Beating the Computer

THE FOLLOWING STORY IS known as Newcomb's Problem. It involves playing a game against an opponent who can predict what you do.

Suppose there are two closed boxes presented to you. You are faced with the following choice. Box A contains either a check for one million dollars or nothing (you don't know which). Box B contains a check for one thousand dollars. You can take the contents of Box A alone, or the contents of both boxes. A very smart computer has been fed information about you; if it predicted you'll take both boxes, it has already put nothing in box A; but if it predicted you'll take only A, it has already put one million dollars in there. The computer has almost always been right in predicting other people in the past.

Here is a table summarizing the situation:

	Computer predicts you'll take only A & puts $1 million in A	Computer predicts you'll take both and puts $0 in A
You take just A	You get $1 milion	You get $0
You take both boxes	You get $1,001,000	You get $1,000

What should you do? You will probably come up with one of these two conflicting answers:

(1) The computer has *already* put either one million dollars or nothing in A. What's in there won't change depending on your choice. If you take just A, you'll get whatever's in there; if you take both, you'll get that plus the thousand in B. Take both boxes.

(2) The computer has almost certainly predicted you correctly; so if you pick both boxes, it probably has put nothing in there, and you'll get only one thousand dollars. If you pick only A, again the computer has almost certainly predicted this, so you'll probably get one million dollars. Pick only A.

Which is the right answer? Do you take one box or two?

The odd thing about this problem is that people seem to divide into dogmatic one-boxers and militant two-boxers. There's really no consensus about who is right. Each faction has a perfectly good argument, which it keeps repeating louder and louder to the other side, to no avail.

A QUESTION TO THINK ABOUT: Which strategy do you think is right? Try to construct an argument to convince the opposition of your view.

FOR FURTHER READING: The first appearance of this problem in print seems to have been in Robert Nozick, "Newcomb's Problem and Two Principles of Choice," in *Essays in Honor of Carl G. Hempel*, ed. Nicholas Rescher (Dordrecht: Reidel, 1969). (Nozick says that it was invented, but not published, by Dr. William Newcomb of the Livermore Radiation Laboratories in California.) Newcomb's Problem is discussed in many places, including Richard C. Jeffrey's *The Logic of Decision*, 2nd ed. (Chicago: University of Chicago Press, 1983).