

Title: Matchbox Educable Noughts and Crosses Engine

URL: [https://en.wikipedia.org/wiki/Matchbox\\_Educable\\_Noughts\\_and\\_Crosses\\_Engine](https://en.wikipedia.org/wiki/Matchbox_Educable_Noughts_and_Crosses_Engine)

PageID: 63983302

Categories: Category:Machine learning, Category:Mechanical computers

Source: Wikipedia (CC BY-SA 4.0).

-----

The Matchbox Educable Noughts and Crosses Engine (sometimes called the Machine Educable Noughts and Crosses Engine or MENACE ) was a mechanical computer made from 304 matchboxes designed and built by artificial intelligence researcher Donald Michie in 1961. It was designed to play human opponents in games of noughts and crosses (tic-tac-toe) by returning a move for any given state of play and to refine its strategy through reinforcement learning . This was one of the first types of artificial intelligence.

Michie did not have a computer readily available, so he worked around this restriction by building it out of matchboxes. The matchboxes used by Michie each represented a single possible layout of a noughts and crosses grid. When the computer first played, it would randomly choose moves based on the current layout. As it played more games, through a reinforcement loop, it disqualified strategies that led to losing games, and supplemented strategies that led to winning games. Michie held a tournament against MENACE in 1961, wherein he experimented with different openings.

Following MENACE's maiden tournament against Michie, it demonstrated successful artificial intelligence in its strategy. Michie's essays on MENACE's weight initialisation and the BOXES algorithm used by MENACE became popular in the field of computer science research. Michie was honoured for his contribution to machine learning research, and was twice commissioned to program a MENACE simulation on an actual computer.

#### Origin

Donald Michie (1923–2007) had been on the team decrypting the German Tunny Code during World War II . [ 1 ] Fifteen years later, he wanted to further display his mathematical and computational prowess with an early convolutional neural network . Since computer equipment was not obtainable for such uses, [ 2 ] and Michie did not have a computer readily available, [ 2 ] he decided to display and demonstrate artificial intelligence in a more esoteric format and constructed a functional mechanical computer out of matchboxes and beads. [ 3 ] [ 4 ]

MENACE was constructed as the result of a bet with a computer science colleague who postulated that such a machine was impossible. [ 5 ] Michie undertook the task of collecting and defining each matchbox as a "fun project", later turned into a demonstration tool. [ 6 ] Michie completed his essay on MENACE in 1963, [ 4 ] "Experiments on the mechanization of game-learning", as well as his essay on the BOXES Algorithm, written with R. A. Chambers [ 6 ] and had built up an AI research unit in Hope Park Square, Edinburgh , Scotland . [ 7 ]

MENACE learned by playing successive matches of noughts and crosses. Each time, it would eliminate a losing strategy by the human player confiscating the beads that corresponded to each move. [ 4 ] It reinforced winning strategies by making the moves more likely, by supplying extra beads. [ 8 ] This was one of the earliest versions of the Reinforcement Loop , the schematic algorithm of looping the algorithm, dropping unsuccessful strategies until only the winning ones remain. [ 4 ] This model starts as completely random, and gradually learns. [ 9 ]

#### Composition

MENACE was made from 304 matchboxes glued together in an arrangement similar to a chest of drawers. [ 10 ] Each box had a code number, which was keyed into a chart. This chart had drawings of tic-tac-toe game grids with various configurations of X , O , and empty squares, [ 4 ] corresponding to all possible permutations a game could go through as it progressed. [ 11 ] After removing duplicate arrangements (ones that were simply rotations or mirror images of other configurations), MENACE used 304 permutations in its chart and thus that many matchboxes. [ 12 ]

Each individual matchbox tray contained a collection of coloured beads. [ 13 ] Each colour represented a move on a square on the game grid, and so matchboxes with arrangements where positions on the grid were already taken would not have beads for that position. Additionally, at the front of the tray were two extra pieces of card in a "V" shape, [ 10 ] the point of the "V" pointing at the front of the matchbox. [ 11 ] Michie and his artificial intelligence team called MENACE's algorithm "Boxes", [ 7 ] after the apparatus used for the machine. The first stage "Boxes" operated in five phases, each setting a definition and a precedent for the rules of the algorithm in relation to the game. [ 14 ]

#### Operation

MENACE played first, as O, since all matchboxes represented permutations only relevant to the "X" player. [ 12 ] [ 17 ] To retrieve MENACE's choice of move, the opponent or operator located the matchbox that matched the current game state, or a rotation or mirror image of it. For example, at the start of a game, this would be the matchbox for an empty grid. The tray would be removed and lightly shaken so as to move the beads around. [ 4 ] Then, the bead that had rolled into the point of the "V" shape at the front of the tray was the move MENACE had chosen to make. [ 4 ] Its colour was then used as the position to play on, and, after accounting for any rotations or flips needed based on the chosen matchbox configuration's relation to the current grid, the O would be placed on that square. Then the player performed their move, the new state was located, a new move selected, and so on, until the game was finished. [ 12 ]

When the game had finished, the human player observed the game's outcome. As a game was played, each matchbox that was used for MENACE's turn had its tray returned to it ajar, and the bead used kept aside, so that MENACE's choice of moves and the game states they belonged to were recorded. Michie described his reinforcement system with "reward" and "punishment". Once the game was finished, if MENACE had won, it would then receive a "reward" for its victory. The removed beads showed the sequence of the winning moves. [ 17 ] These were returned to their respective trays, easily identifiable since they were slightly open, as well as three bonus beads of the same colour. [ 11 ] In this way, in future games MENACE would become more likely to repeat those winning moves, reinforcing winning strategies. If it lost, the removed beads were not returned, "punishing" MENACE, and meaning that in future it would be less likely, and eventually incapable if that colour of bead became absent, to repeat the moves that cause a loss. [ 3 ] [ 8 ] If the game was a draw, one additional bead was added to each box. [ 11 ]

#### Results in practice

##### Optimal strategy

Noughts and crosses has a well-known optimal strategy. [ 18 ] A player must place their symbol in a way that blocks the other player from achieving any rows while simultaneously making a row themselves. However, if both players use this strategy, the game always ends in a draw. [ 18 ] If the human player is familiar with the optimal strategy, and MENACE can quickly learn it, then the games will eventually only end in draws. The likelihood of the computer winning increases quickly when the computer plays against a random-playing opponent. [ 3 ]

When playing against a player using optimal strategy, the odds of a draw grow to 100%. In Donald Michie's official tournament against MENACE in 1961 [ 4 ] he used optimal strategy, and he and the computer began to draw consistently after twenty games. Michie's tournament [ 19 ] had the following milestones: Michie began by consistently opening with "Variant 0", the middle square. At 15 games, MENACE abandoned all non-corner openings. At just over 20, Michie switched to consistently using "Variant 1", the bottom-right square. At 60, he returned to Variant 0. As he neared 80 games, he moved to "Variant 2", the top-middle. At 110, he switched to "Variant 3", the top right. At 135, he switched to "Variant 4", middle-right. At 190, he returned to Variant 1, and at 210, he returned to Variant 0.

The trend in changes of beads in the "2" boxes runs: [ 19 ]

##### Correlation

Depending on the strategy employed by the human player, MENACE produces a different trend on scatter graphs of wins. [ 4 ] Using a random turn from the human player results in an almost-perfect positive trend. Playing the optimal strategy returns a slightly slower increase. [ 3 ] The reinforcement does not create a perfect standard of wins; the algorithm will draw random uncertain conclusions each time. After the  $j$ -th round, the correlation of near-perfect play runs:

$$\frac{1 - D}{D - D^{(j+2)}} \sum_{i=0}^j D^{(j+1-i)} V_i$$

Where  $V_i$  is the outcome (+1 is win, 0 is draw and -1 is loss) and  $D$  is the decay factor (average of past values of wins and losses). Below,  $M_n$  is the multiplier for the  $n$ -th round of the game. [ 4 ]

## Legacy

Donald Michie's MENACE proved that a computer could learn from failure and success to become good at a task. [ 17 ] It used what would become core principles within the field of machine learning before they had been properly theorised. For example, the combination of how MENACE starts with equal numbers of types of beads in each matchbox, and how these are then selected at random, creates a learning behaviour similar to weight initialisation in modern artificial neural networks . [ 20 ] In 1968, Donald Michie and R.A Chambers made another BOXES-based algorithm called GLEE (Game Learning Expectimaxing Engine) which had to learn how to balance a pole on a cart. [ 21 ]

After the resounding reception of MENACE, Michie was invited to the US Office of Naval Research, where he was commissioned to build a BOXES-running program for an IBM Computer for use at Stanford University . [ 22 ] Michie created a simulation program of MENACE on a Pegasus 2 computer with the aid of D. Martin. [ 4 ] There have been multiple recreations of MENACE in more recent years, both in its original physical form and as a computer program. [ 12 ] Its algorithm was later converged into Christopher Watkin's Q-Learning algorithm. [ 23 ] Although not as a functional computer, in examples of demonstration, MENACE has been used as a teaching aid for various neural network classes, [ 24 ] [ 25 ] [ 26 ] including a public demonstration from University College London researcher Matthew Scroggs. [ 27 ] [ 28 ] A copy of MENACE built by Scroggs was featured in the 2019 Royal Institution Christmas Lectures , [ 29 ] [ 30 ] and in a 2023 episode of QI XL . [ 31 ]

MENACE is referenced in Fred Saberhagen 's 1963 short story Without A Thought , and Thomas J Ryan's 1977 novel The Adolescence of P-1 . [ 32 ] In her 2023 book The Future , author Naomi Alderman includes a fictional lecture with a detailed overview of MENACE.

See also

Hexapawn

Reinforcement learning

References

Sources

Michie, D.; Chambers, R. A. (1968), "BOXES: An Experiment in Adaptive Control", Machine Intelligence , Edinburgh, UK: Oliver and Boyd, S2CID 18229198 – via Semantic Scholar , Michie and R. A Chambers' paper on the AI implications of BOXES and MENACE.

Russell, David W. (2012), The BOXES Methodology: Black Box Dynamic Control , Springer London, ISBN 978-1849965286 , a book on the "Boxes" algorithm employed by MENACE.

External links

Online simulation of MENACE