Introduction

Let us play a game.

We have selected a set of numbers, and you must guess the set that we have in mind. The set consists of every positive integer with a sole exception. Thus the set might be $\{2, 3, 4, 5, \ldots\}$ or $\{1, 3, 4, 5, \ldots\}$ or $\{1, 2, 4, 5, \ldots\}$, etc. We will give you an unlimited number of clues about the set, and you are to guess after each clue. We will never tell you whether you are right.

First clue: The set contains the number 1.

Please guess the set we have in mind. Would you like to guess the set $\{2, 3, 4, 5, \ldots\}$? (That would be unwise.)

Second clue: The set contains the number 3.

Please make another guess. How about $\{1, 2, 3, 4, 5, 6, 8, 9, 10, \ldots\}$ or does that seem arbitrary to you?

Third clue: The set contains the number 4.

Go ahead and guess.

Fourth clue: The set contains the number 2.

Does the fourth clue surprise you? Guess again.

Fifth clue: The set contains the number 6.

Guess.

Sixth clue: The set contains the number 7.

Guess.

Seventh clue: The set contains the number 8.

Guess.

We interrupt the game at this point because we would like to ask you some questions about it.

First question: Are you confident about your seventh guess? Give an example of an eighth clue that would lead you to repeat your last guess. Give an example of an eighth clue that would lead you to change your guess.

Second question: Let us say that a "guessing rule" is a list of instructions for converting the clues received up to a given point into a guess about the set we have in mind. Were your guesses chosen according to some guessing rule, and if so, which one?

Third question: What should count as winning the game? Consider the following criterion: You win just in case at least one of your guesses is right. This criterion makes winning the game too easy. Say why.

Fourth question: We advocate the following criterion: You win just in case you eventually make the right guess and subsequently never change your mind regardless of the new clues you receive. In this case let us say that you "win in the limit." Is it possible to win the game in the limit even though you make one hundred wrong guesses? Is there any number of wrong guesses that is logically incompatible with winning the game in the limit?

Fifth question: Suppose that all the clues we give you are of the form: The set contains the number n. Suppose furthermore that for every positive integer i, we eventually give you a clue of this form if and only if i is in fact contained in the set we have in mind. (So for every number i in our set, you are eventually told that the set contains i; also you receive no false information about the set.) Do not suppose anything about the order in which you will get all these clues. We will order them any way we please. (Recall how we surprised you with the fourth clue.) Now let us call a guessing rule "winning" just in case the following is true. If you use the rule to choose your guesses, then no matter which of the sets we have in mind, you are guaranteed to win the game in the limit. Specify a winning guessing rule for our game.

Sixth question: We make the game harder. This time we are allowed to select any of the sets that are legal in the original game, but we may also select the set $\{1, 2, 3, 4, 5, 6, \dots\}$ of all positive integers. The rules about clues are the same as given in question 5. Play this new game with a friend, and then think about the following question. Is there a winning guessing rule for the new game?

Seventh question: Let us make the last game easier. The choice of sets is the same as in the last game, but we now agree to order our clues in a certain way. For all positive integers i and j, if both i and j are included in the set we have in mind, and if i is less than j, then you will receive the clue "The set contains i" before you receive the clue "The set contains j." Can you specify a winning guessing rule for this version of the game?

Eighth question: Here is another variant. We select a set from the original collection (thus the set $\{1, 2, 3, 4, 5, \dots\}$ of all positive integers is no longer allowed). Clues can be given in any order we please. You get only one guess. You may wait to see as many clues as you like, but your first guess is definitive. Play this game with a friend. Then show that no matter what rule you use to make your guess, you are not guaranteed to be right. Think about what happens if you are allowed two guesses in the game.

The games we have been playing resemble the process of scientific discovery. Nature plays our role, selecting a certain pattern that is imposed on the world. The scientist plays your role, examining an endless series

of clues about this pattern. In response to the clues, the scientist emits guesses. Nature never says whether the guesses are correct. Scientific success consists of eventually offering a correct guess and never deviating from it thereafter. Language acquisition by children can also be construed in terms of our game. The child's parents have a certain language in mind (the one they speak). They provide clues in the form of sentences. The child converts these clues into guesses about the parents' language. Acquisition is successful just in case the child eventually sticks with a correct guess.

The similarity of our game to these and other settings makes it worthy of more careful study. We would like to know which versions of the game are winnable and by what kinds of guessing rules. Research on these questions began in the 1960s by Putnam (1975), Solomonoff (1964), and Gold (1967). These initial investigations have given rise to a large literature in computer science, linguistics, philosophy, and psychology. This body of theoretical and applied results is generally known as *learning theory* because many kinds of learning (e.g., language acquisition) can be construed as successful performance in one of our games.

In this book we attempt to develop learning theory in systematic fashion, presupposing only basic notions from set theory and the theory of computation. Throughout our exposition, definitions and theorems are illustrated by consideration of language acquisition. However, no serious application of the theory is described.

The book is divided into three parts. Part I advances a fundamental model of learning due essentially to Gold (1967). Basic notation, terminology, and theorems are there presented, to be relied on in all subsequent discussion. In part II these initial definitions are generalized and varied in dozens of ways, giving rise to a multitude of learning models and theorems. We attempt to impose some order on these results through a system of notation and classification. Part III explores diverse issues in learning theory that do not fit neatly into the classification offered in part II.