

Computer chess: From idea to DeepMind¹

H. Jaap van den Herik

Leiden Centre of Data Science, Universiteit Leiden

Abstract. Computer chess has stimulated human imagination over some two hundred and fifty years. In 1769 Baron Wolfgang von Kempelen promised Empress Maria Theresia in public: “I will invent a machine for a more compelling spectacle [than the magnetism tricks by Pelletier] within half a year.” The idea of an intelligent chess machine was born. In 1770 the first demonstration was given.

The real development of artificial intelligence (AI) began in 1950 and contains many well-known names, such as Turing and Shannon. One of the first AI research areas was chess. In 1997, a high point was to be reported: world champion Gary Kasparov had been defeated by Deep Blue. The techniques used included searching, knowledge representation, parallelism, and distributed systems. Adaptivity, machine learning and the recently developed deep learning mechanism were only later on added to the computer chess research techniques.

The major breakthrough for games in general (including chess) took place in 2017 when (1) the AlphaGo Zero program defeated the world championship program AlphaGo by 100-0 and (2) the technique of deep learning also proved applicable to chess. In the autumn of 2017, the Stockfish program was beaten by AlphaZero by 28-0 (with 72 draws, resulting in a 64-36 victory). However, the end of the disruptive advance is not yet in reach. In fact, we have just started. The next milestone will be to determine the theoretical game value of chess (won, draw, or lost). This achievement will certainly be followed by other surprising developments.



Fig. 1. Right: Claude Shannon 1916–2001 demonstrating his (home-made) electric chess automation to Chess Grandmaster Edward Lasker.



Fig. 2. Enigma a-320 (WW II) “Digital computers applied to games” 1953.



Fig. 3. Alan Turing 1912–1954 “Computing machinery and intelligence” 1950.

¹This article is an elaboration of my keynote lecture delivered at the 10th International Conference on Computers and Games (CG 2018) in the National Taipei University in Taipei, Taiwan on July 9, 2018.

THE RUN-UP

This article provides an overview of the development of computer chess. What has been required to defeat the human world champion? What is the role of intuition? What have been the major breakthroughs? And which top researchers have spent (much) of their time on computer chess? These are all interesting topics and there is only limited space to discuss them. For an adequate overview I have taken the chronology and the above questions as a guideline and tried to place the most important items in the spot lights. The division of the article is therefore in time periods, as follows: 1770–1940, and then every ten years 1940–1950, 1950–1960, . . . until 2010–2016. Then in the period 2015–2017 we have Intermezzo and in Autumn 2017 the Apotheosis will follow.

1770–1940

“I will invent a machine for a more compelling spectacle within half a year”, Hofrat von Kempelen promised Empress Maria Theresia in public after a presentation and demonstration by Pelletier in 1769 about magnetism. It is clear that von Kempelen did not find the presentation very inspiring. It is not clear from the written reports whether he had already in mind a “chess machine” or that he had the conviction that he certainly would have in a later stage a brilliant idea on an innovation. The fact that the first ‘machine’, which was called “The Turk”, contained a human being does not matter. The idea was born and soon it would conquer the world. Surprisingly, it took until 1834 before the deception came out. Edgar Allen Poe made it public in 1836. Yet, the idea was so attractive that in 1868 a new chess machine was introduced under the name “Ajeeb” and two years later another one under the name “Mephisto”.



Fig. 4. 1769 Wolfgang von Kempelen.

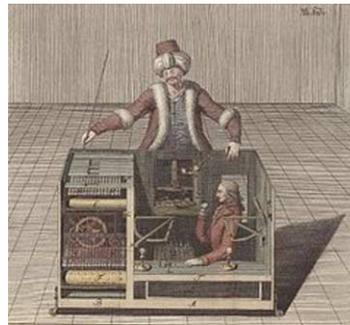


Fig. 5. 1770 De Turk.

Meanwhile, Babbage (1791–1871) developed his difference engine and analytical engine. Already in that time, he believed that the analytical engine if it was properly constructed could play chess. Since the realization of his belief was still far away, Babbage described a Tic-tac-toe playing program. Again, the description was fine, but there was never an actually playing program.

The world had to wait for the substantiation of a really important contribution. Torres y Quevedo (1852–1936) was the first to make such a contribution. He developed around 1890 an electromechanical machine that was able to play the endgame King and Rook against King. In the starting position, the White King (WK) was on a8 and the White Rook (WR) on b7. The Black King (BK) could be placed anywhere. After 24 years (in 2014), Torres significantly improved the mechanism. In 1953

W.L. (Wim) van der Poel transferred Torres' latest algorithm into a computer program. In a demonstration of that program to Max Euwe, the former world champion found a flaw in the algorithm so that the mating procedure required more than 50 moves. This meant that the game was drawn.

In the beginning of the 20th century, there were three mathematicians of fame who also dealt with fundamental ideas about chess, namely Zermelo (1912), König (1927) and Euwe (1929). They investigated applications of set theory on chess. In summary, it came down to the following proposition: If a position is theoretically a won position then the advantage can also be enforced into a win in a finite number of moves.

In the 1920s von Neumann (1903–1957) appeared on the stage. His publication in 1928 titled “Zur Theorie der Gesellschaftspiele” contains the basis for the minimax method. However, the work only gained recognition after it was published in collaboration with Morgenstern in the book “Theory of Games and Economic Behavior” in 1944.

A real breakthrough idea of using a computer for decision making came from the German engineer Konrad Zuse (1910–1995). In 1934 he formulated the first original thoughts about a potential program. Then in the period 1935–1937 he built the first relay calculator that was controlled by a program. Zuse's Z3 computer, which was operational in 1941, is now recognized as Turing complete and is therefore the first digital computer. (Note that the first computer design is attributed to the John Vincent Atanasoff.) Zuse's importance for computer chess lies in the design of the programming language Plankalkül. He wrote about this: “Um klarer zu sehen, wählte ich, u.a., das Schachspiel as Modellfall aus und träumte davon, eines Tages den Schachweltmeister mit einem Computer zu besiegen.”

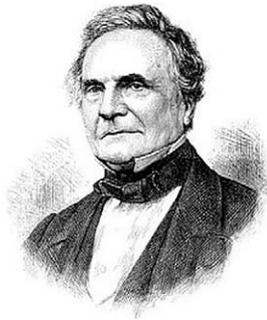


Fig. 6. Charles Babbage 1791–1871, Analytical engine.



Fig. 7. John von Neumann 1903–1957, Minimax method.



Fig. 8. Konrad Zuse 1910–1995, Z3 machine.

1940–1950

The first half of decade 1940–1950 was dominated by the Second World War. No time for chess, or still? But then somewhere in connection with other developments? Indeed, the British government had created a research center in Bletchley Park, with the prerequisite for admission: being smart and creative. A relationship with chess was an implicit recommendation as appeared from the list of residents of cabin 8: C.H.O. d'Alexander, A.M. Turing, I.J. Good, H. Golombek, P. Hilton, and D. Michie (there were more). If no code-breaking was done, there was brainstorming about the question: how a computer could play chess? Jack Good later admitted the following: “This procedure [minimax with evaluation] seemed obvious to us long before Shannon's paper, and I made the mistake of thinking

it was not worth publishing.” So, it could happen that Turing (1912–1954) paid more attention to other issues, even after the war. Turing and Shannon (1916–2001) both spent a year in Princeton as guest of John von Neumann. They even worked at Bell Laboratories at the same time. In that environment they probably spoke to each other some 15 to 20 times during lunch, according to Shannon, but they worked in different departments and were not allowed to talk about their code work because it was a secret (America vs. England). So, they have never exchanged their ideas on computer chess. Shannon remembered that their conversations were mostly subject to the relationship “human brain–mechanical brain”. Shannon’s idea was: “Our brain is basically a machine, although it is complicated to simulate.”

Shannon was the first to publish on computer chess. He had a lecture on computer chess at Bell Labs in 1949 and published in 1950 his famous article “Programming a Computer for Playing Chess” in *Philosophical Magazine* (where else at that time?). It contained a description of three types of strategy (Type A: brute force; Type B: selected search; and chess master-like reasoning (later called Type C)). Moreover, there was a design of evaluation functions with weight factors, stability search, the two-dimensional board representation, and a philosophical consideration of playing a reasonably brilliant game.

Turing followed after three years. He first laid the foundation for artificial intelligence with his article “Computing Machinery and Intelligence” published also in 1950. In 1953 he published: “Digital Computers Applied to Games”. It was included in B.V. Bowden’s book, titled *Faster than Thought*. In addition to the evaluation of the pieces, Turing also described the evaluation of the positional characteristics. Moreover, he emphasized the notion of “quiescence” which was denoted as “stability search” by Shannon. Interestingly, Turing, together with the mathematician David Champernowne, also performed hand simulations. Their program was called TUROCHAMP.

Although Shannon and Turing were, without any discussion, the first two major researchers in the field of computer chess, they were still preceded by the Hungarian researcher Tihamér Nemes (1895–1960) who had similar ideas and published them in *Magyar Sakkvilág* and *Chess* in 1949. Later he published detailed descriptions in *Acta Technica*. However, because of the cold war, it took a long time for this information penetrated to the West.

The ideas by Shannon and also by Turing were partly based on (1) previous experimental-psychological research by A.D. de Groot (1914–2006), who obtained his doctorate in 1946 on the thesis “Het Denken van den Schaker”, as well as on (2) a Note by Norbert Wiener (1894–1964) included in his book *Cybernetics, or control and communication in the animal and the machine*. Wiener believed that it should be possible for a machine to play a ‘good’ game of chess.

1950–1960

After the turbulent flood of ideas on the part of Shannon and Turing it was quiet at the computer chess front. Two smaller examples came from Turing’s environment. Dieter Gunther Prinz was able to implement the first chess program on a MADM computer at the University of Manchester in 1951. The program was able to solve a mate-in-two problem. To speed up the search program, Prinz developed the killer heuristic. That turned out to be a particularly valuable contribution for all programs. Around this time, Christopher Strachey, a talented employee at Turing’s lab was working on draughts (checkers on the 100 squares).

Somewhat later, in Los Alamos, America, a team of researchers on nuclear material became interested in intelligent systems that could play chess. The Institute Director Robert Oppenheimer allowed them

to work on such “frivolities” during the lunch breaks. However, in that time computer restrictions led to a limitation of the board size. So, the 8×8 board was reduced to a 6×6 board (the Bishops disappeared). The consequence was that the relationship between Knight and Rook changed drastically. The first game was played between the program and a secretarial assistant who had learned the course of the pieces only shortly in advance to the game. The program won. The second game was won by the human. Although Martin Krushkal (Princeton) was obliged to give the program the advantage of a Queen, he won the very exciting and long-lasting game. In the report of these two games there is also mention of computer chess investigations by the Russian researcher V.M. Kurochin. That research was later continued in the Steklov Mathematical Institute (see below).

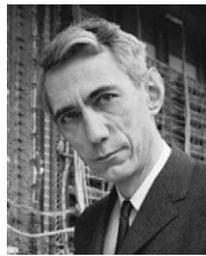
The great breakthrough in this decade, however, was the Dartmouth conference that took place in 1956. John McCarthy was the organizer along with Marvin Minsky, Nathaniel Rochester, and Claude Shannon. Their self-given assignment was to investigate whether a computer in the future could simulate ‘intelligence’ and ‘learning’. Among the participants we see Herb Simon, Allen Newell, Arthur Samuel and Alex Bernstein. In 1958, the Bernstein group (IBM) succeeded in developing a chess program that was able to play a game on an 8×8 board. The shareholders of IBM were “not amused” that their money was spent on such frivolities and T.J. Watson Jr. ordered that the investigation was stopped.



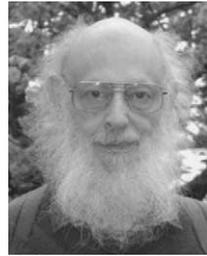
John MacCarthy



Marvin Minsky



Claude Shannon



Ray Solomonoff



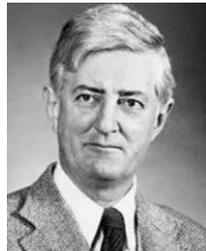
Alan Newell



Herbert Simon



Arthur Samuel



Oliver Selfridge



Nathaniel Rochester



Trenchard More

Fig. 9. Dartmouth Conference 1956, The Founding Fathers of AI.

This stop was temporarily (as we know now) and addressed mainly the relation with business opportunities. The computer chess research at the American universities continued since it was inspired by the Dartmouth conference. The research was intensive and competitive. For instance, there are three groups that claim the discovery of the alpha–beta algorithm (a search algorithm based on the minimax method with a sophisticated pruning mechanism). Newell, Shaw and Simon were the first to publish about it (1958). It later turned out that they had developed the simple cutting procedure. Their NSS program was based on De Groot’s ideas. The program was written in a higher programming language (also new at that time). Together with George Baylor, Herb Simon additionally implemented the null move in the Mater III program to discover if there was mate-in-1. These two techniques (alpha–beta and null move) have become the beginning of a powerful pruning mechanism in the search process.

Yet, the name alpha–beta algorithm originates from McCarthy. He came to the idea when he listened to Bernstein’s lecture at the Dartmouth Conference. Bernstein explained the minimax and McCarthy believed there should be a more advanced method to achieve the same result. In his own words, he implemented the ideas and saw that it worked well and did not find it worthwhile to write it down (the algorithm was in his words “self-evident”). This is a funny observation particularly with respect to the third discoverer. For this story, the reader should know that when Samuel wrote his seminal standard work on AI and checkers (draughts on 64 squares) in 1959, he paid no attention to the alpha–beta algorithm for the same reason. He did so only in 1967, although he had developed the technique already in the late 1940s. In 1967, he in fact produced the first standard work on the alpha–beta algorithm. Still, the algorithm was at that time regarded as a heuristic and not as a “real” algorithm. In the 1959 article, Samuel emphasized learning the weight functions in the evaluation function. That was a breakthrough in thinking about learning (see the ambitions of the Dartmouth conference). With that article, Samuel then laid the foundation for the current successes in machine learning.

1960–1970

In this decade the research shifted from clever searching to knowledge representation. In this respect, the one-dimensional representation of the chess board can be seen as a front runner of the shift. Especially John McCarthy was a supporter of more knowledge in programs. At the MIT in 1961 he found bachelor student Alan Kotok willing to develop a chess program of this type. In addition to his attention for the evaluation function, Kotok implemented also one move generator (instead of two). After completing his studies, Kotok disappeared from the computer chess scene, but McCarthy took Kotok’s program to Stanford and worked on the development of a plausible move generator (Shannon B strategy).

Meanwhile, at the ITEP (Institute for Theoretical and Experimental Physics) in Moscow, the development of a chess program also continued under the direction of George M. Adelson Velsky. The program used an M-20 computer and was developed according to the Shannon A strategy (brute force). On November 22, 1966, a match of four games started between Stanford and ITEP. The games were simultaneously played. The moves were passed on by telegraph. After 40 moves, a break was agreed, and the positions would then be adjudicated. The first victory came on March 10, 1967. ITEP mated Stanford on move 19. ITEP also won the second game. Two games were adjudicated. ITEP had an advantage in both games. The final score was agreed on 3-1 (both games were declared draw). The outcome was seen as a success for brute force.

This historic meeting was followed by a new breakthrough initiated by Richard D. Greenblatt, who was also (like Alan Kotok) a student at MIT. In mid-November 1966, Greenblatt started developing the Mac Hack Six chess program. In February 1967, he submitted the program as a participant in a local chess tournament. It was the first time that a non-human entered the tournament. Mac Hack Six scored 0.5 out of 5 games, accounting for a rating of 1243 (Class D). Four new ideas were implemented in the program: more knowledge in the evaluation function, more knowledge in the plausible set-up generator, an opening library and the implementation of secondary search (a kind of progressive deepening, taken from De Groot).

In this decade there were also two former world chess champions who became involved in computer chess as well as the respected researcher I.J. Good (see Section 1940–1950). First of all, Max Euwe (world champion 1935–1937) took the lead of the major research project Euratom. The intention was to come to an automatic translation program. As a first step, the team aimed to create a program that

could play a reasonable game of chess. However, that was too high a goal and the project ended with the reporting of two dissident opinions. No result is also a result.

The second former world champion was Botvinnik (world champion 1948–1963). Botvinnik had the idea that Shannon's Type-A and Type-B strategy were not the right way. He wanted to imitate the grandmaster's thinking. It was good for the development of ideas about artificial intelligence, but it did not contribute much to the increase of the playing strength.

Then there was the theoretical foundation by I.J. Good (1968): "A Five-Year Plan for Automatic Chess". It was a good and solid article. Although experts saw Turing's hand, it did not lead to anything substantial.

In addition to these three researchers of fame, there were several researchers, let us say aficionados of chess and computers, who contributed. We call Frank Anderson and Bob Cody (endgames), Fischer and Schneider (implementation of the draw rule), Barbara J. Huberman (a mating procedure with Bishop and Knight).

Actually, the decade was characterized by Greenblatt and Dreyfus. Greenblatt with his breakthrough (even Robert J. Fischer has played against the program) and Dreyfus with his opposition towards Artificial Intelligence. Dreyfus started in 1965 with the publication of "Alchemy and Artificial Intelligence" that became the core of his later publication *What Computers Cannot Do; A Critique of Artificial Reason*. Dreyfus challenged Simon's 1957 prediction "Unless the rules bar it from competition, within ten years a computer program will be the world champion" by stating in 1965 "still no chess program can play even amateur chess, and the world championship is only 2 years away." In 1967 Seymour Papert invited the AI critic Dreyfus for a game against Mac Hack Six. Dreyfus lost in 37 moves. Papert: "It says nothing, only that Dreyfus is not good at chess".

1970–1980

The next scientific challenge in this decade was: would it be possible to integrate advanced searching and adequate knowledge representation? And if possible, to what extent should it be executable? Maybe it was too high a goal for the whole game, but in endgames it was hoped that even perfect knowledge was reachable. Thomas Ströhlein thought at least so. After all, it only concerns storing all positions with their corresponding theoretical value (won, drawn, lost), followed by a number that indicates how many moves from that particular position were necessary to be played to reach the end result. This is the simplest form of a database. In 1970 Thomas Ströhlein was the first researcher to introduce this construction into the computer chess world. He did so for the endgames KRK, KQK, KRKB, KRKN, KQKR (K = King, R = Rook, Q = Queen, B = Bishop, N = kNight). Thereafter many researchers followed his example. New metrics emerged, such as distance to mate, and distance to conversion. The subtlety lies in part in the drawing rule that within 50 moves a piece has to be captured or a pawn should have been moved (apart from specifically mentioned exceptions). A famous example is Timman–Velimirović (1978), which was fully analyzed in 1986 with database technology.

Until 1975, the alpha–beta algorithm was considered to be a heuristic, a well-functioning heuristic, but no more than that. Brudno (1963) was actually the first researcher attempting to align the experimental results to that date. There were two ideas in circulation: (1) the algorithm is extremely sensitive to the order of the moves to be investigated (hence, the order should be: first check moves, then capture moves and thereafter killer moves). Then J.J. Scott in 1969 had the idea (2) to let determine the

order of the moves by the algorithm itself. Obviously, this is the seminal idea of the iterative alpha–beta algorithm. It was later rediscovered independently several times, for example by Jim Gillogly (Carnegie Mellon), Slate and Atkin (North Western University) and the Russian Kaissa team.

In a fundamental article published in 1975, Donald Knuth and Ronald Moore showed that the application of the alpha–beta algorithm always yields the same value as the minimax algorithm. By that result, the algorithm was stripped of its heuristic character. It was a big theoretical step. However, the experimental researchers had always thought so. Of course, it was still extremely important for the experimental researchers to be able to examine the best move first. In short, for time saving reasons, heuristics were still necessary. Hence, thinking in the opponent’s time also became a research issue. Thus, after determining a principal variation, you can anticipate the opponent’s move and prepare for your own move. At the very best, you can answer immediately when the opponent plays the expected move. Later on, every researcher worked out this idea and divided the time, e.g., into two and thus anticipated two possible moves by the opponent. In summary, a variety of strategies emerged. Changes were also made in the representation of board and pieces by applying the bit representations (depending on the word length of the computer). The names of Adelson Velsky and his team (1970) are linked to this. Independently it has been rediscovered by Hans Berliner (1974) and also by Slate and Atkin (1977).

All these issues became important because special computer chess tournaments entered the computer chess world. The most important two were the North American Computer Chess Championship (NACCC) organized by the ACM (from 1970 to 1994) and the World Computer Chess Championship (WCCC), which was initially organized by IFIP (1974–1977) and then by the ICCA/ICGA (1980–today). First once every three years, thereafter with the intention to organize a tournament every year.

Furthermore, from 1980 to 2001 there was a World Microcomputer Chess Championship (WMCC) and since 2010 a World Chess Software Championship (WCSC) has been organized. For the results we refer to the link of the WCCC. Since 2010 there is also an Unofficial WCCC which is organized under the name of Top Chess Engine Championship (TCEC).

For the Russian Kaissa team, winning the first World Championship for computers in Stockholm in 1974 was a great underlining of their activities in this area. The most important contributions for this success came from Mikhail Donskoy and Vladimir Arlazarov. The second championship in Toronto 1977 was won by the Chess 4.6 programmed by David Slate and Larry Atkin. At the end of that championship, the International Computer Chess Association was established under the chairmanship of Benjamin Mittman, who also became Editor-in-Chief of the ICCA Newsletter. Ben Mittman sadly passed away in 2018.

A year after the first World Championship, the first Advances in Computer Chess Conference (ACC) was held in London in 1975. This was the beginning of a very fruitful exchange of ideas about computer chess techniques. Initially they were organized with an interval of three years; later this was every two years with in between the Conferences on Computers and Games (CG) series. For a complete list of books and proceedings we refer to the abundant literature.

The conference meetings significantly accelerated the exchange of information and increased integration of AI techniques and computer chess applications. Especially the work by Arthur Samuel found its way. That is how the work by J.R. Quinlan (1979) “Discovering rules by induction from large collections of examples” published in D. Michie (Ed.), *Expert systems in the microelectronic age*, Edinburgh University Press, received many positive reactions from researchers who wanted to understand (human intelligible) rules derived from end-game databases (see Ströhlein 1970).

1980–1990

In 1980 the knowledge explosion became visible through its recognition by the general AI researchers. Moreover, we saw a further growth of computer chess activities and a large number of publications of new ideas. Many of these ideas found their way into the computer chess programs. In this decade, the emphasis was on parallelism. The top researcher from that time has been without any doubt Kenneth Lane Thompson (born 1943). In the period 1969–1985 he enriched the world with the development of and contributions to the operating system Unix (together with Dennis Ritchie). He designed the language B and was involved in the development of the language C. Together Joe Condon he built the chess system Belle (it was not just a program but a technological entity). Belle was the first program that officially received the US Chess Master title in 1983 and was awarded the first Intermediate Fredkin Prize of \$ 5,000. Thompson also generated many new 5-man end-game databases. Belle gave him the world title in Linz 1980. In 1983, together with Dennis Ritchie, he received the Turing Award for their above-mentioned contributions during the World Computer Chess Championship in New York and fifteen years later the National Medal of Technology from Bill Clinton in 1998. It was a big surprise that he did not prolongate Belle's world title in 1983. Belle was succeeded by Cray Blitz. This suddenly made the hidden connection clear to everyone: Computer Chess is a combination of Hardware and Software. Robert Hyatt, Harry Nelson and Albert Gower had fast transposition tables and even faster multiprocessors. The Cray Supercomputer was born and put into use.

The development of ideas then continued at a furious pace. One idea had not yet been published or the other had already appeared and published. In 1980 Judea Pearl published his Scout algorithm and that was a hit, because good ideas are often the reason for even better ideas. In 1983, as Editor-in-Chief, I converted the ICCA Newsletter into the ICCA Journal. The reception was fantastic. In the first issue I received a contribution from Botvinnik. This was followed by Jonathan Schaeffer's History Heuristic. For the second issue Alexander Reinefeld sent me the article "An Improvement to the Scout Tree Search Algorithm". He called it Negascout. It became a success in many chess programs. In 1987 Don Beal introduced Null Move Quiescence Search (NMQS) to work as efficiently as possible in the ends of the search tree.

From an organizational point of view, much has happened, also in the Netherlands. On October 18, 1980, the Computer Chess Association Netherlands (CSVN) was established. There were already 622 members immediately. The CSVN would from then on hold a Dutch Computer Chess Championship every year. Internationally, it was announced that the ICCA would organize a World Microcomputer Chess Championship every year. That happened from 1980 until 2001.

In the meantime, the Delft University of Technology had also become fascinated by computer chess. This was among others due to my own research: PhD in 1983 on thesis Computer Chess, Chess World and Artificial Intelligence (supervisors H.J.M. Lombaers, A.D. de Groot and S.J. Doorman). Moreover, the study association Christiaan Huygens considered computer chess a beautiful subject for its fifth lustrum in 1982. A public debate in 1981 on the back of the national daily paper NRC-Handelsblad between Grandmaster Donner and Van den Herik led on 5 March 1982 to the Man-Machine meeting Donner-Belle. It was a fascinating meeting not in the least so because of the press releases. Donner won in 56 moves, but the trend was set. Donner compared the gameplay of Belle with that of his niece of four and van den Herik saw it as a step forward in the direction of a match with the world champion. In addition to the Pion Program (Derksen and Huisman), the Delft research team (Dekker, Nakad, van den Herik) also focused on creating special five- and six-man end-game databases, such as two Knights against a Pawn and the Timman-Velimirović end-game. (Rook against Bishop and on both sides an a-Pawn.) Harry Nefkens developed the opening library for the program Pion and chose as one of the first researchers to store positions instead of moves and variations. In

this way, he could add many grandmaster games to the opening library. It inspired Dap Hartmann to build the so-called Dap Tap, a mechanism for recognizing patterns that can support the search process. This technique was later (in the 1990s) adopted in Deep Blue. From a psychology point of view, De Groot (1986) published “Intuition in Chess”. His definition of intuition was as follows: “Intuition is having judgments (or making decisions) in a manner that cannot be made explicit”. He thought that intuition was an essential, intrinsic part of being a chess grandmaster/world champion. According to him, it was thus *impossible to program* intuition. Consequently, a computer could never beat the world champion.

At the same time, there was a great deal of activity at Carnegie-Mellon University. Since 1969 Hans Berliner (1929–2017), the fifth world champion correspondence chess (1965–1968), worked on computer chess topics. He obtained his doctorate in 1974 under supervision of Allen Newell on “Chess as problem solving: the development of a tactics analyzer”. Still, after the work had been published, he did not have really sufficient grip on the evaluation function and decided to apply his new ideas to Backgammon. The result was that his program BKG9 (based on fuzzy logic) in July 1979 beat world champion Luigi Villa by 7-1. It was the first time that a world champion lost in any sport of a computer program. Then Berliner developed the B*-algorithm. These and other ideas formed the basis for HiTech. With his team (including Carl Ebeling and Murray Campbell) he worked day and night on the new program, which was based on multiprocessing, special circuits and knowledge expansions through transitions.

An important contribution at the time was also “Using time wisely” by Bob Hyatt (1984). By this publication he gave the reader a good impression of how time should be divided over the various tasks that take place within a program. Another important contribution was Hash Tables. They were very noteworthy in Cray Blitz (Nelson, 1985) and in the development of new Search Tables (Warnock and Wendroff, 1986). In addition, new search techniques were developed such as Parallel alpha–beta search (Newborn 1982, Bal and Van Renesse 1989) and Conspiracy Number Search (Mc Allester 1987).

After two world championships of Cray Blitz (1983 and 1986), Deep Thought (like HiTech from Carnegie Mellon) won the world title in 1989. The battle between the two Carnegie-Mellon teams was breathtaking, especially as Murray Campbell worked for both teams. He was the *trait d’union* that carried all the secrets. In the final battle he sat with Hans Berliner behind the pieces of HiTech opposite to Feng-hsiung Hsu and Thomas Anantharaman, who played the moves for Deep Thought.

Still, it is good to stop here at this internal Carnegie-Mellon competition. Hans did well, he was a learned high performer. But Feng-hsiung Hsu was a rising star who wanted to break through. He knew everything better and that was true. He was intelligent, designed Chiptest in 1985 (called Deep Thought in 1988) and gathered smart people around him. Hans raced for the Fredkin Intermediate Prize of \$ 10,000. As a token of recognition, we mention that HiTech beat Grandmaster Denker (who had played a draw against Botvinnik in Groningen in 1946) in 1989, but a year earlier Deep Thought had won a serious game from the Danish Grandmaster Bent Larsen in a strong grandmaster tournament. It was the first time that a world-class grandmaster was defeated by a program. That is why in 1988 the second Intermediate Fredkin Prize was awarded to Deep Thought.

1990–2000

In the last decade of the last century, all of the above activities turned out to be actual sources of inspiration for two Dutch programmers, who were not academic researchers. They were professionals



Fig. 10. Ken Thompson, Claude Shannon and David Slate at the 6th World Computer Chess Championship in Edmonton (Alberta), 1989. Source: Monroe Newton.

of the highest level, of world class level. In Vancouver 1991, Ed Schröder won the WMCC with the program Gideon, albeit shared with Mephisto. In Madrid 1992, Ed Schröder also won the World Championship (WCCC) in all categories with the program The ChessMachine Schröder. It was a microcomputer, but with a formidable strength. Schröder had already been in the spotlights in Cologne 1986 with the program Rebel by being a contender for the title up to the last round. Then, in 1992 the time had come. The victory earned him an incredible amount of fame and recognition. It was the begin of the Dutch story.

After their 1989 victory, the Deep Thought team had set as its new goal developing a program that would beat the human world champion. For that milestone IBM had offered its support. The new team was called Deep Blue. They were very eager and dedicated to beat the world champion. Therefore, Deep Blue withdrew from the normal leagues, such as the 1992 WCCC. Consequently, after six years the team showed up at the 1995 WCCC in Hong Kong. The tournament was one of the strongest held so far. For instance, from MIT, Star Socrates participated. It was a program by Don Dailey, Christopher Joerg and Charles Leiserson, with technical chess support from Larry Kaufman. They eventually became second. To make a long story short, in the last round the Fritz program of the other Dutch top programmer Frans Morsch (then 3 out of 4) played against Deep Blue (then 3.5 out of 4). In a very exciting game, Fritz won convincingly, without giving Deep Blue a single chance. So, Fritz was World Champion. Big applause and many compliments to Frans Morsch were given. Here, Frans saw many years of incredible efforts resulting in a place in the hall of fame of computer chess. The CSVN was proud to have two world champions in one decade. Star Socrates was second and Deep Blue third. Hence, IBM's main question was: What are we going to do now?

After a long and intensive discussion, the company still decided to reveal its plans at the end of the tournament. So, there was a press conference to announce that IBM was going to try with Deep Blue to challenge World Champion Kasparov in 1996 with an improved version of the current program. That was the introduction to a new phase in the history of computer chess (Zuse's dream had come very close now).

Scientifically, many things happened at that time. Although IBM kept its claims as secretly as possible, we know that the Dap Tap was further professionalized, many grandmaster games were incorporated into the Deep Blue program and much emphasis was placed on fast hardware ("brute force" in the popular vernacular). Fast, fast, fast was the adage that applied especially to the development of RS 6000 chips. Parallelism, distributed systems and acceleration of the hardware were the technological key issues. In addition, ordinary research took place at universities and institutes. But there was more because there were various chess computer companies that also wanted to keep their products advanced (we mention Fidelity, Chess Challenger, Mephisto, Novag, Saitek, and Tasc).

In 1994 Bob Hyatt launched the open source program Crafty. Until then there was only a reasonably strong open source chess program called GNU Chess, it was a loot from the well-known GNU-program set. Now there was a program that could serve as an opponent, practice partner, and even as a participant in the world championships (a version improved by an Indonesian team played in Jakarta (1996) for the World Championship with the agreement of the participants). Hyatt is a pioneer who has contributed much to the development of computer chess (see earlier). Here we mention: an evaluation cache, transposition tables, bit-board data structures, and rotated bitmaps.

In this decade, there was extensive scientific research in Maastricht into transposition tables (Breuker, Uiterwijk, van den Herik), opponent modeling (Iida, van den Herik, Herschberg, Donkers), proof number search (Allis, van der Meulen and van den Herik), speculative play (Uiterwijk, van den Herik)) and solving 8×8 Domineering (Breuker) as well as Fonorama (Donkers). Afterwards, the attention for games shifted in Maastricht to Lines of Action (Winands) and Go (Van der Werf). In 1993 Frans Morsch suggested the recursive use of the null move. He shared his thoughts with Chrilly Donniger who worked it out in the breakthrough article (1993) Null Move and Deep Search. It was a signal for much and thorough research to refine this mechanism down to the smallest detail.

In Edmonton, Tony Marsland was still very active. But the torch was slowly but surely taken over by Jonathan Schaeffer, who focused at the University of Alberta on building teams that were involved in various techniques and games. Eight of such teams became known for their research in the field of (1) MtD (f) (Plaat, Schaeffer, de Bruin and Pijls), (2) Checkers (Schaeffer, Björnsson, Burch, Kishimoto, Müller, Lake, Lu and Sutphen; the game has been solved, the outcome is a draw), (3) Partial Information Endgame Databases (Björnsson, Schaeffer and Sturtevant), (4) Othello (Michael Buro and Mark Brockington), (5) Go (led by Martin Müller), (6) Poker (led by Darse Billings, Mark Brockington), (7) Hex (led by Ryan Hayward) and (8) Sokoban (Andreas Junghans, Tony Marsland and Jonathan Schaeffer).

All this research was co-stimulated by David Levy's initiative to organize a Computer Olympiad in London with Don Beal in 1989. From each successive computer Olympiad, the researchers learned two things: (1) computer chess was still the incentive to explore other games, and (2) perhaps the most difficult game in the world, Go, was a worthy research successor of the noble chess game. It was clear that there was a shift of attention. This led the ICCA to officially change its name from ICCA to ICGA (International Computer Games Association) in 2002.

The biggest revolution in the decade 1990–2000 was the Kasparov's defeat. It was preceded by an exercise in waiting and improving. In the first match Kasparov–Deep Blue (1996), Deep Blue played

on an RS/6000 SP machine with 36 processors. The first game was played on 11 February 1996 and seemed destined to usher in a new era, as Deep Blue won convincingly. However, that turned out not to be the case, as Kasparov decided the match with 4–2 in his favor. A year later (1997) there was a second match. The hardware and the structure of the program remained virtually the same. There was more training with and by grandmasters (Larry Christiansen and Michael Rohde), although we know that three other grandmasters had also played test games for the first match.



Fig. 11. 1997 May 11th, AI achieves its long-standing goal. DEEPBLUE (IBM) wins from Kasparov by $3\frac{1}{2}-2\frac{1}{2}$.

The second match started with a huge bang for the IBM team, Kasparov won convincingly. The Deep Blue team licked their wounds and took new courage. In the second game, Deep Blue showed strong opening play (closed Spanish) and put Kasparov in a difficult position. If Deep Blue had chosen a pawn win then the game would have become very complicated and tactically. Deep Blue, however, played the beautiful move 37. Le4, which decided the game in a higher sense. It is true that Deep Blue later in the game played a weak move with 44. Kf1, but in the end the program won and that made the score 1-1. Thereafter, it became 2.5-2.5. The match would thus be decided in the sixth game. Kasparov played too challenging in the opening and ended up in a lost position via a piece sacrifice. A deserved victory by Deep Blue. Zuse's dream was fulfilled. (Please, note that the photo above was taken earlier in the match.) Sales of the IBM computers rose, but Kasparov played his last card. He explained at press conferences, in publications and at ordinary conferences that grand master (i.e., human) aid was the basis of 37. Le4. Similarly, 44. Kf1 was a human error. Both moves could never have been played by an actually intelligent thinking computer. The first move was too good and the second too bad. Over the years Kasparov has understood that computers can do strange things. He eventually admitted the loss of 1997, although he had previously demanded a rematch. That was in the rules of the first match but not in the rules of second match. For their performance, the Deep Blue team received the Fredkin Prize of \$ 100,000. IBM did not allow Deep Blue to play anymore because of publicity considerations. It was the end of a wonderful journey. Nevertheless, the developments continued.

2000–2010

The big question of the first decade in the new century was: when can microcomputers beat the human world champion? A second related question was: when does a computer (large or micro) pass the ELO barrier of 3000 points? During the WCCC 1999, the microcomputer Shredder was able to defeat the CilkChess program (playing on a supercomputer) in a decision-making game for the world title. Hardware and speed are indeed of utmost importance, but smart algorithms also play their part. In 2001 the last WMCC was played. In 2003, Kasparov played a match of six games against frequent WMCC winner Deep Junior. It ended in 3-3. In 2005 Hydra played against Michael Adams (once number 4 in the world ranking): the result was 5.5-0.5. Deep Fritz then played a match of six games against former world champion Kramnik. Deep Fritz won by 4-2. This answered the first question. A microcomputer was also stronger than the world champion. By that victory, a completely new situation came into effect. From now on, the public had to sit in a different room at the human world championships. The obvious reason was that the two strongest human players in the world did not know what the best move would be in the current position, while the audience could see on their pocket computer what the best move was according to a computer program.

In this decade, the globalization of the WCCC and the Olympiad began. There were meetings in Israel (2004), Taiwan (2005, Olympiad), Reykjavik (2005, WCCC), Turin (2006), Amsterdam (2007) and Beijing (2008). The playing strength increased with the year and the developments continued in parallel, i.e., more and more powerful supercomputers and ever smarter microcomputers. Thus, the ICGA (so called in 2002) decided in Pamplona (2009) to organize a World Computer Software Championship (WCSC), with the characteristic that the programs had to play on a uniform platform. The first official WCSC took place in Kanazawa 2010 and was won by Shredder. For other tournaments we refer to the WCSC link.

Thanks to the combination of WCCC, Olympiad and conference, there has always been a large exchange of ideas among the researchers. Through this interaction it happened that studying the game of Go further advanced in the interest of researchers and organizers. Initially, the development happened slowly because the alpha–beta minimax method was not very suitable for Go (it should be noted that the difficulty was not in the branching factor, the large search space or the complexity as many people thought, but in making the correct evaluation (see also the problems that Berliner had in 1974–1979)). The French researcher Bruno Bouzy had a solution: avoid all evaluations and look at the end of the variant and do so for a large (limited) collection of moves chosen randomly.

The implementation of the idea was done in the program Indigo. In the Olympiads where Indigo played, it finished with this idea always at positions in the middle of the pack. For Bruno Bouzy the idea still gave him much recognition by publications in the ICGA Journal and by his Habilitation thesis (2004). Additionally, after some time the idea turned out to be the launch of a whole new research area. At the CG conference (Turin 2006), Rémy Coulom introduced the idea to use Monte Carlo search techniques to traverse the search space. Everyone in the hall in Turin felt: this is the direction we should go for Go. Two important contributors to the further development of these techniques are (1) Kocsis and Szepesvári (2005) who designed the UCT formula and (2) Chaslot, Winands, Uiterwijk, Van den Herik and Bouzy (2008), who designed the Monte Carlo Tree Search technique. By Monte Carlo Tree Search (MCTS), Coulom's ideas were accommodated in a dedicated framework which enabled further development of specialties. In the fiercely evolving Go world, computer chess went its own way: to perfection of the playing strength and to perfection of efficient algorithms.

In 1999 Shredder had broken the hegemony of the supercomputer. From that moment on we see that the microcomputers Shredder and Deep Junior dominate the WCCC. This line of development

changed in 2005. Two programs came from nowhere and governed the WCCC in Reykjavik with a strong hand. Zappa (Anthony Cozzie) became world champion with 10.5 out of 11 and Fruit (with Fabien Letouzey) was second with 8.5 out of 11. It seemed like a new time would arrive, but things went differently. Zappa disappeared from the arena after 2007 and Fruit had publicly made his program available to play against the interested public. It encouraged people to obtain ideas for their own inspiration and research work, but not to copy modules from them. Letouzey himself started to do something fully different (see his Wiki page). In later years he returned to our community and played in the Computer Olympiad with the Draughts program Scan. He won the tournament convincingly. After the interbellum, Junior (without Deep) and Shredder resumed their former positions and in Turin (2006) Junior became World Champion. Shredder and Rajlich (later Rybka) shared the second place in Turin. In Amsterdam 2007, Rybka showed itself the strongest. That remained so in Beijing 2008, Pamplona 2009 and Kanazawa 2010.

A similar trend was observed in the Go world. There, the MoGo program attracted all attention. Moreover, the power of MCTS increased continuously. The main reason was that the proper adding of knowledge to the search process was found; it was still not perfect, but sufficient to steer the random effect to efficient and effective results. According to some Go watchers, the coming decade would be the final piece of the development, according to others it was certainly two decades further on or even more. The world of games followed the developments with excitement and kept in mind the relation between Go and chess.

2010–2016

Chess thus remained in the spotlights and computer chess even more. Although the human world champion was defeated, the competition between humans and computers continued in comparing their playing strengths (Elo ratings). Many researchers interested in technology made their “own” programs from other programs. It led to super strong programs such as Houdini and Stockfish. However, the ICGA was of the opinion that these programs could not participate in the World Championship, because they should be original programs, that is, the participating program had to be designed and implemented by the programmer (or a team of programmers and developers). The rules for admission became stricter and occasionally a program was refused or even disqualified due to plagiarism. It was a complicated clash of science and technology, of amateurs and professionals, of inventors and thieves. The rules became more complex and the pressure from the professional enthusiasts (no matter where the algorithms originate) grew.

In 2010 the Top Chess Engine Championship (TCEC) was organized for the first time. It was an online tournament. The players could play from home. They considered the tournament as the unofficial WCCC. Just like Crafty in the earlier times, Houdini and later also Stockfish were publicly available. For more information about TCEC we refer to the link on TCEC.

On 23 January 2011, the Rybka-Fruit case came to a public clash. Fabian Letouzey wrote an open letter to the ICGA board with the announcement that the programmer of Rybka (Vasik Rajlich) had plagiarized his work. This led to in-depth research and a technical committee of the highest caliber (including Ken Thompson). The committee did its work very carefully and came to the conclusion that the complaint by Letouzey had to be assigned (there were even pieces of code found with original comments from Letouzey). This meant that Rybka’s titles and all podium places in the WCCCs and the WCSCs were declared invalid. New world champions and new runner ups were appointed. The computer chess world was turned upside down. The developments, however, went quickly. The

computer chess world resumed its research activities. Johannes Zwanzger joined the top after years of hard work with the program Jonny. Shredder maintained and Junior disappeared from the scene. The newly coming program was Komodo. It was a program from Don Dailey's MIT stable, which was maintained and renewed by Mark Leffler. Moreover, new chess ideas had been provided by Larry Kaufman. They dominated the years 2015–2017, witness the results of the WCCCs and WCSCs in Leiden (2015, 2016, 2017).

On January 29, 2016 I was privileged to give my valedictory address at the Tilburg University under the title "Intuition is programmable". As a tribute to my supervisor Adrian de Groot, I focused on intuition and computer chess and the possibility of programming intuition in general (mental, physical, emotional, and environmental intuition). My guideline for computer chess was Donald Michie, who defined intuition as follows: "Intuition is simply a name for rule-based behavior where the rules are not accessible for consciousness." Here, I note that I agree with De Groot that (1) intuition is not irrational and (2) intuition is not infallible (and generates a reasonable degree of correctness). However, I had a dispute with him on the point that (3) intuition cannot be programmed. The current results show that the current top programs are even better than being of world championship caliber. Two paths to a conclusion are possible. Path 1 reads: "Intuition is not an essential (intrinsic) part of chess, because a computer program without intuition was able to defeat the world champion." Path 2 reads: "Intuition is a form of knowledge that we are not able to interpret correctly. Intuition is implicit in the chess knowledge that was incorporated into the program."

Meanwhile the chess world was helped by the Go world (see intermezzo below). The company DeepMind was able to create a neural network with the help of deep learning, which, by playing against itself, was able to chart the details of the game in an unprecedented way.

2015–2017 INTERMEZZO

This intermezzo of the then current computer chess research is on machine learning and its application to Go. That is why we present only in brief the results achieved with machine learning and deep learning in the Go world. An asterisk (*) means: increased playing strength compared to the previous version.

- In the Autumn of 2015, AlphaGo beats the European champion Fan Hui 5-0
- In March 2016, AlphaGo * beats the multiple world champion Lee Sedol 4-1
- In May 2017, AlphaGo ** beats world champion Ke Jie 3-0
- In the Autumn of 2017, AlphaGo Zero beats the AlphaGo ** program by 100-0

AUTUMN 2017 APOTHEOSIS

Here we return to chess. In the Autumn of 2017 the deep learning technique has been applied to chess. AlphaZero is a program that only knows the movements of the pieces and the rules of the game (such as castling, *en passant* capturing and the 50-move rule). It learns by deep learning and by self-play. Many details are described by David Silver et al. (2016, 2017). After sufficient training, it was decided by the Deep Mind team that the program was sufficiently strong to play with the very strongest chess program of the world. As opponent, the Deep Mind team chose to invite the program Stockfish (on the basis of the TCEC competition). The estimated Elo rating of Stockfish in 2017 was above 3400. So, AlphaZero played Stockfish in a match of 100 matches. It defeated Stockfish convincingly by 28-0

(and 72 draws), in chess terminology the victory was by 64-36. This is the end of story. I leave the commentary to the reader. There is one more research challenge: to determine the game theoretical value of chess.

WORD OF THANKS

I would like to thank all the people (and those are incredible many) with whom I have had the pleasure to work in the past forty years. Without them, this article would not have been possible. The article itself is an expanded result of my keynote lecture at the 10th International Conference on Computers and Games (CG 2018) in the National Taipei University in Taipei, Taiwan, on July 9, 2018, organized by I-Chen Wu and his team. A brief Dutch version appeared after many encouragements by Robbert Fokkink, Editor-in-Chief of the *Nieuw Archief van de Wiskunde*. Really, without his tenacity this historical overview would not have been written. Finally, I thank the editorial team of the *ICGA Journal* and IOS Press who supported me, together with my personal second Ron de Goeij.