

This book is not a compendium of all the diverse examples in which limitations of reason are found. Rather, our goal is to understand why these boundaries arise and why reason cannot extend beyond them. Several representative limitations in each area are selected and discussed in depth.

Rather than just listing the limitations, I aim to explain them or at least provide the intuition of why a particular area is beyond reason. It is important to realize that this book is not meant to be speculative or to have a New Age orientation. Nor is it a history book in which I gloss over the meaning of ideas in order to focus on their chronological development. This is a popular science book that will gradually and clearly explain the ideas presented.

Since I accept Stephen Hawking's dictum that every equation halves the number of readers, very few equations are found in this book. However, I do believe in the power of diagrams, charts, and graphs to simplify complex ideas. My goal is clarity.

Each chapter deals with a different area: science, mathematics, language, philosophy, and so on. These chapters are arranged from concrete to abstract. I start with simple problems of everyday language and move on to straightforward philosophical questions, ending with the abstract world of mathematics. For the most part, the chapters are independent of each other and can be read in any order. Readers are encouraged to begin with topics that most interest them. (The unifying theme of self-referential paradoxes is found in chapters 2, 4, 6, and 9.)

Acknowledgments

In a sense, this book is a collaborative effort with my friends and colleagues in the Computer and Information Science Department at Brooklyn College. They have read the chapters, corrected my mistakes, chided me when I was being silly, and encouraged me when I was stuck. They have given me the warm intellectual environment that made this book possible. I thank them!

Many Brooklyn College faculty members were kind enough to read and comment on several of these chapters: Jonathan Adler, David Arnow, George Brinton, Samir Chopra, Jill Cirasella, Dayton Clark, Eva Cogan, Jim Cox, Scott Dexter, Keith Harrow, Danny Kopec, Yedidyah Langsam, Matthew Moore, Rohit Parikh, Simon Parsons, Michael Sobel, Aaron Tenenbaum, and Paula Whitlock. Their comments have made this a far better book.

Some chapters were used in a course given at Brooklyn College. Great benefit was derived from student participation and the give-and-take of a classroom dialog. I thank the students for listening, arguing, helping, and correcting. Many students at Brooklyn College and the Graduate Center of the City University of New York read and commented on earlier drafts: Firat Atagün, Can Başkent, Hubert Bennett, Greg Benson, Be Birchall, Raila C. Brejt, Fatemeh Choopani, Simon Dexter, Aline Elmann, Madelene Feingold, Terri Grosso, Miriam Gutherc, Jay Jankelewicz, Matthew P. Johnson, Joel Kammet, Tatiana Kedel, Wai Khoo, Karen Kletter, Erdal Köse, Michael Lampis, Shalva Landy, Holly Lo Voi, Jon Lo Voi, Matthew Meyer, Valia Mitsou, Jordi Navarrette, Shoshana Neuburger, Hadassah Norowitz, Nicole Reilly, Artur Sahakyan, Connor Savage, Angela Shatashwili, Alex Sverdlov, Stanislav Turzhavskiy, Fredda

Weinberg, and Karol Wysocki. I am indebted to them all.

Numerous other people looked over parts of the book and have been helpful: Ros Abramsky, Samson Abramsky, Marcia Barr, Michael Barr, Rebecca Barr, Adam Brandenburger, Richard Churchill, Melvin Fitting, Leopold Flatto, Robert J. Fogelin, Chaim Goodman-Strauss, Ariel Halpert, Eliyahu Hershfeld, Ellen Hershfeld, Faigy Hershfeld, Pinchas Hershfeld, Shai Hershfeld, Yitzchok Hershfeld, Michael Hicks, Joshua Honigwachs, Roman Kossak, Klaas Landsman, André Lebel, Raphael Magarik, Camille Martin, Jolly Mathen, Rochel Moskowitz, Larry Moss, Naftoli Neuburger, Janos Pach, Carol Parikh, Suri Raber, N. Raja, Barbara Rifkind, Andrei Rodin, Ariel Ropek, Evan J. Siegel, Michael Vitz, Sharon Yanofsky, and Mark Zelcer. Their criticism, comments, and helpful ideas are deeply appreciated.

I thank Robbert Dijkgraaf for permission to use his artwork in figure 8.6. I also thank C. Goodman-Strauss for permission to use many of his diagrams in section 9.3. My beautiful and talented daughter Hadassah was very helpful with some of the diagrams.

James DeWolf, Marc Lowenthal, Marcy Ross, and the whole MIT Press team have been very helpful in getting this book into shape. Thank you.

Karen Kletter remains the world's greatest editor and proofreader. Thank you, Karen!

However, at the end of the day, I am the sole cause of any errors that may remain.

Several other, more general debts should be acknowledged. I am grateful to my friend and research partner Ralph Wojtowicz of Baker Mountain Research Corporation for supporting my other research while I was working on this book.

In the spring of 1987 I had a chance encounter with Dr. Avi

Rabinowitz on a street corner in Jerusalem. Avi is a brilliant physicist bristling with creativity and enthusiasm. We eventually became traveling companions and good friends. There are few topics that Avi cannot discuss in depth. Conversations with him usually proceed at the speed of light. We've had many intense conversations while climbing mountains in Greece and watching ridiculous science fiction movies. He is a true mentor and friend. Every page in this book contains ideas that I have discussed with him over the years. (He would probably disagree with most of what I wrote.) His influence is immense and I am forever appreciative.

Over the past few years, three people who enriched my life have passed away. They added much to my education and hence much to this book.

During my senior year at Brooklyn College, Professor Chaya Gurwitz supervised me in a research project, thereby introducing me to the rigors of higher mathematics and computer science. She taught me how to read an academic paper, to put my ideas into action, and to analyze the results. This experience piqued my interest in attending graduate school. She graciously invited me to her home for many meals, where I also became friends with her husband and eight wonderful children. She continued to guide me as a graduate student, as a teacher, as a colleague, and as a person until her untimely passing in 2008. I am truly indebted to her.

My dissertation advisor, *mon maître*, Alex Heller, was a distinguished professor in the Department of Mathematics at the Graduate Center of the City University of New York. He was a kind, gentle man. Although I graduated in 1996, we continued to meet once or twice a week until a few days before his sad passing in 2008. (In a sense, he gave me twelve years of postdoctoral research.) Our conversations meandered from mathematics, to

politics, to morality, to philosophy, to history, and so on. He was an amazing genius and his range of knowledge was astonishing; in mathematics, however, it was particularly striking. While speaking with him, one got the impression that he had a magnificently clear vision of the entire structure of mathematics before him. It was a privilege to study under him and to be befriended by him.

A word must be said about Professor Heller's unique method of mentoring. After learning much from him during two years of coursework, I was disinvited from attending any more of his classes. He said that I had gained enough from him. From then until his passing, even though we talked about mathematics constantly, he never *taught* me a scintilla of mathematics. My job was to present my work or what I was studying. His job was to find flaws in my presentation or my understanding. He would harangue me about developing a correct definition, indicate where my proofs had failed, and point out when I was not exact. Although a gentleman—always with a kind word—his method of mentoring was intimidating and disheartening, to say the least. He never articulated this sink-or-swim philosophy. Nevertheless, he had a valid point: it was my job to learn mathematics and I had to struggle with it on my own. I am forever grateful for his confidence in me and for the independence he insisted I develop. He was the greatest of teachers.

I had the benefit of having a world-class mathematician as a neighbor. Professor Leon Ehrenpreis lived a few blocks from me and I took advantage of it by visiting him on a regular basis. My Friday-night visits were always greeted with a warm, welcoming smile. Besides being a first-rate mathematician, he was also a rabbinic scholar, marathon runner, handball player, classical pianist, and father of eight. This Renaissance man's breadth of knowledge was truly astounding. I have many fond memories of sitting at his kitchen table chatting about the subtleties of Hebrew grammar, the edge-of-the-wedge theorem, raising children, the consequences of the Kochen-Specker theorem, the role of cows in the Book of Genesis, hypergeometric functions, and many other topics. Professor Ehrenpreis was blessed with the most pleasant disposition and always had a kind, encouraging word. I learned much from him. He passed away in August of 2010.

All three are painfully missed.

This book is dedicated to my wife, Shayna Leah, whose warmth and loving support made this work possible, and to my daughters, Hadassah and Rivka, who fill our home and hearts with laughter and joy. My love and gratitude toward them are limitless.

1

Introduction

Human reason, in one sphere of its cognition, is called upon to consider questions, which it cannot decline, as they are presented by its own nature, but which it cannot answer, as they transcend every faculty of the mind.[1]
—Immanuel Kant (1724–1804)

As the circle of light increases, so does the circumference of darkness.[2]
—Attributed to Albert Einstein

Zorba: Why do the young die? Why does anybody die?
Basil: I don't know.
Zorba: What's the use of all your damn books if they can't answer that?
Basil: They tell me about the agony of men who can't answer questions like yours.
Zorba: I spit on this agony!
—*Zorba the Greek* (1964)

A civilization can be measured by how much progress its science and technology have made. The more advanced their science and technology are, the more advanced their civilization is. Our civilization is deemed more advanced than what we call primitive societies because of all the technological progress we have made. In contrast, if an alien civilization visited Earth, we would be considered primitive, almost by definition, since they have mastered interstellar space travel while we have not. The reason for using science and technology as a measuring stick is that these activities are the only aspect of culture that builds on itself. What was done by one generation is used by the next generation. This was expressed nicely by one of the greatest scientists of all time, Isaac Newton (1643–1727), who is quoted as saying, “If I have seen further it is only by standing on the shoulders of giants.” This constant accumulated progress makes science a good measuring stick to compare civilizations. In contrast to science and technology, other areas of culture, such as the arts, human relations, literature, politics, morality, and so on, do not build on themselves.[3]

Another way to measure a civilization is by the extent to which it has banished unscientific and irrational ideas. We are more advanced today because we have cast alchemy into the wastebasket of silly dreams and study only chemistry. Centuries of treatises on astrology have been deemed nonsense while we retain our study of astronomy. As a civilization progresses, it subjects its beliefs and mythologies to logical analysis and disregards what is not within the bounds of reason.

The tool a civilization uses to make this progress is reason. Rationality and reason are the methodologies used by a society to advance. When a culture acts reasonably it will progress. When it deviates from reason, or steps beyond the limits of reason, it stagnates or regresses.

Reason comes in many forms. In broad (and perhaps inexact) terms, science is the language that we use to describe and predict the physical and measurable universe. The more abstract mathematics can be split into two areas: applied mathematics is the language of science, and pure mathematics is the language of reason. Logic is also a language of reason. Since science, technology, reason, rationality, logic, and mathematics are all intimately connected to each other, much of what I say about one will usually be true about all. At times I will just use the word *reason* to describe them all.

Philosophers have reflected and argued for centuries about what humans can and cannot know. The branch of philosophy that deals with human knowledge and its limitations is called *epistemology*. While the ideas of such philosophers are fascinating, their work will not be our central focus. Instead, we will be interested in what scientists, mathematicians, and current researchers have to tell us about the limits of human knowledge and reason.

One of the most amazing aspects of modern science, mathematics, and rationality is that they have matured to the level where they are able to see their own limits. As of late, scientists and mathematicians have joined philosophers in discussing the limitations of man's ability to know the world. These scientific limitations of reason are the central subject of this book.

The following is a cute little puzzle that gives a taste of what it means for reason to describe a limitation.[4] The puzzle is loads of fun, is worth pondering, and is also strongly recommended as a challenge at any cocktail party. Take a normal 8-by-8 chessboard and some dominoes that are of size 2-by-1. Try to cover the chessboard with the dominoes. There are sixty-four squares on the chessboard and each domino covers two squares, so thirty-two dominoes will be needed. There are millions of ways to perform this task. Figure 1.1 shows how we might start the process.

Figure 1.1
Covering a chessboard with dominoes

That was pretty easy. Now let's try something a little more challenging. Put two queens on the opposite corners of the chessboard. Try to cover all the squares except the ones with queens, as in figure 1.2. There are sixty-two squares that need to be covered, which means thirty-one dominoes will be required. Try it!

Figure 1.2
Covering a chessboard minus two opposing corners

After trying this problem for a while and not being able to cover every square, you might consider showing it to others—in particular, puzzle fans. They will have a similar experience. You might want to get a computer to work on the problem since a machine can quickly try many possibilities. There are millions, if not billions, of possible ways to try to start placing the dominoes on the board. Nevertheless, there is no way anyone or any computer will ever finish this task.

The reason why this simple problem of placing thirty-one dominos on a chessboard seems so hard is because *it cannot be done.* It is not a hard problem; it is an *impossible* problem. It is actually easy to explain why. Every domino is 2-by-1 and hence must cover a black-and-white square on the chessboard. The original board in figure 1.1 had thirty-two black squares and thirty-two white squares that needed to be covered. There was total symmetry on

(For example, "Second-hand smoke is not so bad for you." "Democracy is not always the best form of government.")

• An apparently absurd or self-contradictory statement or proposition, or a strongly counter-intuitive one, which investigation, analysis or explanation may nevertheless prove to be well-founded or true. (For example, "In the long-run, the stock market is a bad place to invest." "Standing is more strenuous than walking.")

To us, the most important definition will be

• An argument, based on (apparently) acceptable premises and using (apparently) valid reasoning, which leads to a conclusion that is against sense, logically unacceptable, or self-contradictory.

Such paradoxes will be our main concern. Here one has a premise or makes an assumption and using valid logic derives a falsehood. We might envision this paradox or derivation as

assumption ⇒ falsehood.

Since falsehoods cannot occur and since our derivation followed valid logic, the only conclusion is that our assumption was not true. In a way, the paradox is a test to see if an assumption is a legitimate addition to reason. If one can use valid reason and the assumption to derive a falsehood, then the assumption is wrong. The paradox shows that we have stepped beyond the boundaries of reason. A paradox in this sense is a pointer to an incorrect view. It points to the fact that the assumption is wrong. Since the assumption is wrong, it cannot be added to reason. This is a limitation of reason.

The type of falsehood that we will mostly encounter is a contradiction. By a contradiction I mean a fact that is shown to be both true and false. This is written as

assumption ⇒ contradiction.

Since the universe does not have contradictions, there must be something wrong with the assumption. For example, in chapter 6, we will see that if we assume that a computer can perform a certain task, then we can derive a contradiction about certain computers. Since there are no contradictions about physical objects like computers, there must be something wrong with our assumption.

Such paradoxes work the same way as a commonly found mathematical proof. A "proof by contradiction" or in Latin, *reductio ad absurdum* ("reduction to the absurd"), is as follows. If you want to show that some statement is true, simply assume that the statement is false and derive a contradiction:

statement is false ⇒ contradiction.

Since contradictions are not permitted in the exact world of mathematical reasoning, it must be that the assumption was incorrect, and the statement is, in fact, true. A simple example is the mathematical proof that the square root of 2 is not a rational number (section 9.1). If we assume that the square root of 2 *is* a rational number, then we derive a contradiction. From this we conclude that the square root of 2 is not a rational number. In section 4.3 I show that if we assume two particular sets are the same size, we can derive a contradiction. From this we conclude that one of the sets is larger than the other. Proofs by contradiction are ubiquitous.

One need not derive a full-fledged contradiction for a paradox. All that is needed is to derive a fact that is different from observation or simply false:

assumption ⇒ false fact.

Once again, because we derived something false, our assumption must be in error. Zeno's paradoxes are examples of this type (section 3.2). Zeno assumes something and then proceeds to show that movement is impossible. Anyone who has ever walked down the street knows that movement occurs all the time and hence the assumption is false. The difficulty with Zeno's paradoxes is to identify the bad assumptions.

Many times paradoxes arise and highlight previously hidden assumptions. It could be that these assumptions are so deep within us that we do not even consider them (for example, that space is continuous and not discrete, or that physical objects have exact definitions). Such paradoxes will be a challenge to our intuitions about the universe we live in. By showing that our intuitions are false, we can disregard them and be propelled forward. The American philosopher Willard Van Orman Quine (1908–2000) eloquently wrote:

> The argument that sustains a paradox may expose the absurdity of a buried premise or of some preconception previously reckoned as central to physical theory, to mathematics or to the thinking process. Catastrophe may lurk, therefore, in the most innocent-seeming paradox. More than once in history the discovery of paradox has been the occasion for major reconstruction at the foundation of thought.[5]

This method of exploring paradoxes and looking for their assumptions will be one of our focuses throughout the book.

Particular types of paradoxes play a major role in the tale we tell. Self-referential paradoxes are paradoxical situations that come from a system where the objects of the system can deal with / handle / manipulate themselves. The classic example of a self-referential paradox is the so-called *liar paradox*. Consider the

English sentence:

"This sentence is false."

If the sentence is true, then the sentence is, in fact, false because it says so. If the sentence is false, then since the sentence expresses its own falsehood, the sentence is true. This is a genuine contradiction. The problem arises from the fact that English sentences have the ability to describe true and false statements about themselves. For example, "This sentence has five words" is a legitimate English sentence that expresses something true about itself. In contrast, "This sentence has six words" is a false statement about itself. We will see that whenever a system can discuss properties about itself, a paradoxical situation can occur. We will find that language, thought, sets, logic, math, and computers are all systems with the ability to deal with themselves. Within each of these areas, the potential for self-reference will lead to paradoxes and hence some type of limitation. The amazing fact is that although these areas are very different, the form of the paradoxes are the same.

Another method of describing a limitation is by piggybacking on an already established limitation. Before I explain what this is all about, let's discuss some mountain climbing. Mount Everest is 29,000 feet high and Mount McKinley is "only" 20,000 feet high. The following fact seems obvious: if you can climb Mount Everest, then you can most definitely (*a fortiori*) climb Mount McKinley. We write this as

climbing Everest ⇒ climbing McKinley.

If you are able to climb Mount McKinley, you would feel great pride. We write this as

climbing McKinley ⇒ pride.

Putting the two implications together, we get

climbing Everest ⇒ climbing McKinley ⇒ pride,

which leads to the obvious conclusion that if you are able to climb Mount Everest, you would feel great pride. Now let us look at the dark side of mountain climbing. Suppose your doctor told you that bad things might happen to you if you try to climb Mount McKinley. We write this as

climbing McKinley ⇒ bad.

This is expressing a limitation of your abilities: you should not climb Mount McKinley. Combining this implication with the first one, gives us

climbing Everest ⇒ climbing McKinley ⇒ bad.

This states the obvious fact that if you should refrain from climbing Mount McKinley, then you most definitely should refrain from climbing Mount Everest. In other words, the obvious implication that

climbing Everest ⇒ climbing McKinley

can be used to transfer or piggyback a known limitation about climbing Mount McKinley into a limitation about climbing Mount Everest. I use these simple ideas in the following pages.

Now let us use this intuition about mountain climbing to understand the general concept of one limitation piggybacking on another limitation. Imagine that a limitation was established by a contradiction as follows:

assumption A ⇒ contradiction.

unusual for a person to express a desire to be thin while having another piece of cake.[6]

When we meet a paradox in the physical world and derive a contradiction, we know that there must be something wrong with the assumption of the paradox. However, when we meet a contradiction in the realm of human thought or in human language, then we need not abandon the assumption. More subtlety is possible. Why not permit the contradiction? Consider the liar paradox discussed earlier. Why not simply say that the sentence

This sentence is false.

is both true and false or perhaps meaningless? It is only an English sentence and many English sentences express contradictions. Similarly, the belief

This belief is false.

is both true and false. Why not permit such contradictory beliefs in our already-confused minds?

The relationship between the contradiction-free universe and our feeble human minds and languages raises many more interesting questions. How is it that the human mind can understand any part of the universe? How can a language formulated by human beings describe the universe? Why does science work? Why is mathematics so good at describing science and the universe? Do the laws of science have an external existence or are they only in our mind? Can there be a final description of the universe—that is, will science ever complete its mission and end? Are the truths of science and mathematics time dependent or culturally dependent? How can human beings tell when a scientific theory is true? As Albert Einstein wrote, "The eternal mystery of the world is its comprehensibility.[7] These and a host of other questions from the philosophy of science and mathematics are addressed in chapter 8.

Between the contradiction-free universe and the contradiction-laden human mind, a landscape full of vagueness exists:

- A person who stands in the doorway of a room is both in the room and not in the room.
- How many hairs does a man have to lose in order to be considered bald? Depending on which way the wind blows, he is sometimes considered bald and sometimes considered not bald.
- Is 42 a small or large number?

Human beings use vague ideas all the time. Our mindset and our concomitant human language are full of vague statements:

- Sometimes we say people in a doorway are in the room and sometimes we say they are not in the room.
- We call certain people with a few hairs bald and others not bald.
- If our bank account contains only $42, we say that 42 is a small number, but if we are talking about the number of diseases a person has, 42 is a large number.

Because vague ideas are outside the pristine world of science and mathematics, we cannot rely on some of the usual tools in addressing these ideas. Vagueness plays a major role in our discussions in chapter 3.

As a slight aside, special types of jokes are of interest for our discussion. We have seen that paradoxes are ways of showing that one has gone too far with reason. Violating a paradox means you stepped beyond the boundaries of reason and entered the land of the absurd. There are jokes that also play on the fact that we are taking reason too far. Such jokes take logic and reason to places where they were not intended. They start off with concepts that are well understood and then go farther or beyond their usual meaning. Consider the following:

- Woody Allen cheated on his metaphysics exam by looking into the soul of the boy sitting next to him.

- Steven Wright said he would kill for a Nobel Peace Prize.

- Groucho Marx didn't care to belong to any club that would have him as a member.

In all of these jokes, normal ideas are taken too far. Cheating on an exam, desiring a Nobel Peace Prize, or resigning your club membership in disgust are all common ideas. However, these great thinkers have taken these usual concepts where they do not belong: to the silly and ridiculous.

Even puns fall into this category. A pun is a joke where the meaning of a word or phrase is taken into an area where it was not intended:

- "Have you heard about the guy whose whole left side was cut off? He's all right now."

- "I'm reading a book about antigravity. It's impossible to put down."

- "Did you hear about the par-a-dox? . . . Doctor Shapiro and Doctor Miller."

Groan! (Sorry. The only thing worse than a pun is an analysis of a pun. Let us move on.)

I close this introduction with a few questions about the nature of reason and its limitations. Read the book with these questions in mind. I return to these issues in the last chapter and perhaps get closer to the answers using some of the ideas presented in the book.

I would be remiss in writing a book titled *The Outer Limits of Reason* without giving a definition of *reason*. After all, how can we say something is beyond the limits of reason if we do not define reason? What is a reasonable process to determine facts? Are there different levels of reason? How do we draw the line between alchemy and chemistry? Between astrology and astronomy? Why are some actions deemed reasonable and others not? Why does it make sense to check your blood pressure while it is ludicrous to check your horoscope? What thought processes are reasonable and will avoid contradictions?

The *Oxford English Dictionary* gives sixteen classes of definitions for the word *reason*. The definition closest to the one we want is the following: "The power of the mind to think and form valid judgments by a process of logic; the mental faculty which is used in adapting thought or action to some end; the guiding principle of the mind in the process of thinking. Freq. contrasted with *will, imagination, passion,* etc. Often personified." But this definition just raises more questions. What is a "valid judgment"? When is something a logical process as opposed to an illogical process? When is thinking part of the will and when is it reason? This definition is unsatisfying. Other purported definitions are not much better.

There is something self-referential in our entire enterprise. We are using reason to find limitations of reason. If reason is limited, how are we to use reason to discover those limitations? What are the limits to our limit-showing abilities?

Let's hold these questions in abeyance and return to them in chapter 10, when we conclude our explorations of the limits of reason.

Further Reading

Other books that discuss limitations of reason are Barrow 1999, Dewdney 2004, and Poundstone 1989. Sorensen 2003 is a wonderful history of paradoxes.

The liar paradox is found in many different forms. For example, we can denote a sentence L_1 and then say that L_1 asserts its own falsehood:

L_1: L_1 is false.

Again, if L_1 is true, then it is false. And if L_1 is false, then it is true. Other variations of the liar paradox have sentences that are not directly self-referential. Consider the following two sentences:

L_2: L_3 is false.

L_3: L_2 is true.

If L_2 is true, then L_3 is false, which would mean that "L_2 is true" is false and hence L_2 is false. In contrast, if L_2 is false, then L_3 is true and L_3 asserts that L_2 is true. Buzz! That's a contradiction.

It is important to note that just because sentences refer to themselves and their falsehoods does not mean there is a contradiction. Consider these two sentences:

L_4: L_5 is false.

L_5: L_4 is false.

Let's assume that L_4 is false. Then L_5 is true and L_4 is false. Similarly, if you start with the premise that L_4 is true, you get that L_5 is false, and hence L_4 is true. Neither assumption leads you to a contradiction.

There are many other forms of the liar paradox:

• <u>The only underlined sentence on this page is a total lie.</u>

• **The boldface sentence on this page is a blatant falsehood.**

• The sentence after the boldface sentence on this page is not true.

Are they true or false?

The liar paradox has been around for over 2,500 years and philosophers have devised many different ways of avoiding such contradictions. Some philosophers try to avoid these linguistic paradoxes by saying that the liar sentences are neither true nor false. After all, not every sentence is true or false. Questions such as “Your place or mine?” and commands such as “Go directly to jail!” are neither true nor false. One usually thinks of declarative sentences like “Snow is white” as either true or false, but the liar sentences show that there are some declarative sentences that are also neither true nor false.

There are those who say that the sentence “This sentence is false” is not even grammatically correct. After all, what does “This sentence” refer to? If it refers to something, we should be able to replace “This sentence” with whatever it refers to. Let’s give it a try:

“This sentence is false” is false.

This is grammatically correct and it might be true or false. But it is not self-referential and not equivalent to the original liar sentence. This is similar to the sentence

“This sentence is false” has four words.

which is true, while

"This sentence is false" has five words.

is false. It would be nice to have a grammatically correct English sentence that is a self-referential paradox. W. V. O. Quine came up with a clever way around these problems. Consider the following *Quine's sentence*:

"Yields falsehood when preceded by its quotation"

yields falsehood when preceded by its quotation.

First notice that this is a legitimate English sentence. The subject is the phrase in quote marks and the verb is *yields*. Now, let us ask ourselves if it is true. If it is true, then when you attach the subject to the rest of the sentence, as we did, we get falsehood. So the sentence is false. In contrast, what if the sentence is false? That means that when you attach the subject to the sentence, you do not get a falsehood; rather, you get a true sentence. So if you assume that Quine's sentence is false, you derive that it is true. This is a grammatically correct English sentence that is self-contradictory.

Another potential solution to paradoxical sentences is to restrict language so as to avoid such sentences. Some have said that language should be stratified into different levels. They have declared that sentences cannot talk about other sentences of their own level or higher. For example, at the lowest level there will be sentences like "Grass is green" and "My pen is blue." The next level will be sentences about sentences on the lowest level. So we might have

"Grass is green" is an obvious sentence.

or

"My pen is blue" has four words in it.

One goes on to higher-level phrases like

""My pen is blue' has four words in it" is a dumb fact.

By restricting the types of sentences, we will be avoiding sentences of the form

The sentence in italics on this page is grammatically correct.

This is a sentence dealing with itself and hence is a sentence on its own level. It is declared not kosher—that is, not a legitimate part of language. Every sentence is only permitted to talk about sentences that are "below" it. If a sentence does talk about a sentence that is on its own level, that sentence is proclaimed meaningless. This stratification will ensure that there are no self-references and hence no contradictions. With such restrictions in place, linguists are fairly certain that they have banned most paradoxical linguistic sentences. However, this solution is somewhat artificial. Common human language has always dealt with some type of self- reference without problem:

- Someone says, "Oh! I am groggy today and I do not know what I am talking about." Is he aware of saying this sentence?

- Carly Simon sings a song with the lyrics "You're so vain, you probably think this song is about you." But this song *is* about him!

- "Every rule has an exception except one rule: this one."

- "Never say 'never'!"

- "The only rule is that there is no rule."

In all of these cases—and many more—human language is violating the restriction of only dealing with sentences that are "below" it. In each case, a sentence discusses itself. And yet, somehow, all these examples are a legitimate part of human language.

Another possible solution to paradoxical sentences was mentioned in chapter 1, namely, human language is a product of the human mind and, as such, subject to contradictions. Human language is not a perfect system that is free of discrepancies (in contrast to perfect systems like mathematics, science, logic, and the physical universe). Rather, we should simply accept the fact that human language is faulty and has contradictions. This seems reasonable to me.

2.2 Self-Referential Paradoxes

The cause of the problem with the liar paradox is that language can be used to describe language. In particular, one can have a sentence that discusses its own truthfulness. This ability of language to describe language is a form of self-reference. Paradoxes that arise from such self-reference are the subject of this section. While these paradoxes are not linguistic paradoxes per se, they are similar to the liar paradox and will help us understand the true nature of self-reference.

The British philosopher Bertrand Russell described a delightful little paradox that has come to be known as the *barber paradox.* Imagine a small isolated village in the Austrian Alps that has only one barber. Some villagers shave themselves and some go to the barber. Everyone in the village abides by the following rule: all those who do not shave themselves must go to the only barber and all those who do shave themselves do not go to the barber. This seems like a pretty innocuous rule. After all, if they can save some money by shaving themselves, why go to the barber? And if they go to the barber, why shave themselves? Now, simply ask yourself:

French is not French ⇒ *French* is heterological,

so too

heterological is not heterological ⇒ *heterological* is heterological.

We have come to the conclusion that *heterological* is heterological if and only if it is not heterological. Buzz! This is a contradiction and troublesome.

We can again envision this self-referential paradox as figure 2.2.

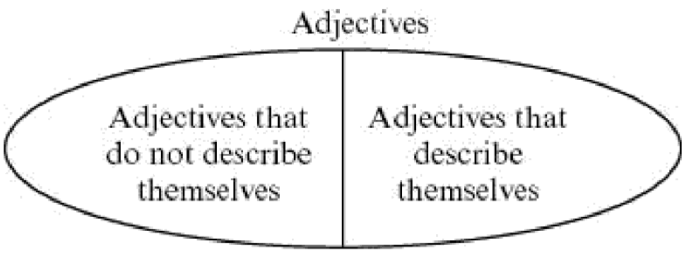

Figure 2.2
Which subset does *heterological* belong to?

This paradox also seems to have a simple solution: there is no word *heterological*, or if the word does exist, it has no meaning. We saw that if one defines *heterological*, then we come to a contradiction. This is similar to saying that the village in the barber paradox does not exist.

However, we cannot simply solve all problems by waving our hand and declaring that the word *heterological* does not exist or has no meaning. The problem is too deeply rooted in the very nature of language. Rather than dealing with the word *heterological*, consider the related adjective phrase "not true of itself." Simply ask if the phrase "not true of itself" is true of itself. It is true if and only if it is not true. Are we simply to posit that "not true of itself" is not a legitimate adjectival phrase? There are no problems with any of the words in the phrase. There is nothing about the phrase that is weird like the word *heterological.* Nevertheless, we come to a contradiction if we use it.

The *reference-book paradox* is very similar to the heterological paradox. A reference book is a book that lists books in different categories. There are many reference books that list books of many different types. There are reference books that list antique books, anthropology books, books about Norwegian fauna, and so on. Certain reference books list themselves. For example, if one were to publish a reference book of all books published, that reference book would contain itself. There are also certain reference books that would not list themselves. For example, a reference book on Norwegian fauna would not list itself. Consider the reference book that lists all reference books that do not list themselves. Now ask yourself the following simple question: Does this book list itself? With a little thought, it is easy to see that this book lists itself if and only if it does not list itself. We conclude that no such reference book with such a rule for its content can exist. (I leave to the reader the task of drawing a diagram similar to figures 2.1 and 2.2 for this paradox.)

Bertrand Russell used the barber paradox to explain a more serious paradox called *Russell's paradox.* This is more abstract than the other self-referential paradoxes we saw and is worth pondering. Consider different sets or collections of objects. Some sets just contain elements and some sets contain other sets. For example, one can look at a school as a set containing different grades, where each grade is the set of students in the grade. Some sets even contain copies of themselves. The set of all sets described in this book contains itself. The set of all sets with more than five elements contains itself. There are, of course, many sets that do not contain themselves. For instance, consider the set of all red apples. This does not contain itself since a red apple is not a set. Russell would like us to consider the set R of all sets that do not contain themselves. Now pose the following question:

Does R contain itself?

If R does contain itself, then, by definition of what belongs to R, it is not contained in R. If, on the other hand, R does not contain itself, then it satisfies the requirement of belonging to R and is contained in R. We have a contradiction. This can be visualized in figure 2.3.

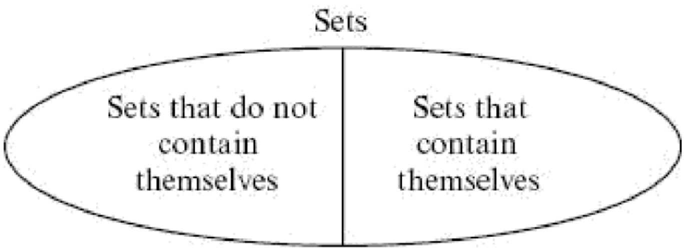

Figure 2.3
Which part contains R?

This paradox is usually "solved" by positing that the collection R does not exist—that is, that the collection of all sets that do not contain themselves is not a legitimate set. And if you do deal with this illegitimate collection, you are going beyond the bounds of reason. Why should one not deal with this collection R? It has a perfectly good description of what its members are. It certainly looks like a legitimate collection. Nevertheless, we must restrict ourselves in order to steer clear of contradictions. The obvious (and seemingly reasonable) notion that for every clearly stated description there is a collection of those things that satisfy that description is no longer obvious (or reasonable). For the clearly stated description of "red things," there is a nice collection of all red things. However, for the seemingly clear description of "all sets that do not contain themselves," there is no collection with this property. We must adjust our conception of what is obvious.[5]

Russell's paradox should be contrasted with the other paradoxes. There are simple solutions to the barber paradox and the reference-book paradox: those physical objects simply do not exist. And there is a simple solution to the heterological paradox: human language is full of contradictions and meaningless words. We are, however, up against a wall with Russell's paradox. It is hard to say that the set R simply does not exist. Why not? It is a well-defined idea. A collection is not a physical object, nor is it a human-made object. It is simply an idea. And yet this seemingly innocuous idea takes us out of the bounds of reason.

The liar paradox was summarized by one sentence:

This sentence is false.

It can also be summarized by the following description:

The sentence that denies itself.

Similarly, the other four self-referential paradoxes can be summarized by the following four descriptions:

- "The villager who shaves everyone who does not shave themselves."
- "The word that describes all words that do not describe themselves."
- "The reference book that lists all books that do not list themselves."
- "The set that contains all sets that do not contain themselves."

As you can see, all these descriptions have the exact same structure (as do figures 2.1 through 2.3). Every time there is self-reference, there are possibilities for contradictions. Such contradictions will have to be avoided and will require a limitation. We explore such limitations throughout the book.

Before moving on to the next section, there is an interesting result that demands further thought. One might think that every language paradox has some form of self-reference. That is, there must be some chain of reasoning that is circular and returns to where it started. This was the common belief until Stephen Yablo came up with a clever paradox called *Yablo's paradox*. Consider the following infinite sequence of sentences:

K_1 K_i is false for all $i > 1$

K_2 K_i is false for all $i > 2$

K_3 K_i is false for all $i > 3$

⋮ ⋮

K_m K_i is false for all $i > m$

K_{m+1} K_i is false for all $i > m + 1$

⋮ ⋮

K_n K_i is false for all $I > n$

⋮ ⋮

to $1^3 + 12^3$ but it is also equal to $9^3 + 10^3$. Since 1729 is the smallest number for which this can be done, 1729 is an "interesting" number.[7]

This tale brings to light the *interesting-number paradox*. Let's take a tour through some small whole numbers. 1 is interesting because it is the first number. 2 is the first prime number. 3 is the first odd prime. 4 is a number with the interesting property that $2 \times 2 = 4 = 2 + 2$. 5 is a prime number. 6 is a perfect number—that is, a number whose sum of its factors is equal to itself (i.e., $6 = 1 \times 2 \times 3 = 1 + 2 + 3$, etc.). The first few numbers have interesting properties. Any number that does not have an interesting property should be called an "uninteresting number." What is the smallest uninteresting number? The smallest uninteresting number is an interesting number. We are in a quandary.

What went wrong here? The contradiction came about because we thought we could split all numbers into two groups: interesting numbers and uninteresting numbers. This is false. There is no way to define what an interesting number is. It is a vague term and we cannot say when a number is interesting and when it is uninteresting.[8] "Interesting" is a feeling that a person gets sometimes and hence is a subjective property. We cannot make a paradox out of such a subjective property.

A more serious and related paradox is called the *Berry paradox*. The key to understanding this paradox is that in general the more words one uses in a phrase, the larger the number one can describe. The largest number that can be described with one word is 90. 91 would demand more than one word. Two words can describe ninety trillion. Ninety trillion + 1 is the first number that demands more than two words. Three words can describe ninety trillion trillion. The next number (ninety trillion trillion + 1) would demand more than three words. Similarly, the more letters in a word, the larger the number you can describe. With three letters, you can describe the number 10 but not 11.

Let us stick to number of words. Call a phrase that describes numbers and has fewer than eleven words a *Berry phrase*. Now consider the following phrase:

the least number not expressible in fewer than eleven words.

This phrase has ten words and expresses a number, so it should be a Berry phrase. However, look at the number it purports to describe. The number is not supposed to be expressible in fewer than eleven words. Is this number expressible in eleven words or less? This is a real contradiction.

We may also talk about other measures of how complicated an expression is. Consider

the least number not expressible in fewer than fifty syllables.

This phrase has fewer than fifty syllables. Another phrase,

the least number not expressible in fewer than sixty letters,

has fifty-nine letters. Do these descriptions describe numbers or not? And if they do describe numbers, which ones? They describe a certain number if and only if they do not describe that number. But why not? Each certainly seems like a nice descriptive phrase.

Yet another interesting paradox about describing numbers is *Richard's paradox*. Certain English phrases describe real numbers between 0 and 1. For example,

- "pi minus 3" = 0.14159
- "the chance of getting a 3 when a die is thrown" = 1/6

- “pi divided by 4” = 0.785

- “the real number between 0 and 1 whose decimal expansion is 0.55555” = 0.55555

Call all such phrases *Richard phrases*. We are going to describe a paradoxical sentence. Rather than just stating the long sentence, let us work our way toward it. Consider the phrase

> the real number between 0 and 1 that is different from any Richard phrase.

If this described a number, it would be paradoxical since the phrase would describe a number and yet it would not be a Richard phrase. However, there are many real numbers that are different from all Richard phrases. Which one is it? The problem is that this phrase does not really describe an exact number. Let us try to be more exact. The set of Richard phrases are a subset of all English phrases, and as such, they can be ordered like names in a telephone book. We can first order all Richard phrases of one word, then the phrases of two words, and so on. With such an ordered list we can talk about the *n*th Richard sentence. Now consider

> the real number between 0 and 1 whose *n*th digit is different from the *n*th digit of the *n*th Richard phrase.

This is just showing how the number described is different from all the Richard phrases, but it still does not describe an exact number. The number described by the forty-second Richard number might have an 8 as the forty-second digit. From this, we know that our phrase cannot have an 8 in the forty-second position. But should our number have a 9 or 6 in that position? Let us be exact:

> the real number between 0 and 1 defined by its *n*th digit being 9 minus the *n*th digit of the *n*th Richard phrase.

That is, if the digit is a 5, this phrase will describe a 4. If the digit is an 8, this phrase will describe a 1. And if the digit is a 9, this phrase will describe a 0. This phrase is a legitimate English phrase that precisely describes a number between 0 and 1, yet it is different from every single Richard phrase. The phrase does describe a number if and only if it does not describe a number. What to do?[9]

These last two paradoxes can be seen as self-referential paradoxes. In a sense, they can be summarized by the following two descriptions:

- "the Berry phrase that is different from all Berry phrases"
- "the Richard phrase that is different from all Richard phrases"

From this point of view, they are simple extensions of the liar paradox. Self-reference is very common and we must be careful with it.

Further Reading

Many of the paradoxes can be found in places such as Quine 1966, Hofstadter 1979, 2007, Barrow 1999, and Poundstone 1989. Sorenson 2003 is a clear and well-written introduction to paradoxes. Chapter 5 of Sainsbury 2007 covers the liar paradox and other forms of self-reference. Chapter 3 of Paulos 1980 provides a humorous look at all self-referential paradoxes. Yablo's paradox is found in Yablo 1993.

A formal version of self-referential paradoxes can be found in Yanofsky 2003, which is derived from Lawvere 1969.

3

Philosophical Conundrums

> Moreover, although these opinions appear to follow logically in a dialectical discussion, yet to believe them seems next door to madness when one considers the facts. For indeed no lunatic seems to be so far out of his senses.
> —Aristotle (384–322 BC), *On Generation and Corruption*, 325a15

> All are lunatics, but he who can analyze his delusion is called a philosopher.
> —Ambrose Bierce, *The Collected Works of Ambrose Bierce*

> It depends on what the meaning of the word "is" is.
> —William Jefferson Clinton

Long before modern scientists took up the task of investigating the limits of reason, philosophers were analyzing the complexities of our world and our knowledge of it. In this chapter I explore some of the ancient and contemporary philosophical aspects of reason's limitations.

In section 3.1, I begin by discussing some very fundamental questions about concrete and abstract objects and the way we define them. In section 3.2, the very nature of space, time, and motion are analyzed using some of Zeno's paradoxes. The section ends with a short discussion of time-travel paradoxes. Section 3.3 is concerned with vagueness. Section 3.4 is centered on the very notion of knowing and having information. These sections are ax. A certain museum wanted to preserve the ax of the founding father of the United States. The ax consists of two parts: a handle and a head. As time went on, the wooden handle would rot and the metal head would rust. When needed, each of these two parts was replaced. As the years passed, the head was changed four times and the handle was replaced three times. Is it still Washington's ax? Notice that here there is no question of the change being gradual. Every time a change is made, half the parts of the ax are replaced.

Our discussion is not limited to ships and axes. A tree is lush and green in the summer and bare and brown during the winter. Mountains rise and fall. Cars and computers get refurbished. Any physical object changes over time. This is the content of Heraclitus' famous dictum that you cannot step into the same river twice. For Heraclitus, the river changes at every instant.

Physical objects are not the only things that change. Businesses, institutions, and organizations are also dynamic entities that constantly change and evolve. Barings Bank was in existence from 1762 through 1995. In that time, the owners, workers, and customers all changed. The Brooklyn Dodgers have been around since 1883. Their players, managers, owners, and fans have definitely changed. What remains the same about a baseball team? After heartlessly betraying their city of birth, the Dodgers cannot even claim that they play in the same city as they originally did. In colleges, the students change every four years. Even the professors change over the years. The only real heart and soul of a college are the beloved secretaries. But, alas, even they change. Political parties are also not immune to change. The Democratic Party was founded in the 1790s to support states' rights over federal rights, the opposite of their current platform. Everything changes!

We are not only talking about change. Rather, we are discussing what it means for an object to be that object. What does it mean for a certain institution to be that institution? When we say that a certain object changes, we mean that it had a certain property beforehand and after the change it does not. In the beginning, the ship of Theseus had planks that Theseus himself touched. At the end, there were planks that he did not touch. That is a change in the properties of the ship. Our fundamental question is: What are the core properties of the ship of Theseus? We have shown that there are no clear answers to this question.

This discussion becomes far more interesting when we stop talking about ancient ships and start talking about human beings. Every person changes over time. We grow from infants to old people. What properties does a three-year-old have in common with their eighty-three-year-old self? These philosophical questions are called the *problems of personal identity*. What are the properties that make up a particular human being? We are not the same person we were several years ago. Nevertheless, we are still considered the same person.

Philosophers usually fall into one of several camps on this question. Some thinkers push the notion that a person is essentially their body. We each have different bodies and can say that every person is identified with their body. By postulating that a human being is their body, we are subject to the same insoluble questions that we faced with the ship of Theseus and other physical objects. Our bodies are in constant flux. Old cells die and new cells are constantly being born. In fact, most of the cells in our body are replaced every seven years. This leads to hundreds of questions that philosophers have posed over the centuries. Why should a person stay in jail after seven years? After all, "he" did not perform the crime. It was someone else. Should a person own anything after seven years? The old person bought it. In what sense is a person the same after having a limb amputated? Science fiction writers are adept at discussing challenging questions like cloning, mind transfers, identical twins, conjoined twins, and other interesting topics related to the notion that a person is the same as their body. When an ameba splits, which is the original and which is the daughter? When your body loses cells it loses atoms. These atoms can go on to belong to others. Similarly, other peoples' atoms can become part of your body. What about death? We usually think in terms of the end of a person's existence when they are dead even though the body is still there. Sometimes we use sentences like "She is buried there" as if "she" were still a person. And sometimes we use sentences like "His body is buried there" as if there is a difference between "him" and his body. In short, it is problematic to say that a human being is identified with their body.

Other thinkers favor the notion that a person is really their mental state or psyche. After all, human beings are not simply their bodies. A person is more than a physical object because there is thought. To such philosophers, a person is a continuous stream of consciousness—they are memories, intentions, thoughts, and desires. This leads us to ask other insoluble questions: What if a person has amnesia? Are they the same person? Doesn't a person's personality change over time? Who is the real you: the one who is madly in love with someone or the one who is bored with the same person two months later? Literally hundreds of questions can be posed about change in a person's thoughts, memories, and desires. Again, philosophers and science fiction writers have become quite adept at describing interesting scenarios that challenge our notion of a human being as a continuous stream of mental states. These scenarios are concerned with Alzheimer's disease, amnesia, personality changes, split-brain experiments, multiple personality disorders, computers as minds, and so on. There are also many questions along the lines of the mind-body problem. How much is the mind—that characterizes a human being—independent of the brain, which is a part of the body?

One of the more interesting challenges to the position that continuity of mental states characterizes a human being is the question of *transitivity of identity*. My mental states are essentially the same as they were ten years ago. That means I am the same person I was ten years ago. Furthermore, ten years ago, my mental states were essentially the same as they were ten years earlier. Hence the person I was ten years ago is the same as the person I was twenty years ago. However, at present, I do not have similar mental states to those I had twenty years ago. So how can it be that I am the same person I was ten years ago, and that person is the same as I was twenty years ago, but I am *not* the same as I was twenty years ago?

Yet another option is that everyone has a unique soul that determines who they are. Avoiding the questions of the definition or existence of a soul, let us concentrate instead on how this answers our question of the essential nature of a human being. Assuming the existence of a soul, what is the relationship between the soul and the body? What is the relationship between a soul and a person's actions, psyche, and personality? If there is no connection, then in what sense is one soul different from another soul? How can you differentiate between souls if they have no influence over any part of you? What would the purpose of a soul be? If, on the other hand, a connection exists, then does the soul change when the body, actions, psyche, or personality changes? Is the soul in flux? If the soul does change, we are back to the same questions we had previously asked: Who is the real you? Are you the one with the soul prior to the change or are you the one with the changed soul?

Most people probably have an opinion representing some hybrid version of all three ideologies: a person is a composite of body, mind, and soul. Nevertheless, all schools of thought are somewhat problematic.

Rather than answering all the questions posed in this section, let us try to resolve the issues by meditating on why none of the questions have clearcut answers. Why is it that when we pose these questions to different people, we get so many different answers?

Examine the way people learn to recognize different objects, make definitions, and create distinctions. In the beginning, babies are bombarded with many different sensations and stimuli. As toddlers grow, they learn to recognize objects in the world. For example, when they see a shiny silver thing covered with brown gooey stuff coming toward them, they have to learn that it is applesauce on a spoon and that they should open their mouth. By learning to recognize that the physical stimulus of silver covered with brown gooey stuff is applesauce, they are able to handle life better. Human beings need to classify objects. We learn how to tell things apart and determine when they are the same. We learn that an object still exists even when it is out of sight ("object permanence"). Children learn after a while to recognize their mother. A few months later, they learn that even though she is wearing makeup—that is, even when she looks different—she is still the same person. Children have to learn that their mother is the same even when she is wearing perfume and smells totally different. Here toddlers are acting as philosophers and learning how to deal with different questions of personal identity. With all these skills, children are imposing order and structure on the complicated world they have entered. Before these skills are mastered, they are showered with an incomprehensible stream of stimuli and sensations. With these classification abilities the children can comprehend and start to control their environment. If they fail to learn the classification skills, they will be overburdened with external stimuli and unable to deal with their surroundings.

With enough sophistication, children also learn to classify abstract entities. For example, they might learn what it means to be a family. Their mother is a family member. Their father and siblings are also part of the family. What about first cousins? Second cousins? These are a little vague. Sometimes they are part because there is no such thing as objective aesthetics. It's a matter of taste. Similarly, whether changing a plank of a ship changes the ship cannot be given a definitive answer because there is no such thing as an objective ship of Theseus.

One can safely argue with what is posed here and claim that objects really do have an existence outside of the human mind and that what children are learning to do is classify and name those entities. They are learning to associate names of entities with physical stimuli. Weathered, rotting wood that looks like a ship in the port of Athens should be associated with the "ship of Theseus." This ideology might be called *extreme Platonism* (see figure 3.1). Classical Platonism is the belief that abstract entities have real existence outside of the human mind. The number 3 really exists. There is an exact idea when one refers to the U.S. government. An idea of a chair exists. However, classical Platonism takes no stand about concrete physical entities. In contrast, extreme Platonism is the belief that even a concrete physical object has some type of unchanging platonic entity associated with it. To someone who maintains this position, some platonic notion of "ship-of-Theseus-ness" exists and when a question is posed about a change to the ship of Theseus, all one has to do is somehow connect to the platonic notion and see if the changed ship still satisfies the definition. Extreme Platonism demands a fairly advanced metaphysics, and we cannot really say that as metaphysics, it is true or false. It is impossible to show that no such abstract entity exists. Nevertheless, as with all metaphysical notions, there is no real reason to posit such an existence.[4] If you claim that a name or a definition of an object is some type of "tag" on the object, then we can ask where the tag is. Why is it that people disagree so vociferously about the tag on the ship of Theseus?

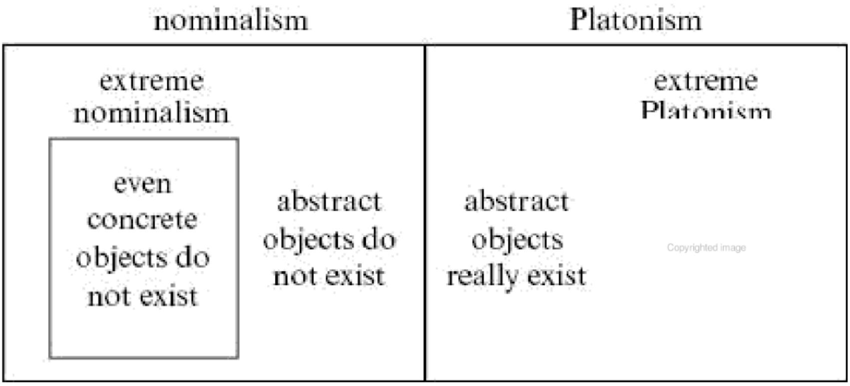

Figure 3.1
Different philosophical schools of thought

In this chapter, I am promoting an idea that might be termed *extreme nominalism.* The philosophical position of classic nominalism is the belief that abstract entities really do not exist outside the human mind. To a nominalist, abstract ideas like the number 3, the idea of the U.S. government, and the idea of a chair or "chairness" do not really exist outside the minds of those who discuss them. Have you ever met a 3? Can you stub your toe against a 3? Can you point to the U.S. government? A classic nominalist would say that these entities only exist in the human mind. Since we share a similar education and social structures, we can banter about these different names and concepts with our neighbors. However, a classical nominalist does not have a position on the question of concrete physical entities.

Extreme nominalism takes nominalism a step further. It is the belief that even physical objects exist *as those physical objects* only in name. They do not have an external existence outside of a human mind. A particular chair is a chair because we call it a chair, not because it has properties of being a chair. The ship of Theseus is whatever people call the ship of Theseus. There are no exact, agreed-on definitions of the ship of Theseus. I believe that extreme nominalism is correct because of the fact that there is so much disagreement about what constitutes a particular object. If there were exact definitions, presumably people would know about them. Another reason for believing in (extreme) nominalism is that any form of Platonism demands unnecessarily complicated metaphysics. Why do we need the supposed existence of an abstract entity or “tag” for every physical object? Such abstract entities serve no purpose.

From the view afforded by extreme nominalism, it becomes apparent that the reason we cannot answer questions about the ship of Theseus or changes to human beings has nothing to do with linguistic limitations. It is not that we lack the right words or definitions of these concepts. There is also no epistemological problem—that is, it is not a lack of knowledge of the exact definition of the real ship of Theseus. Nor is it a problem of having some type of deeper knowledge of the ship of Theseus beyond its physical stimuli.[5] Rather, we are dealing with a question of existence. In philosophical parlance this is an ontological problem. A real ship of Theseus need not exist.

It is interesting to note that with extreme nominalism, certain abstract objects, such as the number 42, have a clearer existence than physical objects such as ships. After all, we all agree about the many different properties of the number 42. If you take 42 and you subtract 1, you get 41 rather than 42. This is in stark contrast to subtracting planks from a ship.

I have shown that the ship of Theseus is part of our culturally constructed universe. There are other objects in this constructed universe such as Mickey Mouse and unicorns. In fact, more people know about Mickey Mouse than about Theseus’ silly boat. Our friendly mouse is introduced to nearly every child, whereas only classics majors, philosophy majors, and privileged readers of this book know about Theseus. Furthermore, one can go to Disney World and actually see a physical manifestation of Mickey. You can even stub your toe against him (such actions are not recommended). In contrast, at present, we cannot find any trace of Theseus' ship in the port of Athens. We are left with the obvious question: In what way is the ship more existent than Mickey Mouse?

The resolution of the problems presented in this section is a challenge to the usual view of the universe. Most people believe that there are certain objects in the universe and that human minds call those objects by names. What I am illustrating here is that those objects do not really exist. What do exist are physical stimuli. Human beings classify and name those different stimuli as different objects. However, the classification is not always strict and vagueness prevails.[6]

3.2 Hangin' with Zeno and Gödel

Zeno of Elea (about 490–430 BC) was a great philosopher who was a student of Parmenides (early fifth century BC). Being a devoted student, Zeno promoted and protected his teacher from all criticism. Parmenides had the philosophical and mystical belief that the world was "one" and that change and motion were merely illusions that a person could see through with enough training. To demonstrate that Parmenides' ideas are correct, Zeno proposed several thought experiments or paradoxes that showed that it is illogical to actually believe that the world is a "plurality" and not "one," or that change and motion actually happen. In this section I will concentrate on four of those thought experiments that demonstrate that motion is an illusion. Since motion occurs within space and time, Zeno's paradoxes will challenge our intuition of these obvious concepts.

Unfortunately, most of Zeno's original writings have been lost. Our knowledge of the paradoxes largely comes from people who wanted to prove him wrong. Aristotle briefly sets up some of

Zeno's ideas before knocking them down. Because Zeno's ideas were given short shrift, it is not always clear what his original intentions were. This should not deter us since our central interest is not what Zeno actually said; rather, we are more interested to know if something is wrong with our intuition and how it can be adjusted. These ideas should not be taken lightly. They have bothered philosophers for almost 2,500 years. Regardless of whether one agrees with Zeno or not, he cannot be ignored.

The first and easiest of Zeno's paradoxes of motion is the *dichotomy paradox.* Imagine an intelligent slacker waking up in the morning. He tries to get from his bed to the door in his room (see figure 3.2).

Figure 3.2
Zeno's dichotomy paradox

To get the whole way to the door, he must reach the halfway point. Once he reaches that point, he still must go a quarter of the way more. From there he has an eighth of the way to go. At every point, he must still go halfway more. It seems that this slacker will never be able to reach the door. In other words, if he does want to get to the door, he will have to complete an infinite process. Since one cannot complete an infinite process in a finite amount of time, the slacker never gets to the door.

Our slacker can further justify his laziness with more logical reasoning. To reach the door, one has to go halfway. To reach the halfway point, he must first get to the quarter-way point, and

A better solution is to say that the problem with Zeno's reasoning is that he assumes that space is continuous. That means that space looks like the real-number line and is infinitely divisible —that is, between every two points lies an infinite number of points. Only with this assumption can one describe the dichotomy paradox. In contrast, imagine that we are watching the slacker go to the door in an old-fashioned television made up of millions of little pixels. Then as he is moving, he is crossing the pixels. He crosses half of the pixels and then he crosses half of the rest of the pixels. Eventually the TV slacker will be one pixel away from the door and then he will be at the door. There are no half pixels to cross. A pixel is either crossed or not crossed. On the TV screen there is no problem with the slacker getting to his destination and Zeno's paradox evaporates. Maybe we can say the same thing with the real world. Perhaps space is made up of discrete points each separated from its neighbor and that between any two points there is at most a finite number of other points. In that case we would not have to worry about the dichotomy paradox. If we assume such a discrete space, then we can understand why our lazy slacker makes it to the door: he only has a finite number of points to cross. At a certain point, the intervals could no longer be split into two. Objects move in this type of space by going from one discrete point to the next without passing between them.

In the language of chapter 1, we can say that this is a paradox because we are assuming that space is continuous:

Space is continuous ⇒ movement is impossible.

Since there is definitely movement in this world, and our assumption led us to a false fact, we conclude that space is not continuous. Rather, it is discrete, or separated into little "space atoms."

Such ideas of discrete space are familiar to people who study quantum mechanics.[8] Physicists discuss something called *Planck's* *length*, which is equal to 1.6162×10^{-35} meters. Something smaller in length cannot be measured. To some extent, nothing smaller than that exists. Physicists assure us that objects go from one Planck's length to another. In high school chemistry it is taught that electrons fly in shells around a nucleus of an atom. When energy is added to an atom, the electrons make a "quantum leap" from one shell to the next. They do not pass in between the shells. Perhaps our lazy slacker also makes such quantum leaps and hence can finally reach the door.

Let us reconsider figure 3.3. The square is infinitely divvied up as illustrated. But this is only possible if we think of the square as a mathematical object. In mathematics every real number that represents a distance can be split into two, hence we can continue chopping forever. In contrast, let us think of the square as a piece of paper. We can start cutting paper into smaller and smaller pieces using finer and finer scissors. This will work for a while, but eventually we will reach the atomic level where no further cutting will be possible. This is true for any physical object made of atoms. We are forced to conclude that the square depicted in figure 3.3 is not a good model for the physics associated with the paper square. The real numbers can be infinitely divided but the paper cannot be. What Zeno is forcing us to do is to ask the question of whether space (which is not made of atoms) can be infinitely divvied up. If it can be, the slacker will not reach his goal. If it cannot be, there must be discrete "space atoms," and continuous real-number mathematics is not a proper model for space.[9]

We cannot, however, be so flippant about asserting that space is discrete and not continuous. The world certainly does not look discrete. Movement has the feel of being continuous. Much of mathematical physics is based on calculus, which assumes that the real world is infinitely divisible. Outside of some quantum theory and Zeno, the continuous real numbers make a good model for the physical world. We build rockets and bridges using mathematics that assumes that the world is continuous. Let us not be so quick to abandon it.[10]

Zeno's second paradox of motion is the story of *Achilles and the Tortoise*. Achilles was the ancient Greek version of the modern D.C. Comics character The Flash and was the fastest runner in town. One day he had a race with a slow Tortoise. To make the race more interesting (and because Achilles had a warm heart), Achilles gave the Tortoise a head start, as shown in the top line of figure 3.4.

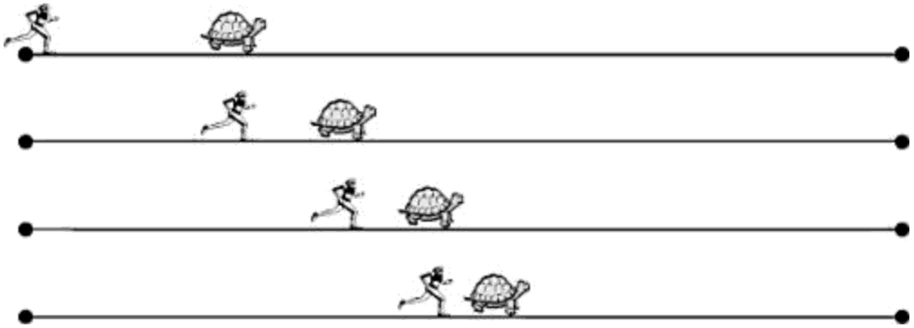

Figure 3.4
Achilles not catching up to the Tortoise

The problem is, in order for Achilles to overtake the Tortoise, he must first pass the point where the Tortoise started (as in the second line in figure 3.4). At that point, the Tortoise has already moved further. Once again, in order for Achilles to overtake the Tortoise, he must get to the point where the Tortoise moved. At each point, Achilles is getting closer and closer to the pesky Tortoise, but he will never be able to reach him, let alone beat him.

Again there is a mathematical analogy to this. In calculus we say that the limit of $1/x$ as x goes to infinity is zero. That is, the larger x gets, the closer $1/x$ gets to zero. Since infinity is not a number, x

can never get to infinity and $1/x$ can never get to zero. But the concept of a limit makes it meaningful. Similarly, the distance between Achilles and the Tortoise will never really be zero but the *limit* of the distance does get to zero. Again, we can find fault with this analogy. The concept of a mathematical limit is a type of trick. For no finite number will $1/x$ actually equal zero and at no time period will Achilles actually reach the Tortoise.

This paradox would also melt away if we assume that the racetrack is made out of discrete points. The fact that Achilles runs faster than the Tortoise simply means that he covers more of the discrete points in the same time. So eventually Achilles will overtake the Tortoise. Discrete space would answer the paradox, but again, we have to be careful. We should abandon the notion of continuous space with great trepidation since that mathematical model works so well in general physics.

In the third paradox, Zeno is not interested in determining whether a motion can be completed. Here he attacks the very idea of any motion whatsoever. In the *arrow paradox*, we are asked to think of an arrow flowing through space. At every instant in time, the arrow is in some particular position. If we think of time as a continuous sequence of "nows" that separates "pasts" from "futures," then for each "now," the arrow is in one particular position. At each point in time the arrow is in a definitive position and not moving. The question is, when does the arrow move? If it does not move at each of the "nows," when does it?

This paradox can also be solved if we introduce discrete ideas into the mix. But rather than saying that space is discrete, here we say that time is discrete. At each separate point in time there is no motion. But time leaps from one separated point to another and motion happens at that instant leap. In other words, say that time is discrete and not continuous. We do not see these magical leaps for the same reason we think we see continuous motion when we are watching a movie. In fact, a movie is made out of numerous discrete frames and there is no motion between them. Because the separate time points are so close to each other and there are so many, there is an illusion of continuity.

This paradox basically describes the following derivation:

Time is continuous $\Rightarrow$ movement is impossible.

Once more, since it is an obvious fact that movement is possible, we conclude that time is not continuous but is discrete.

There is also a mathematical analogy to this paradox. Consider the real-number line. Think of the real line as time. Each point of the real line corresponds to a "now." And yet each "now" has no thickness. In ninth grade you learned that the real line is made of an infinite number of points. Each point has zero length. So how can a finite line be made of points that do not have any thickness? Zero times anything is zero. Was your ninth-grade teacher lying to you? Does the real-number line make sense? Should we abandon it?

Again, there is a problem abandoning the notion of continuous time for discrete time. Modern physics and engineering are based on the fact that time is continuous. All the equations have a continuous-time variable usually denoted by t. And yet, as Zeno has shown us, the notion of continuous time is illogical.

The fourth and final paradox against motion is the *stadium paradox*. Zeno wants us to imagine three marching bands as in figure 3.5.

The two great achievements of twentieth-century physics are relativity theory and quantum theory.[11] These two revolutionary sciences essentially describe most of the phenomena in the physical universe. Relativity theory deals with gravity and large objects, while quantum theory deals with the other forces and small objects. However, these two theories are in conflict with each other. One of the main reasons for their conflict is that relativity theory considers space and time to be continuous while quantum theory believes space and time to be discrete. For the most part, since the theories deal with different realms, the conflict does not bother us. Nonetheless, the conflict is apparent with certain phenomena such as black holes, which are termed the "edge of space." Since we cannot have conflicting physical theories, it must be that we do not know the final story. The jury is still out regarding the structure of space and time.

The most amazing aspect of Zeno's paradoxes is that they are 2,500 years old and they deal with such simple topics. What is the nature of space, time, and motion? It is doubtful that we have heard the last of our Elean friend.

Since we are discussing the relationship of space, time, and logic, let us talk about time-travel paradoxes. We first have to ask ourselves what it means to travel back[12] through time. What would it mean for me to go back to the Continental Congress held in Philadelphia in 1776, in order to witness the signing of the Declaration of Independence? If I am miraculously transported back there and see the signing, then the very fact that I am in the room on that hot day in July means that it is not the original Continental Congress. After all, I was not there during the original. In other words, if there were 150 people present at the original Continental Congress, when I go back there will be 151 people present. That is not the original. It is a major difference between what I was transported to and the original. What exactly am I being transported to? One thing is certain: not to the Continental Congress of 1776.[13] This conundrum shows how hard it is to understand the very basic concepts of time travel.

Be that as it may, let us imagine for a moment that we understood what traveling through time actually means, and furthermore, let us imagine that such a process was, in fact, possible. If time travel was possible, a time traveler might go back in time and shoot his bachelor grandfather, ensuring that the time traveler was never born. If he was never born, then he could not have shot his grandfather. Homicidal behavior is not necessary to achieve such paradoxical results. The time traveler might just ensure that his parents never have children,[14] or he might simply go back in time and make sure that he does not enter the time machine. These actions would entail a contradiction and hence cannot happen. The time traveler should not shoot his own grandfather (moral reasons notwithstanding) because if he shoots his own grandfather, he will not exist and will not be able to travel back in time to shoot his own grandfather. So by performing an action he is ensuring that the action cannot be performed. The event is self-referential. Usually, one event affects other events, but here an event affects itself. In the language of chapter 1, we are showing that

Time travel ⇒ contradiction.

Since the universe does not permit contradictions, we must somehow avoid this paradox. Either time travel is impossible, or even if it was possible, one would still not be able to cause a contradiction by killing an earlier version of oneself. Which impossibility should we prefer?

Albert Einstein's theory of relativity tells us that the usual way that we conceive the universe makes time travel impossible. In 1949, Einstein's friend and Princeton neighbor, Kurt Gödel, did some moonlighting as a physicist and wrote a paper on relativity theory. Gödel constructed a mathematical way of looking at the universe in which time travel would be possible. In this "Gödel universe," it would be very hard, but not impossible, to travel back in time. Gödel, the greatest logician since Aristotle, was well aware of the logical problems of time travel. The mathematician and writer Rudy Rucker tells of an interview with Gödel in which Rucker asks about the time-travel paradoxes. The relevant passage is worth quoting: "Time-travel is possible, but no person will ever manage to kill his past self.' Gödel laughed his laugh then, and concluded, 'The *a priori* is greatly neglected. Logic is very powerful.'"[15] Gödel replies that the universe simply will not allow you to kill your past self. Just as the barber paradox shows that certain villages with strict rules cannot exist, so too the physical universe will not allow you to perform an action that will cause a contradiction.

This leads us to even more mind-blowing questions. What would happen if someone took a gun back in time to shoot an earlier version of himself? How will the universe stop him? Will he not have the free will to perform the dastardly deed? Will the gun fail to shoot? If the bullet fires and is properly aimed, will the bullet stop short of his body? It is indeed bewildering to live in a world that does not permit contradictions.

3.3 Bald Men, Heaps, and Vagueness

At what point does a man lose enough hair that he is considered bald? Do we have to be able to see his scalp? What if his hair is long but thin? Does that make a difference? When is someone considered tall? Is there a difference between a "pile" of toys and a "heap" of toys? Is that color red or maroon? All these questions are based on concepts that are somewhat vague. There does not seem to be universal agreement on when someone is bald and when someone is not bald. Nor is there a generally agreed on use of the terms *tall* and *short*. Even your interior decorator might have a hard time distinguishing dark red from maroon. In this section I explore the pervasive element of vagueness in our language and thought.

One of our core ways of describing limitations of reason is by finding contradictions. As I stressed in chapter 1, there are no contradictions in the physical universe. In contrast to the physical universe, in human language and thought there can be contradictions. Humans are not perfect beings. Our language and thought are rife with contradictory statements and beliefs. When we want to reason and talk about the physical world we must ensure that our language and thought do not have contradictions. There are, however, times when we are ostensibly thinking about or discussing the physical world and our meaning is not clear. This happens when there is vagueness. In contrast to contradictions where a statement is both true and false, a vague statement can be thought of as neither true nor false.

Vagueness is applied to terms that are not always perfectly defined. For example, a five-year-old is clearly a child. In contrast, a twenty-five-year-old is definitely not a child. At what point is a person no longer considered a child? There are *borderline cases* where someone is neither a child nor older than a child. Such terms with borderline cases are vague. Other terms with borderline cases are *tall*, *smart*, and *red*. Where does red end and maroon begin? How about scarlet, cardinal, crimson, cherry, puce, pink, ruby, and fuchsia?[16]

One must make a distinction between vague statements and *ambiguous statements*. An ambiguous statement is one in which the subject of the statement is unclear. For example, "Jack is above six feet" is ambiguous since you do not know which Jack is being discussed. Jack Baxter is above six feet, but Jack Miller is below six feet. However, this statement is not vague since six feet is an exact amount. Of course, we can make a statement that is both vague and ambiguous: "Jack is tall."

One must also make a distinction between vague statements and *relative statements*. "Jack Baxter is smart" might be true or not depending on who he is being compared to. If you are comparing Jack to the other people in his class, then he might very well be considered smart; however, the class might not be the smartest class. The truth of a relative statement can be determined by looking at the context of the statement. Who are we talking about? One can imagine the salutatorian at a Harvard University graduation legitimately being called stupid . . . by the valedictorian.

In both ambiguous cases and relative cases there is a lack of specificity. In other words, there is missing information. Usually, if one adds more information, then the statements can be clearly understood. If one identifies the subject of an ambiguous statement or the context of a relative statement, then we can determine if the statement is true or false. In contrast, vague statements usually cannot be tweaked by adding more information. There is no more information to add. When is a person considered bald? The answer is "blowin' in the wind." There is no real answer.

Vagueness is not necessarily a bad thing. Sometimes vagueness is a necessity. Biologists use vague characteristics to describe different species.[17] Many lawyers are employed to work with vagueness (and to obfuscate the truth). Diplomats are vague when they make treaties with foreign countries so that they are not caught by their own words. When a woman asks if a certain dress makes her look fat, it might be wise to be vague in your response.

Philosophers are usually split as to why there is vagueness. Some philosophers promote *ontological vagueness*—that is, the reason some terms do not have an exact meaning is that an exact back to ancient Greek times and is called a *sorites paradox* (from the Greek word *soros* for "heap"). Eubulides of Miletus (fourth century BC) is usually credited with being the first to formulate this puzzle.[19] He asked how many grains of wheat form a heap. Is one grain of wheat considered a heap? Obviously not. How about adding one grain to it? Are two grains considered a heap? Still not. After all, we only added one grain. We can formulate the following law:

If n grains are not a heap, then $n + 1$ grains are also not a heap.

Following a similar analysis of the bald man, we come to the obviously wrong conclusion that no amount of grains form a heap. What went wrong?

Let us carefully analyze the argument given. We start with the obvious statement:

1 grain is not a heap.

We also use the n-grain rule for $n = 1$ to get

If 1 grain is not a heap, then 2 grains are also not a heap.

Combining these two rules using modus ponens, we get

2 grains are not a heap.

Furthermore, combining this with

If 2 grains are not a heap, then 3 grains are also not a heap.

gives us:

3 grains are not a heap.

Continuing on with this shows us that for any n, no matter how large,

n grains are not a heap.

This is obviously false.

We can also go the other way. Consider a heap with 10,000 grains of wheat. If we take off one little grain are we to come to the conclusion that 9,999 grains are not a heap? Obviously they are still a heap. A rule can be formulated:

If n grains are a heap, then $n - 1$ grains are also a heap.

Using this rule and applying the modus ponens rule many times, we arrive at an obviously false conclusion that a collection of 1 grain is also a heap. A similar argument can show that a man with 1 hair, or even no hairs, is not bald.

Another sorites-type paradox is the *small-number paradox* (also called *Wang's paradox*). 0 is a small number. If n is a small number, then so is $n + 1$. We conclude with the apparent false fact that any number is considered a small number. There are many other types of sorites paradoxes. Is a person tall if we add another centimeter to their height? Does a person become heavy if they add one more pound? Similarly, for any other vague terms like *rich*, *poor*, *short*, *clever*, and so on, one has an associated sorites-type paradox.

How is one to understand such paradoxes? Some philosophers say that the sorites paradoxes show us that there is something wrong with the logical rule of modus ponens. By following modus ponens we came to a false conclusion, so modus ponens cannot be trusted. This seems a little too harsh. The modus ponens rule works so perfectly in most logic, math, and reasoning. Why should we abandon it? Other philosophers (who believe that all vagueness is epistemic—i.e., they believe exact boundaries exist that we are not aware of) assume that the rule

If n grains are not a heap, then $n + 1$ grains are also not a heap.

is simply false. For them, there is some n for which n grains do not form a heap but $n + 1$ grains do form a heap. We mortals are not aware of which n this is but it nevertheless exists. For such philosophers modus ponens is true, but this implication is simply not valid and so cannot be used in a modus ponens argument. As noted above, to us it seems that vagueness is not an epistemic but an ontological problem. There are no exact boundaries and the implication from n to $n + 1$ grains is, in fact, always true.

Rather than saying that there is something wrong with the obvious rule of modus ponens, I prefer to say that this amazing rule is perfect but cannot always be applied. In particular, one should not use modus ponens with vague terms. Although modus ponens seems to work with the first few applications of the rule (i.e., that 2, 3, and 4 grains do not make a heap), for many more applications of the rule we come to obvious false conclusions. We must restrict ourselves to using modus ponens only with exact terms. We will not be able to use modus ponens with vague terms because that will take us beyond the bounds of reason.

It makes sense that these logical and mathematical tools do not work with vague terms since these tools were formulated with exact terms in mind. One needs exact terms to do science, logic, and mathematics. When we leave the domain of exact definitions—that is, when we talk about baldness, tallness, and redness—we are necessarily leaving the boundaries where logic and math can help us. Vagueness is beyond the boundaries of reason. While we all freely live and communicate with such terms on a daily basis we must, nevertheless, be careful about crossing the outer limits of reason.

As shown above, when it comes to vague statements, mathematicians and logicians are somewhat at a loss. Their usual tools in their toolbox do not work. However, since these vague terms are ubiquitous, we simply cannot ignore them. Researchers have developed a number of different methods to make sense of the vague world. Here I will highlight several of them.

Logic usually deals with terms that are either true or false. *Fuzzy logic* is a branch of logic that deals with terms that can have any intermediate value between true and false. Say that true is 1 and that false is 0. Rather than dealing with the two-element set {0,1}, fuzzy logic deals with the infinite interval [0,1] of all real numbers between 0 and 1. With this we can give different values in different cases. Telly Savalas and Yul Brynner are both totally bald and hence would have the value 0. People with full heads of hair would get a 1. People in the middle will get middle values. 0.1 means almost bald, while 0.5 is halfway there. Someone might get the value of 0.7235. With these different values set up, researchers have gone on to develop different operations similar to AND and OR to work in this logic.

Similar to fuzzy logic is a related field of logic called *three-valued logic.* Rather than saying that a statement is either true or false, say that a statement is true, false, or indeterminate. These branches of logic are used extensively in the field of artificial intelligence, which tries to make computers act more like human beings. If we are going to have computers interacting with human beings, then they are going to have to deal with vague terms like humans beings. These multivalued logics have been very successful in dealing with vague predicates.

Another method used to deal with vague terms is to restrict logic. Consider a man who is halfway between being bald and being hairy. Rather than saying he is neither bald nor not bald, say that he is both bald and not bald. In classical logic if a statement and its negation are both true, we have a contradiction and the system is shows what would happen if you switch. Using the staying strategy gets you the car one out of three times, while the switching strategy has you winning two out of three times. You should indeed switch.

What's going on here? Why does switching help? The answer is that when Monty Hall opens the other door, he is giving you more information. Monty knows where the car is and is not going to open the door with the car. By avoiding the other door he is giving information that the other door was avoided. When he gives you information, the probabilities of what is behind each door change.

The way to see this more clearly is by imagining that Monty presents twenty-five doors to you and tells you that the car is behind one of the doors and there are goats behind the other twenty-four doors. You choose one of the doors and then Monty proceeds to open twenty-three other doors. Each door he opens reveals a goat, as in figure 3.8.