CHAPTER EIGHT

A Matchbox Game-Learning Machine

I knew little of chess, but as only a few pieces were on the board, it was obvious that the game was near its close. . . . [Moxon's] face was ghastly white, and his eyes glittered like diamonds. Of his antagonist I had only a back view, but that was sufficient; I should not have cared to see his face.

THE QUOTATION is from Ambrose Bierce's classic robot story, "Moxon's Master" (reprinted in Groff Conklin's excellent science-fiction anthology, *Thinking Machines*). The inventor Moxon has constructed a chess-playing robot. Moxon wins a game. The robot strangles him.

Bierce's story reflects a growing fear. Will computers someday get out of hand and develop a will of their own? Let it not be thought that this question is asked today only by those who do not understand computers. Before his death Norbert Wiener anticipated with increasing apprehension the day when complex government decisions would be turned over to sophisticated game-theory machines. Before we know it, Wiener warned, the machines may shove us over the brink into a suicidal war.

The greatest threat of unpredictable behavior comes from the learning machines: computers that improve with experience. Such machines do not do what they have been told to do but what they have *learned* to do. They quickly reach a point at which the programmer no longer knows what kinds of circuits his machine contains. Inside most of these computers are randomizing devices. If the device is based on the random decay of atoms in a sample radioactive material, the machine's behavior is not (most physicists believe) predictable even in principle.

Much of the current research on learning machines has to do with computers that steadily improve their ability to play games. Some of the work is secret—war is a game. The first significant machine of this type was an IBM 704 computer programed by Arthur L. Samuel of the IBM research department at Poughkeepsie, New York. In 1959 Samuel set up the computer so that it not only played a fair game of checkers but also was capable of looking over its past games and modifying its strategy in the light of this experience. At first Samuel found it easy to beat his machine. Instead of strangling him, the machine improved rapidly, soon reaching the point at which it could clobber its inventor in every game. So far as I know no similar program has yet been designed for chess, although there have been several ingenious programs for nonlearning chess machines.

A few years ago the Russian chess grandmaster Mikhail Botvinnik was quoted as saying that the day would come when a computer would play master chess. "This is of course nonsense," wrote the American chess expert Edward Lasker in an article on chess machines in the Fall 1961 issue of a magazine called *The American Chess Quarterly*. But it was Lasker who was talking nonsense. A chess computer has three enormous advantages over a human opponent: (1) it never makes a careless mistake; (2) it can analyze moves ahead at a speed much faster than a human player can; (3) it can improve its skill without limit. There is every reason to expect that a chess-learning machine, after playing thousands of games with experts, will someday develop the skill of a master. It is even possible to program a chess machine to play continuously and furiously against itself.

Its speed would enable it to acquire in a short time an experience far beyond that of any human player.

It is not necessary for the reader who would like to experiment with game-learning machines to buy an electronic computer. It is only necessary to obtain a supply of empty matchboxes and colored beads. This method of building a simple learning machine is the happy invention of Donald Michie, a biologist at the University of Edinburgh. Writing on "Trial and Error" in *Penguin Science Survey 1961*, Vol. 2, Michie describes a ticktacktoe learning machine called MENACE (Matchbox Educable Naughts And Crosses Engine) that he constructed with three hundred matchboxes.

MENACE is delightfully simple in operation. On each box is pasted a drawing of a possible ticktacktoe position. The machine always makes the first move, so only patterns that confront the machine on odd moves are required. Inside each box are small glass beads of various colors, each color indicating a possible machine play. A V-shaped cardboard fence is glued to the bottom of each box, so that when one shakes the box and tilts it, the beads roll into the V. Chance determines the color of the bead that rolls into the V's corner. First-move boxes contain four beads of each color, third-move boxes contain three beads of each color, fifth-move boxes have two beads of each color, seventh-move boxes have single beads of each color.

The robot's move is determined by shaking and tilting a box, opening the drawer and noting the color of the "apical" bead (the bead in the V's apex). Boxes involved in a game are left open until the game ends. If the machine wins, it is rewarded by adding three beads of the apical color to each open box. If the game is a draw, the reward is one bead per box. If the machine loses, it is punished by extracting the apical bead from each open box. This system of reward and punishment closely parallels the way in which animals and even humans are taught and disciplined. It is obvious that the more games MENACE plays, the more it will tend to adopt winning lines of play and shun losing lines. This makes it a legitimate learning machine, although of an extremely simple sort. It does not make (as does Samuel's checker machine) any self-analysis of past plays that causes it to devise new strategies.

Michie's first tournament with MENACE consisted of 220 games over a two-day period. At first the machine was easily trounced. After seventeen games the machine had abandoned all openings except the corner opening. After the twentieth game it was drawing consistently, so Michie began trying unsound variations in the hope of trapping it in a defeat. This paid off until the machine learned to cope with all such variations. When Michie withdrew from the contest after losing eight out of ten games, MENACE had become a master player.

Since few readers are likely to attempt building a learning machine that requires three hundred matchboxes, I have designed hexapawn, a much simpler game that requires only twenty-four boxes. The game is easily analyzed—indeed, it is trivial—but the reader is urged *not* to analyze it. It is much more fun to build the machine, then learn to play the game while the machine is also learning.

Hexapawn is played on a 3 × 3 board, with three chess pawns on each side as shown in Figure 43. Dimes and pennies can be used instead of actual chess pieces. Only two types of move are allowed: (1) A pawn may advance straight forward one square to an empty square; (2) a pawn may capture an enemy pawn by moving one square diagonally, left or right, to a square occupied by the enemy. The captured piece is removed from the board. These are the same as pawn moves in chess, except that no double move, *en passant* capture or promotion of pawns is permitted.

The game is won in any of three ways:

1. By advancing a pawn to the third row.
2. By capturing all enemy pieces.
3. By achieving a position in which the enemy cannot move.

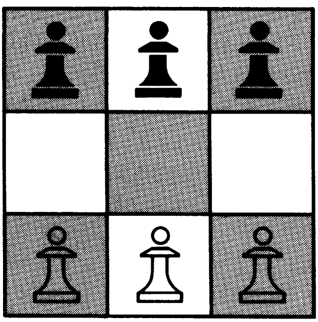

Figure 43
The game of hexapawn

Players alternate moves, moving one piece at a time. A draw clearly is impossible, but it is not immediately apparent whether the first or second player has the advantage.

To construct HER (Hexapawn Educable Robot) you need twenty-four empty matchboxes and a supply of colored beads. Small candies that come in different colors—jujubes for example—or colored popping corn also work nicely. Each matchbox bears one of the diagrams in Figure 44. The robot always makes the second move. Patterns marked "2" represent the two positions open to HER on the second move. You have a choice between a center or an end opening, but only the left end is considered because an opening on the right would obviously lead to identical (although mirror-reflected) lines of play. Patterns marked "4" show the eleven positions that can confront HER on the fourth (its second) move. Patterns marked "6" are the eleven positions that can face HER on the sixth (its last) move. (I have included mirror-image patterns in these positions to make the working easier; otherwise nineteen boxes would suffice.)

Inside each box place a single bead to match the color of each arrow on the pattern. The robot is now ready for play. Every legal move is represented by an arrow; the robot can therefore make all possible moves and only legal moves. The robot has no strategy. In fact, it is an idiot.

The teaching procedure is as follows. Make your first move. Pick up the matchbox that shows the position on the board. Shake the matchbox, close your eyes, open the drawer, remove one bead. Close the drawer, put down the box, place the bead on top of the box. Open your eyes, note the color of the bead, find the matching arrow and move accordingly. Now it is your turn to move again. Continue this procedure until the game ends. If the robot wins, replace all the beads and play again. If it loses, punish it by confiscating only the bead that represents its *last* move. Replace the other beads and play again. If you should find an empty box (this rarely happens), it means the machine has no move that is not fatal and it resigns. In this case confiscate the bead of the preceding move.

Keep a record of wins and losses so you can chart the first fifty games. Figure 45 shows the results of a typical fifty-

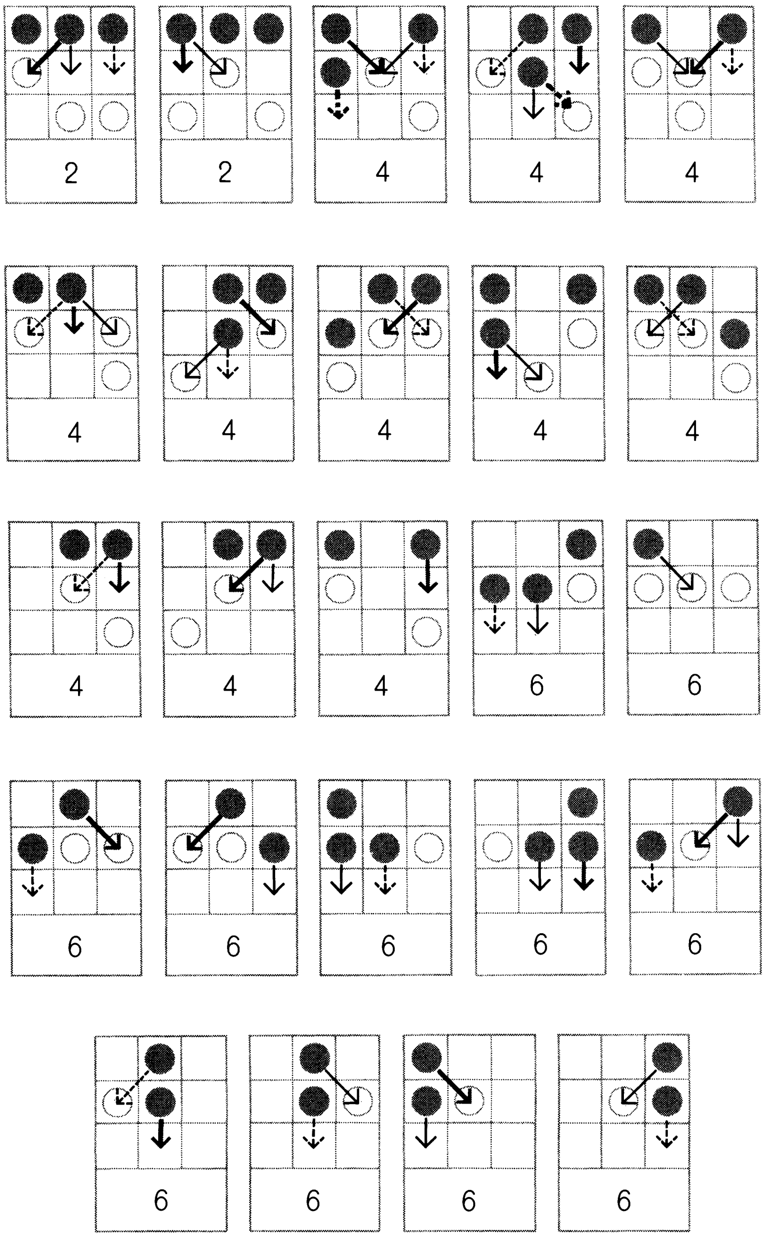

Figure 44
Labels for HER matchboxes. (The four different kinds of arrows represent four different colors.)

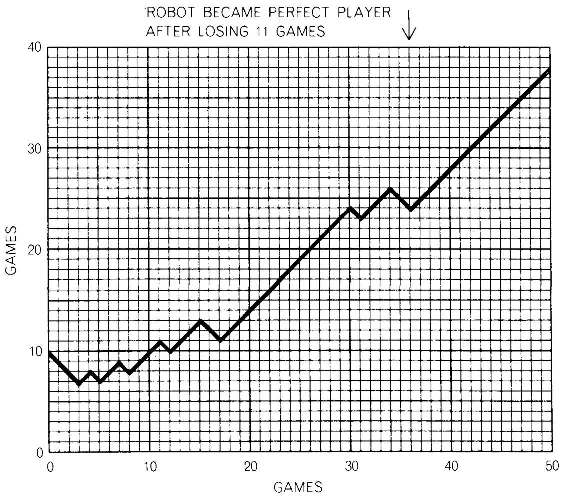

Figure 45
Learning curve for HER's first fifty games (downslant shows loss, upslant a win)

game tournament. After thirty-six games (including eleven defeats for the robot) it has learned to play a perfect game. The system of punishment is designed to minimize the time required to learn a perfect game, but the time varies with the skill of the machine's opponent. The better the opponent, the faster the machine learns.

The robot can be designed in other ways. For example, if the intent is to maximize the number of games that the machine wins in a tournament of, say, twenty-five games, it may be best to reward (as well as punish) by adding a bead of the proper color to each box when the machine wins. Bad moves would not be eliminated so rapidly, but it would be less inclined to make the bad moves. An interesting project would be to construct a second robot, HIM (Hexapawn Instructable Matchboxes), designed with a different system of reward and punishment but equally incompetent at the start of a tournament. Both machines would have to be enlarged so they could make either first or second moves. A tournament could then be played between HIM and HER, alternating the first move, to see which machine would win the most games out of fifty.

Similar robots are easily built for other games. Stuart C. Hight, director of research studies at the Bell Telephone Laboratories in Whippany, New Jersey, recently built a matchbox learning machine called NIMBLE (Nim Box Logic Engine) for playing Nim with three piles of three counters each. The robot plays either first or second and is rewarded or punished after each game. NIMBLE required only eighteen matchboxes and played almost perfectly after thirty games. For an analysis of the game of Nim, see Chapter 15 of my *Scientific American Book of Mathematical Puzzles & Diversions.*

By reducing the size of the board the complexity of many familiar games can be minimized until they are within the scope of a matchbox robot. The game of go, for example, can be played on the intersections of a 2 × 2 checkerboard. The smallest nontrivial board for checkers is shown in Figure 46. It should not be difficult to build a matchbox machine that would learn to play it. Readers disinclined to do this may enjoy analyzing the game. Does either side have a sure win or will two perfect players draw?

When chess is reduced to the smallest board on which all legal moves are still possible, as shown in Figure 46, the complexity is still far beyond the capacity of a matchbox machine. In fact, I have found it impossible to determine which player, if either, has the advantage. Minichess is recommended for computer experts who wish to program a simplified chess-learning machine and for all chess players who like to sneak in a quick game during a coffee break.

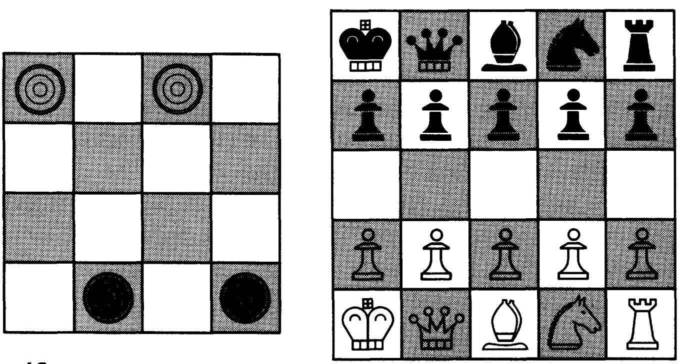

Figure 46
Matchbox machines can be built for minicheckers [left] *but not for minichess* [right]

ADDENDUM

Many readers who experimented with matchbox learning machines were kind enough to write to me about them. L. R. Tanner, at Westminster College, Salt Lake City, Utah, made good use of HER as a concession at a college carnival. The machine was designed to learn by rewards only, so that customers would always have a chance (though a decreasing one) of winning, and prizes to winners were increased in value as HER became more proficient.

Several readers built two matchbox machines to be pitted against each other. John Chambers, Toronto, called his pair THEM (Two-way Hexapawn Educable Machines). Kenneth W. Wiszowaty, science teacher at Phillip Rogers Elementary School, Chicago, sent me a report by his seventh-grade pupil, Andrea Weiland, on her two machines which played against each other until one of them learned to win every time. John House, Waterville, Ohio, called his second machine RAT (Relentless Auto-learning Tyrant), and reported that after eighteen games RAT conceded that HER would win all subsequent games.

Peter J. Sandiford, director of operations research for Trans-Canada Air Lines, Montreal, called his machines Mark I and Mark II. As expected, it took eighteen games for Mark I to learn how to win every time and Mark II to learn how to fight the longest delaying action. Sandiford then devised a devilish plan. He arranged for two students, a boy and a girl from a local high school mathematics club, who knew nothing about the game, to play hexapawn against each other after reading a handout describing the rules. "Each contestant was alone in a room," writes Sandiford, "and indicated his moves to a referee. Unknown to the players the referees reported to a third room containing the jellybean computers and scorekeepers. The players thought they were playing each other by remote control, so to speak, whereas they were in fact playing independently against the computers. They played alternately black and white in successive games. With much confusion and muffled hilarity we in the middle tried to operate the computers, keep the games in phase, and keep the score."

The students were asked to make running comments on their own moves and those of their opponent. Some sample remarks:

"It's the safest thing to do without being captured; it's almost sure to win."

"He took me, but I took him too. If he does what I expect, he'll take my pawn, but in the next move I'll block him."

"Am I stupid!"

"Good move! I think I'm beat."

"I don't think he's really thinking. By now he shouldn't make any more careless mistakes."

"Good game. She's getting wise to my action now."

"Now that he's thinking, there's more competition."

"Very surprising move . . . couldn't he see I'd win if he moved forward?"

"My opponent played well. I guess I just got the knack of it first."

When the students were later brought face to face with the machines they had been playing, they could hardly believe, writes Sandiford, that they had not been competing against a real person.

Richard L. Sites, at M.I.T., wrote a FORTRAN program for an IBM 1620 so that it would learn to play Octapawn, a 4×4 version of hexapawn that begins with four white pawns on the first row and four black pawns on the fourth row. He reports that the first player has a sure win with a corner opening. At the time of his writing, his program had not yet explored center openings.

Judy Gomberg, Maplewood, New Jersey, after playing against a matchbox machine that she built, reported that she learned hexapawn faster than her machine because "every time it lost I took out a candy and ate it."

Robert A. Ellis, at the computing laboratory, Ballistics Research Laboratories, Aberdeen Proving Ground, Maryland, told me about a program he wrote for a digital computer which applied the matchbox-learning technique to a ticktacktoe-learning machine. The machine first plays a stupid game, choosing moves at random, and is easily trounced by human opponents. Then the machine is allowed to play two thousand games against itself (which it does in two or three minutes), learning as it goes. After that, the machine plays an excellent strategy against human opponents.

My defense of Botvinnik's remark that computers will some day play master chess brought a number of irate letters from chess players. One grandmaster assured me that Botvinnik was speaking with tongue in cheek. The interested reader can judge for himself by reading a translation of Botvinnik's speech (which originally appeared in *Komsomolskaya Pravda,* January 3, 1961) in *The Best in Chess,* edited by I. A. Horowitz and Jack Straley Battell (New York: Dutton, 1965), pages 63–69. "The time will come," Botvinnik concludes, "when mechanical chessplayers will be awarded the title of International Grandmaster . . . and it will be necessary to promote two world championships, one for humans, one for robots. The latter tournament, naturally, will not be between machines, but between their makers and program operators."

An excellent science-fiction story about just such a tournament, Fritz Leiber's "The 64-Square Madhouse," appeared in *If*, May 1962, and has since been reprinted in Leiber's *A Pail of Air* (New York: Ballantine, 1964). Lord Dunsany, by the way, has twice given memorable descriptions of chess games played against computers. In his short story "The Three Sailors' Gambit" (in *The Last Book of Wonder*) the machine is a magic crystal. In his novel *The Last Revolution* (a 1951 novel about the computer revolution that has never, unaccountably, been published in the United States) it is a learning computer. The description of the narrator's first game with the computer, in the second chapter, is surely one of the funniest accounts of a chess game ever written.

The hostile reaction of master chess players to the suggestion that computers will some day play master chess is easy to understand; it has been well analyzed by Paul Armer in a Rand report (p-2114–2, June 1962) on *Attitudes Toward Intelligent Machines*. The reaction of chess players is particularly amusing. One can make out a good case against computers writing top-quality music or poetry, or painting great art, but chess is not essentially different from ticktacktoe except in its enormous complexity, and learning to play it well is precisely the sort of thing computers can be expected to do best.

Master checker-playing machines will undoubtedly come first. Checkers is now so thoroughly explored that games between champions almost always end in draws, and in order to add interest to such games, the first three moves are now chosen by chance. Richard Bellman, writing "On the Application of Dynamic Programming to the Determination of Optimal Play in Chess and Checkers," *Proceedings of the National Academy of Sciences,* Vol. 53 (February 1965), pages 244–47, says that "it seems safe to predict that within ten years checkers will be a completely decidable game."

Chess is, of course, of a different order of complexity. One suspects it will be a long time before one can (so goes an old joke in modern dress) play the first move of a chess game against a computer and have the computer print, after a period of furious calculation, "I resign." In 1958 some responsible mathematicians predicted that within ten years computers would be playing master chess, but this proved to be wildly overoptimistic. Tigran Petrosian, when he became world chess champion, was quoted in *The New York Times* (May 24, 1963) as expressing doubts that computers would play master chess within the next fifteen or twenty years.

Hexapawn can be extended simply by making the board wider but keeping it three rows deep. John R. Brown, in his paper "Extendapawn—An Inductive Analysis," *Mathematics Magazine,* Vol. 38, November 1965, pages 286–99, gives a complete analysis of this game. If n is the number of columns, the game is a win for the first player if the final digit of n is 1, 4, 5, 7 or 8. Otherwise the second player has the win.

ANSWERS

The checker game on the 4×4 board is a draw if both sides play as well as possible. As shown in Figure 47, Black has a choice of three openings: (1) C5, (2) C6, (3) D6.

The first opening results in an immediate loss of the game when White replies A3. The second opening leads to a draw regardless of how White replies. The third opening is Black's strongest. It leads to a win if White replies A3 or B3. But White can reply B4 and draw.

With respect to the 3×3 simplified go game, also mentioned as suitable for a matchbox learning machine, I am assured by Jay Eliasberg, vice-president of the American Go Association, that the first player has a sure win if he plays on the center point of the board and rationally thereafter.

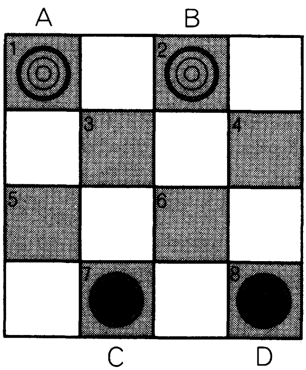

Figure 47
Checker game is drawn if played rationally

The 4 × 4 checker game is trivial, but when the board is enlarged to 5 × 5 the result is both challenging and surprising. Robert L. Caswell, a chemist with the United States Department of Agriculture, wrote to me about this minigame, which he said had earlier been proposed to him. The game begins with three white checkers on the first row, three black checkers on the fifth row. All standard rules obtain, with black moving first. One might guess the game to be drawn if played rationally, but the absence of "double corners" where kings can move back and forth makes this unlikely. Caswell discovered that not only does one side have a sure win but, if the loser plays well, the final win is spectacular. Rather than spoil the fun, I leave it to the reader to analyze the game and decide which player can always win.