PREFACE

This book, like hardly any others, explains the pioneering discoveries of the amazing logician Kurt Gödel through recreational logic puzzles. Its title is based on the fact that almost all the puzzles of the book are centered around Gödel's celebrated result.

In the first quarter of the twentieth century there were some mathematical systems in existence that were so comprehensive that it was generally assumed that every mathematical statement could either be proved or disproved within the system. In 1931 Gödel astounded the entire mathematical world by showing that this was not the case [Gödel, 1931]: For each of the mathematical systems in question, there must always be mathematical statements that can be neither proved nor disproved within the system. Indeed, Gödel provided an actual recipe for exhibiting in each such system, a sentence which must be true, but not provable, in the system. This famous result is known as *Gödel's Theorem*.

The essential idea behind Gödel's proof is this:

Gödel assigned to each mathematical sentence of the system a number, now known as the *Gödel number* of the sentence. He then constructed a most ingenious sentence S that asserted that a certain number n was the Gödel number of a sentence that was not provable in the system. Thus this sentence S was true if and only if n was the Gödel number of an unprovable sentence. But the amazing thing is that n was the Gödel number of the very sentence S! Thus S asserted that its own Gödel number was the Gödel number of an unprovable sentence. Thus in effect, S was a self-referential sentence that asserted its own non-provability. This meant that either S was true and not provable or S was false but provable. The latter alternative seemed completely out of the question since it was obvious from the nature of the system in question that only true sentences could be proved in the system. Thus Gödel's sentence S was true but not provable in the system. Its truth was known only by going outside the system and noting some of its properties.

How did Gödel manage to construct such an ingenious sentence? It is the purpose of this book to explain how, in terms that are completely comprehensible to the general public—even those with no background at all in mathematical logic. I have written this book so that it should be perfectly comprehensible to any reasonably bright high-school student. This is actually the first of my popular logic puzzle books in which I give a *complete* proof of Gödel's theorem for one particularly important mathematical system, as well as provide a host of generalizations that have never been published before, and should therefore be of interest, not only to the general reader, but to the logical specialist as well. These generalizations can be found in Chapters XIII, XIV and XV

On the whole, I have written this book in a very informal style. After the chatty introductory Chapter I, which consists mainly of personal anecdotes and jokes, the remaining chapters of Part I consist of puzzles, paradoxes, the nature of infinity (which often seems more paradoxical than it really is) and some curious systems related to Gödel's theorem. Part II is the real heart of this book, and could be read quite independently of Part I. The first three of its chapters contain my generalized Gödel theorems, which are unusual in that they do not involve the usual machinery of symbolic logic! I have deferred symbolic logic—the logical connectives and quantifiers—to the last three chapters, which begin by explaining its basics and what is known as *first-order arithmetic* followed by a presentation of the famous axiom system known as *Peano Arithmetic*. And I give a complete proof there of Gödel's celebrated result that there are sentences of Peano Arithmetic that cannot be proved or disproved within that axiom system.

Gödel's discoveries have led to an even more important result: Is there any purely mechanical method of determining which mathematical statements are true and which are false? This brings us to the subject of *decision theory*, better known as *recursion theory*, which today plays such a vital role in computer science. Chapter XVI of this book explains some basics of this important field. It turns out that in fact there is no purely mechanical method of deciding which mathematical statements are true and which are not! No computer can possibly settle all mathematical questions. It seems that brains and ingenuity are, and always will be, required. In the witty words of the mathematician Paul Rosenbloom, this means that "Man can never eliminate the necessity of using his own intelligence, regardless of how cleverly he tries!"

I would like to thank Dr. Sue Toledo for her very helpful editing work on this book.

TABLE OF CONTENTS

PART I

PUZZLES, PARADOXES, INFINITY AND OTHER CURIOSITIES

CHAPTER I

A CHATTY PERSONAL INTRODUCTION

Let me introduce myself by what might be termed a *meta-introduction*, by which I mean that I will tell you of three amusing introductions I have had in the past.

1. The first was by the logician Professor Melvin Fitting, formerly my student, whom I will say more about later on. I must first tell you of the background of this introduction. In my puzzle book "What is the Name of this Book?" I gave a proof that either Tweedledee or Tweedledum exists, but there is no way to tell which. Elsewhere I constructed a mathematical system in which there are two sentences such that one of them must be true but not provable in the system, but there is no way to know which one it is. [Later in this book, I will show you this system.] All this led Melvin to once introduce me at a math lecture by saying, "I now introduce Professor Smullyan, who will prove to you that either he doesn't exist, or you don't exist, but you won't know which!"

2. On another occasion, the person introducing me said at one point, "Professor Smullyan is unique." I was in a mischievous mood at the time, and I could not help interrupting him and saying, "I'm sorry to interrupt you Sir,

but I happen to be the only one in the entire universe who is not unique!"

3. This last introduction (perhaps my favorite) was by the philosopher and logician Nuel Belnap Jr., and could be applicable to anybody. He said, "I promised myself three things in this introduction: First, to be brief, second, not to be facetious, and third, not to refer to this introduction."

I particularly liked the last introduction because it involved self-reference, which is a major theme of this book.

I told you that I would tell you more about Melvin Fitting. He really has a great sense of humor. Once when he was visiting at my house, someone complained of the cold. Melvin then said, "Oh yes, as it says in the Bible, many are cold but few are frozen." Next morning I was driving Melvin through town, and at one point he asked me, "Why are all these signs advertising slow children?"

On another occasion, we were discussing the philosophy of solipsism (which is the belief "I am the only one who exists!"). Melvin said, "Of course I know that solipsism is the correct philosophy, but that's only one man's opinion." This reminds me of a letter a lady wrote to Bertrand Russel, in which she said, "Why are you surprised that I am a solipsist? Isn't everybody?"

I once attended a long and boring lecture on solipsism. At one point I rose and said, "At this point, I've become an anti-solipsist. I believe that everybody exists except me."

Do you have any rational evidence that you are now awake? Isn't it logically possible that you are now asleep and dreaming all this? Well, I once got into an argument with a philosopher about this. He tried to convince me that I had no rational evidence to justify believing that I was now awake. I insisted that I was perfectly justified in being certain that I was awake. We argued long and tenaciously, and I finally won the argument, and he conceded that I did have rational evidence that I was awake. At that point I woke up.

Coming back to Melvin Fitting, his daughter Miriam is really a chip off the old block. When she was only six years old, she and her father were having dinner at my house. At one point Melvin did not like the way Miriam was eating, and said, "That's no way to eat, Miriam!" She replied, "I'm not eating Miriam!" [Pretty clever for a six-year old, don't you think?]

One summer, Melvin, who was writing his doctoral thesis with me, was out of town. We corresponded a good deal, and I ended one of my letters saying, "And if you have any questions, don't hesitate to call me collect and reverse the charges." [Get it?]

I would like to tell you now of an amusing lecture I recently gave at a logic conference in which I was the keynote speaker. The title of my talk was "Coercive Logic and Other Matters." I began by saying, "Before I begin speaking, there is something I would like to say." This got a general laugh. I then explained that what I just said was not original, but was part of a manuscript of the late computer scientist Saul Gorn about sentences which somehow defeat themselves. He titled this collection "Saul Gorn's compendium of rarely used clichés." It contains such choice items as:

1. Half the lies they tell about me are true.
2. These days, every Tom Dick and Harry is named "John."
3. I am a firm believer in optimism, because without optimism, what is there?
4. I'm not leaving this party till I get home!

5. If Beethoven was alive today, he would turn over in his grave!
6. I'll see to it that your project deserves to be funded.
7. This book fills a long needed gap.
8. A monist is one who believes that anything less than everything is nothing.
9. A formalist is one who cannot understand a theory unless it is meaningless.
10. The reason that I don't believe in astrology is because I'm a Gemini.

The last one was mine. I used that line frequently in the days that I was a magician. In those days, people often asked me whether I had ever sawed a lady in half. I always replied that I have sawed dozens of ladies in half, and I'm learning the second half of the trick now.

Next, I told the logic group that I had prepared two different lectures for the evening, and I would like them to choose which of the two they would prefer. I then explained that one of the lectures was very impressive and the other was understandable. [This got a good laugh].

Next, I said that I would give a test to see if members of the audience could do simple propositional logic. I displayed two envelopes and explained that one of them contained a dime and the other one didn't. On the faces of the envelopes were written the following sentences:

1. The sentences on the two envelopes are both false.
2. The dime is in the other envelope.

I explained that each sentence is, of course, either true or false, and that if anyone could deduce from these sentences where the dime was, he could have the dime. But for the privilege of taking this test, I would charge a nickel. Would anyone volunteer to give me a nickel for the privilege of doing this? I got a volunteer. I then told him, "You are not allowed to just *guess* where the dime is; you must give a valid *proof* before the envelope is opened." He agreed. I said, "Very well. Where is the dime, and what is your proof?" He replied, "If the first sentence, the sentence on Envelope 1, were true, then what it says would be the case, which would mean that both sentences are false, hence the first sentence would be false, which is a clear contradiction. Therefore the first sentence can't be true; it must be false. Thus it is false that both sentences are false, hence at least one must be true, and since it is not the first, it must be the second, and so the dime must be in the other envelope, as the sentence says."

"That sounds like good reasoning," I said. "Open Envelope 1." He did so, and sure enough there was the dime.

After congratulating him, I said that the next test would be a little bit more difficult. Again I showed two envelopes with messages written on them, and I explained that one of them contained a dollar bill and the other was empty. The purpose now was to determine from the messages which envelope contained the bill. Here are the messages:

1. Of the two sentences, at least one is false.
2. The bill is in this envelope.

I then explained that if the one taking this test could correctly prove where the bill was, he or she could keep it, but for the privilege of taking this test, I would charge 25¢. After some thoughts, one man volunteered. I then asked him where the bill was, and to prove that he was right. He said, "If Sentence 1 were false, it would be true that at least one was false, and you would have a contradiction. Therefore Sentence 1 must be true, hence at least one of the sentences is false, as Sentence 1 correctly says. Therefore Sentence 2 is false, and so the bill is really in Envelope 1." I said, "Very well, open Envelope 1." He did so, and it was empty! He then opened Envelope 2, and there was the bill!

At this point, he, and other members of the audience looked puzzled. I then asked, "How come the bill was in Envelope 2 instead of Envelope 1?" One member of the audience yelled, "Because you obviously were lying!" I assured the audience that at no time did I lie, and indeed I never did! So given the fact that I did not lie, what is the explanation?

Problem 1. What is the explanation of why the bill was in Envelope 2, despite the volunteer's purported proof that the bill was in Envelope 1? What was wrong with the proof he gave? [Answers to problems are given at the end of chapters. Realize, though, that sometimes there are more ways than one to arrive at the solution to a given problem.]

At this point, the volunteer owed me 25¢. I then told the audience that I felt a little bit guilty about having won a quarter by such a trick. And so I said to the volunteer, "I want to give you a chance to win your money back, so I'll play you for double or nothing." [This got a general laugh.] "In fact," I continued, "I'll be even more generous!" I then handed him two $10 bills and told him that he could have his quarter back and even keep some of the money I just gave him, but he would have to agree to something first. I told him I was about to make a statement. If he wanted the deal I was proposing, he had to promise to give me back one of the bills if the statement was false. But if the statement turned out to be true, then he must keep both bills. "That's a pretty good deal, isn't it?" I asked. "You are bound to get at least $10, and possibly $20!" He agreed. I then made a statement such that in order for him to keep to the agreement, the only way was to pay me $1000!

Problem 2. What statement would accomplish this?

At this point the poor fellow owed me a thousand dollars. Later I will tell you how I gave him a chance (sic!) to regain his thousand dollars, but first I wish to tell you of a related incident (which I also told the audience): Many years ago, when I was a graduate student at Princeton, I would frequently visit New York City. On one of my visits I met a very charming lady musician. On my first date with her, I asked her to do me a favor. I told her that I would make a statement in a moment, and I asked her whether she would give me her autograph if the statement turned out to be true. She replied, "I don't see why not." And I said that if the statement was false, she should not give me her autograph. She agreed. I then made a statement such that in order for her to keep her word, she had to give me, not her autograph, but a kiss!

Problem 3. What statement would work?

Now, the statement I gave in the solution to the last problem had to be false, and she had to give me a kiss. However, there is another statement I could have made which would have had to be true, after which she would also have had to give me a kiss.

Problem 4. What statement could that be?

There is still another statement I could have made (a more interesting one, I believe) which could be either true or false, but in either case, she would have to give me a kiss. [There is no way of knowing whether the statement is true or false before the lady acts.]

Problem 5. What statement would accomplish this?

Anyway, whatever statement I would have made, it was a pretty sneaky way of winning a kiss, wasn't it? Well, what happened next was even more interesting. Instead of collecting the kiss, I suggested we play for double or nothing. She, being a good sport, agreed. And so she soon owed me two kisses, then with another logic trick four, then eight, then sixteen, then thirty-two, and things kept doubling and escalating and doubling and escalating and before I knew it, we were married! And I was married to Blanche, the charming lady musician, for over 48 years.

Once at breakfast I had the following conversation with Blanche:

Ray	Is *NO* the correct answer to this question?
Blanche	To what question?
Ray	To the question I just asked. Is *NO* the correct answer to that question?
Blanche	No, of course not!
Ray	Aha, you answered *NO*, didn't you!
Blanche	Yes.
Ray	And did you answer correctly?
Blanche	Why, yes!
Ray	Then *NO is* the correct answer to the question.
Blanche	That's right.
Ray	Then when I asked you what the correct answer is, you should have answered *YES*, not *NO!*
Blanche	Oh yes, that's right! I should have answered *YES*.
Ray	No, you shouldn't! If you answered *YES*, you would be affirming that *NO* is the correct answer so why would you give the incorrect answer *YES*?
Blanche	You're confusing me!

Fortunately, Blanche did not divorce me for this!

It's sometimes annoying for a wife to have an overly rational husband, isn't it? The following dialogue from my book "This Book Needs No Title" well illustrates this:

Wife	Do you love me?
Husband	Well of course! What a ridiculous question!
Wife	You *don't* love me!
Husband	Now what kind of nonsense is this?
Wife	Because if you *really* loved me, you couldn't have done what you did!
Husband	I have already explained it to you that the reason I did what I did was *not* that I don't love you, but because of such and such.
Wife	But this such and such is only a *rationalization*! You really did it because of so and so, and this so and so would never be if you really loved me.

Etc., etc.!

Next Day

Wife	Darling, do you love me?
Husband	I'm not so sure!
Wife	What!
Husband	I thought I did, but the argument you gave me yesterday proving that I don't is not too bad!

I already told you how on my first date with Blanche, I won a kiss using logic. Here is another way of winning a kiss: I say to a lady, "I'll bet you that I can kiss you without touching you." After giving a precise definition of kissing and of touching, she realizes that it is logically impossible, and takes the bet. I then tell her to close her eyes. She does so, I then give her a kiss and say, "I lose!"

This is reminiscent of the prank in which you go into a bar with a friend who orders a martini. You place a tumbler on the martini and say, "I'll bet you a quarter that I can drink the martini without removing the tumbler." He accepts the bet. You then remove the tumbler, drink the martini and give him a quarter!

This is reminiscent of the story of a programmer and an engineer sitting next to each other on an airplane. The following conversation ensued:

Programmer	Would you like to play a game?
Engineer	No, I want to sleep.
Programmer	It's a very interesting game!
Engineer	No, I want to sleep.
Programmer	I ask you a question. If you don't know the answer, you pay me five dollars. Then you ask me a question and if I don't know the answer, then I pay you five dollars.
Engineer	No, no, I want to sleep.
Programmer	I'll tell you what! If you don't know the answer to my question, you pay me five dollars, but if I don't know the answer to your question, I'll pay you fifty dollars!
Engineer	O.K. Here's a question. What goes up the hill with four legs and comes down with five legs?

The programmer then took out his portable computer and worked on the question for an hour, but got nowhere. And so he handed the engineer fifty dollars. The engineer said nothing, but put the fifty dollars in his pocket. The programmer, a bit miffed, said, "Well, what's the answer?" The engineer then handed him five dollars.

Coming back now to my lecture and the guy who owed me a thousand dollars, I said to him, "I really feel sorry for you, and so I will give you back your thousand dollars on condition that you answer a yes/no question truthfully for me." He agreed. I then asked him a question such that the only way he could keep his word was by paying me, not a thousand dollars, but a million dollars!

Problem 6. What question would work?

At this point, I said to him, "I am now in a very generous mood, and so I'll tell you what I'm going to do! I'll give you back your million dollars on condition that you give me the answer to another yes/no question, but this time you don't have to answer truthfully! Your answer can be either true or false; you have the option! There is obviously no way I can trick you now, right?" He agreed that it was obviously impossible for me to con him under the given conditions, and so he accepted. Ah, but there was a way I could con him! The next question I asked was such that he had to pay me, not a million dollars, but a billion dollars!

Problem 7. How in the world was that possible?

Next, I told him that I was very sorry that he owed me a billion dollars, and so I would give him a 50% chance to win it back again, but for this privilege I would charge a nickel extra. "Isn't it worth a nickel," I asked, "for a fifty percent chance of winning back a billion dollars?" He agreed. I then wrote something on a piece of paper, folded it and handed it to someone so I could not use any slight of hand. I then explained that I had written a description of an event which would or would not take place in the room sometime in the next fifteen minutes. His job was to predict whether or not it would take place. "Your chances of predicting correctly is fifty percent, isn't it?" He agreed that it was. I then handed him a pen and a blank piece of paper and told him that if he believed that the event would take place, he should write "yes," otherwise write "no." He then wrote something on the paper. I asked, "Have you written down your prediction?" He said he had. I said, "Then you have lost!"

Problem 8. What could I have written such that regardless of whether he wrote "yes" or "no", he was bound to lose?

At this point, he still owed me a billion dollars. I then said to him, "I'll tell you what. I'll trade you the whole billion dollars for one kiss from your lovely wife!" [This got a real good laugh!]

I am incorrigible, you say? I certainly am! Indeed my epitaph will be:

IN LIFE HE WAS INCORRIGIBLE. IN DEATH HE'S EVEN WORSE!

Solutions to the Problems of Chapter I

1. In the first test I gave, I said that each sentence was either true or false. I never said that in the second test! If I don't say anything about the truth or falsity of the sentences involved, I can write anything I like and put the bill wherever I want! The fact is that in the last test, the sentence on Envelope 1 couldn't be either true or false. If it were false, you would have a logical contradiction. If it were true, you wouldn't have a *logical* contradiction; it would imply an empirically false fact—namely where the bill is. Thus the first sentence cannot be either true or false. [The second sentence, incidentally, is in fact true]. I use this puzzle as a dramatic illustration of Alfred Tarski's discovery that the very notion of truth is not well defined in various languages such as English. Later in this book I will give you a far more formal account of Tarski's theorem.

2. The statement I made was, "You will give me either one of the bills or a thousand dollars." If the statement was false, he would have to give me one of the bills, but doing so would make it *true* that he gives me either one of the bills or a thousand dollars, and we would have a contradiction. Hence my statement can't be false; it must be true. Therefore it is true that he must give me either one of the bills or a thousand dollars, but he can't give me one of his bills, because our agreement was that if my statement was true, he is to keep both bills! Therefore he has to give me a thousand dollars.

3. The statement I made was, "You will give me neither your autograph nor a kiss." If the statement were true, she would have to give me her autograph as agreed, but doing so she would make it false that she gives me neither her autograph nor a kiss, and we would have a contradiction.

Therefore the statement couldn't be true; it must be false. Since it was false that she was going to give me *neither*, then she had to give me *either*—either her autograph or a kiss. But she couldn't give me her autograph for a false statement, for that was the rule to which she had agreed! Hence she owed me a kiss!!

4. Here is a statement which has to be true and is such that she must give me a kiss. "Either you will not give me your autograph or you will give me a kiss."

I am asserting that one of the following two alternatives holds:

(1) You will not give me your autograph.

(2) You will give me a kiss.

If my assertion were false, making both alternatives false, then neither (1) nor (2) would hold, hence (1) would not hold, which means that she *would* give me her autograph, contrary to our agreement that she does not give me her autograph for a false statement. Therefore my statement cannot be false. Thus it is true that either (1) or (2) holds, but then she must give me her autograph, since the statement is true, which means that (1) cannot hold, and therefore (2) must hold. Thus she must give me a kiss (as well as her autograph).

5. A statement that works is, "You will give me either both an autograph and a kiss, or neither one." Thus I am asserting that one of the following alternatives holds:

(1) You will give me both.

(2) You will give me neither.

Suppose the statement is true. This means that one of the two alternatives really does hold, but it can't be (2), since she must give me her autograph for a true statement, hence it must be (1), and so she must give me both her kiss and her autograph.

Now suppose the statement is false. The only way it can be false is that she gives me one but not the other—either a kiss and no autograph, or an autograph and no kiss. The latter possibility is ruled out, because she cannot give me an autograph for a false statement. Therefore she must give me a kiss.

In summary, if the statement is true, she must give me both her autograph and a kiss, and if the statement is false, she must give me a kiss, but not her autograph. The interesting thing is that there is no way of knowing whether the statement is true or false, until the lady acts. It is actually up to her whether the statement is true or false! In either case, she must give me a kiss, but she has the option of giving me her autograph or not. If she does, that would make the statement true, and if she doesn't, that would make the statement false.

6. The question I asked was, "Will you either answer NO to this question or pay me a million dollars?" [Equivalently, I could have asked him the question, "Will you pay me a million dollars if you answer yes to this question?"]

I am asking whether one of the following two alternatives holds:

(1) You will answer NO

(2) You will pay me a million dollars.

If he answers NO, then he is claiming that neither alternative (1) or alternative (2) holds, whereas (1) did hold, so NO cannot be a correct answer. Hence to be truthful, he must answer YES. He therefore affirms that either (1) or (2) holds, but now (1) doesn't hold, and so it must be (2). Therefore he owed me a million dollars!

7. I said that he could answer me either truthfully or falsely. I never said that he could answer me paradoxically! Well, one can design a question such that unless he pays me a billion dollars, neither a YES nor a NO answer

could be either true or false, but paradoxical! Such a question is, “Is YES the correct answer to this question if and only if you pay me a billion dollars?” [In other words, is it the case that either yes is the correct answer to this question and you pay me a billion dollars, or *no* is the correct answer and you don’t pay me a billion dollars]. If he doesn’t pay me a billion dollars, then the question reduces to “Is *no* the correct answer to this question” and the answers *yes* and *no* are both neither true or false, but paradoxical, hence the only way he can avoid answering me paradoxically is by paying me a billion dollars.

8. What I wrote was “You will answer NO.” If he wrote YES, then he is affirming that the event will take place, which it didn’t, and if he writes NO, he is denying the event will take place, which it did. In either case he loses!

CHAPTER II

SOME CURIOUS ADVENTURES

Those of you who are familiar with some of my earlier puzzle books know about the place called the *Island of Knights and Knaves,* where knights always tell the truth and knaves always lie and every inhabitant is either a knight or a knave. Well, many years ago, long before I was married, I visited this strange place and had the following curious adventures, all of them leading to fascinating problems I had to solve. I will start with some simple ones.

Problem 1. On one of my visits I was introduced to three inhabitants A, B and C, and was told that at least one was a knight and at least one was a knave and that one of them had a prize that I could have, if I could determine which one had it. The three made the following statements:

A B doesn't have the prize.
B I don't have the prize.
C I have the prize.

Which one has the prize?

Problem 2. On my next visit to this island I met two natives named *Hal* and *Jal*. Hal uttered a statement of only three words, from which I could deduce that he and Jal were the same type (both knights or both knaves).

What statement could that have been?

Problem 3. On my next visit to this island I came across three natives A, B and C and was reliably informed that one of the three was a magician. They made the following statements:

A	B is not both a knave and a magician.
B	Either A is a knave or I am not a magician.
C	The magician is a knave.

Which one is the magician and what type is each?

Problem 4. A Court Case. I then witnessed a trial. A crime had been committed and three suspects, A, B, and C, were being tried. They made the following statements:

A	I am guilty.
B	I am the same type as at least one of the others.
C	We are all of the same type.

Which one is guilty?

Problem 5. On this particular island, each woman is either a constant liar or a constant truth teller. The men are as usual—knights and knaves. I was introduced to three married couples—the Arks, the Bogs, and the Cogs. One of the three couples is the king and queen of the island. I was reliably informed that in none of the couples are both of them liars. They all made the following statements:

Mr. Ark	I am not the king.
Mrs. Ark	The king was born in Italy.

Mr. Bog	Mr. Ark is not the king.
Mrs. Bog	The king was really born in Spain.
Mr. Cog	I am not the king.
Mrs. Cog	Mr. Bog is the king.

Which one is the king?

Problem 6. Another Court Case In this case, three couples—the Dags, the Eggs and the Fens were interrogated because it was known that one of the three men was a spy, but it was not known which one. A curious fact of this case is that in one of the couples, husband and wife were both truthful, in another, both were liars, and in another, one of the spouses was truthful and the other lied. They all made the following statements:

Mr. Dag	I am not the spy.
Mrs. Dag	Mr. Egg is the spy.
Mr. Egg	Mr. Dag is truthful.
Mrs. Egg	Mr. Fen is the spy.
Mr. Fen	I am not the spy.
Mrs. Fen	Mr. Dag is the spy.

Which one is the spy?

Problem 7. One day I saw an extremely beautiful lady on this island and was immediately smitten with her. I longed to know whether or not she was married, but I did not have the courage to ask her. The next day I came across her two brothers Alfred and Bradford. Alfred then made a statement. From this statement I could not tell whether or not the lady was married. Then to my surprise, Bradford made the same statement, from which I could tell that she was not married.

What statement could that have been?

Problem 8. My next adventure on this island was quite harrowing! I got captured by a ferocious gang of brigands and was shown three natives A, B and C, and was told that one of them was the witch doctor. I was to point to one of them, and if I pointed to the witch doctor, I would get executed, but if I pointed to one who was not the witch doctor I could go free. The three made the following statements:

A I am the witch doctor.
B I am not the witch doctor.
C At most one of us is a knight.

To which of the three should I point?

Problem 9. Actually the last adventure ended quite happily. I pointed to one of them and correctly explained why he couldn't be the witch doctor. The gang was quite pleased with my reasoning and became friendly. One of them said, "He seems like a nice guy; let us give him a reward!" They then showed me a picture of a very beautiful girl, and by Heavens, she was the very one with whom I was smitten! "She has seen you," one of them said, "and is quite fond of you. Tomorrow we will give you another test, and if you pass it, the lady will be yours."

True to their word, the next day I was shown five adjacent rooms and was told that the lady was in one of them. On the door of Rooms 1, 2, 3, 4, 5 were signs 1, 2, 3, 4, 5 respectively. From their signs I was to infer which room contained the lady, and which signs were true. If I succeeded, then the lady would be mine.

Here are the signs:

Sign 1.	The lady is not in Room 2.
Sign 2.	The lady is not in this room.
Sign 3.	The lady is not in Room 1.
Sign 4.	At least one of these five signs is false.
Sign 5.	Either this sign is false, or the sign on the room with the lady is true.

Which room contains the lady, and which of the signs are true?

Solutions to the Problems of Chapter II

1. If C has the prize then all three statements are true, which would mean that all three of the speakers are knights, contrary to what is given. If B has the prize then all three must be knaves, again contrary to what is given. Therefore it must be that A has the prize (and also A and B are knights and C is a knave).

2. What Hal said was, "Jal is truthful." If Hal is a knight, then Jal is truthful, as Hal said, hence also a knight. If Hal is a knave, then contrary to what he said, Jal is not truthful, hence also a knave.

3. From C's statement it follows that C cannot be the magician, because if he is a knight then the magician is really a knave and hence cannot be C. On the other hand if C is a knave, then contrary to his statement, the magician is not a knave but a knight, hence cannot be C who is a knave. Thus in either case, C is not the magician.

Next we will see that A must be a knight. Well, suppose he were a knave. Then his statement is false, which means that B must be both a knave and a magician. Since A is a knave (under our assumption) then it is true that *either* A is a knave *or* (anything else!). Thus it is true that *either* A is a knave *or* B is not the magician, but this is just what B said, and thus the knave B made a true statement which is not possible! Thus the assumption that A is a knave leads to an impossibility, hence A cannot be a knave. Thus A is a knight. Hence his statement is true, which means that B is not both a knave and a magician.

We now know the following:

(1) C is not the magician.

(2) A is a knight.

(3) B is not both a knave and a magician.

Next we will see that B cannot be the magician, for suppose he were. Then it is false that he is not the magician, and it is false that A is a knave (by (2)), hence both alternatives of B's statement are false. Hence B's statement must be false, which makes B a knave. Hence B is then both a knave and a magician, which is contrary to (3)! Thus it cannot be that B is the magician. Also C is not the magician, as we have seen. Thus it must be A who is the magician.

Also, since B is not the magician, what he said is true, hence B is a knight. As for C, what he said cannot be true, since the magician is really a knight (A), not a knave. Hence C is a knave.

In summary, A and B are both knights, C is a knave and the magician is A.

4. Clearly, if we can show that A is a knight, we will know that A is the guilty party, since that is what he claims. Now, if B is a knave, then A must be a knight, since B's telling a lie implies that he is the only knave. On the other hand, if B is a knight, then he really is of the same type as at least one of the others, as he said. Thus either A is a knight or C is a knight. If A is a knight, we are done. But if C is a knight, then all three really are of the same type, making A a knight again. Thus A is a knight, period! So A is unquestionably the guilty one.

5. If Mr. Cog is the king then both Mr. and Mrs. Cog are making false statements, contrary to what is given. Therefore Mr. Cog is not the king.

Now, Mr. Ark and Mr. Bog are in agreement, hence they are both knights or both knaves. If they were both knaves, their wives Mrs. Ark and Mrs. Bog would both be truthful which is impossible, since they can't both be right. Therefore Mr. Ark and Mr. Bog are both knights, hence their statements are both true, which means that

Mr. Ark is not the king. Therefore Mr. Bog is the king and Mrs. Bog is the queen.

6. If Mr. Dag is the spy, then the Dags are both knaves (since their statements are both false) and furthermore the Eggs must also both be knaves, which violates the given conditions. Therefore Mr. Dag is not the spy.

If Mr. Egg is the spy, then the Eggs and the Fens are both mixed couples, because Mr. Egg is then truthful, Mrs. Egg lied, Mr. Fen is truthful and Mrs. Egg lied. Again this cannot be, and so Mr. Egg is not the spy.

Thus the spy is Mr. Fen.

7. What Alfred said was, "Either at least one of us is a knave or she is not married."

If Alfred is a knave, then it would be true that at least one of them is a knave (namely Alfred), hence it would be true that either one of them is a knave or the lady is unmarried, but knaves don't make true statements, hence Alfred must be a knight. Hence it is true that *either* at least one is a knave *or* the lady is unmarried. If Bradford is a knight, then the lady must be unmarried (since it is then false that at least one is a knave), but if Bradford is a knave, then there it cannot be determined whether the lady is married or not. Also, there is no way of knowing whether Bradford is a knave or a knight. Thus until Bradford spoke, there was no way of knowing whether or not the lady was married. But after Bradford said the same thing, he agreed with Alfred who is a knight, hence Bradford must also be a knight, which then settles the case—the lady must be unmarried.

To show that several different approaches can be taken to solve most of these problems, here is another solution to Problem 7, which comes at the problem from a different direction:

What both Alfred and Bradford said was, "Either at least one of us is a knave or she is not married." Now let C be *either* Alfred or Bradford. If C is a knave, then at least one of the two brothers is a knave (namely C). Hence it would be true that either one of them is a knave or the lady is unmarried, which is what C said. But knaves don't make true statements, so C must be a knight. But since C was taken to be either Alfred or Bradford, both the brothers must be knights. Thus it must have been the second part of what each said that was true, and the lady had to be unmarried. Note that when only Alfred had spoken, all that could be deduced was that Alfred was a knight and either Bradford was a knave or the lady was unmarried.

8. I reasoned as follows: C is either a knight or a knave. Suppose he is a knight. Then what he said is true, hence A and B must both be knaves, and since B is then a knave, his statement is false, which means that he is the witch doctor. Thus if C is a knight then B is the witch doctor.

Now suppose C is a knave. Then contrary to what he said, there must be more than one knight, hence A and B are both knights. Since A is then a knight, his statement is true, which means that he is the witch doctor.

In neither case is C the witch doctor, and so I pointed to C.

9. Let us first look at Sign 4. If it were false, then it would be true that at least one of the signs is false (namely Sign 4), which would make Sign 4 true, and we would have a contradiction. Hence Sign 4 cannot be false; it must be true. Since it is true, then like it correctly says, at least one of the signs really is false.

Next let us consider Sign 5. If it were false, then both its claims would have to be false, and the first claim is that sign 5 is false, which would make Sign 5 true, and we would again have a contradiction. Since Sign 5 cannot be false, so it must be true. Since it is true, then, as it correctly says, either it is false or the sign on the room with the lady is true, but the first alternative is out, since the sign is not false, and so it must be the case that the sign on the room with the lady is true. We now know four things:

(1) Sign 4 is true.
(2) Sign 5 is true.
(3) The sign on the room with the lady is true.
(4) At least one of the five signs is false.

From (3) it follows that Sign 2 must be true, because if it were false, then, contrary to what the false sign says, the lady would be in Room 2, hence the sign on Room 2, which is the room with the lady, would be false, which by (3) is not the case. Therefore Sign 2 is true, and as it says, the lady is not in room 2, which makes Sign 1 also true. Thus Signs 1, 2, 4, and 5 are all true, and since at least one of the signs is false, it must be Sign 3. Hence, contrary to what Sign 3 says, the lady is really in Room 1. This solves everything.

Epilogue. And so I won the lady, but despite her beauty, I soon found that she never told the truth, and so we soon broke up, which turned out to be a good thing, since years later I met the very lovely pianist Blanche, with whom I was married for 48 years.