In 1997, IBM Deep Blue beat the current World Chess Champion. Deep Blue was the result of an iterative research effort that had started in the 80s. Lessons learnt from Deep Blue's predecessors matches led to the necessary improvements in Deep Blue described below.

Massively parallel system

Deep Blue's 3-layer parallel system enables a large search capacity. 1 master node searches top levels of the tree, while the other 29 nodes examine a few more levels in parallel. Each processor controls 16 chess chips to search the last levels. These chess chips have a speed of 2-2.5 million positions per second. Deep Blue uses Principal Variation search: using information from previous depth search in iterative deepening, it guesses which move is most likely to be best and searches it fully first. For the other moves, it starts with a quick search. Deep Blue makes the most out of its parallel architecture and customizes the parallelization to each node type.

Combining Hardware and Software search

Deep Blue combines the speed of hardware search and the complexity of software search.
The chess chip search is controlled by a processor, and, because of its simplicity, it is usually limited to shallow depth. Hardware search carries out a fixed-depth null-window search, which includes a quiescence search and some simple extension heuristics. Null-window alpha-beta search speeds up the process to find cut-offs but does not return an exact value. Therefore, when an exact score is needed, the controlling processor has to carry out a binary search. Another limitation on the hardware search is the lack of transposition tables, available to the software search, although chess chips optionally support them.
The evaluation function is implemented in hardware so the time to execute the evaluation function is fixed. The drawback is that new features cannot be added to hardware evaluation and software patches can be problematic.

Non-uniform search

Deep Blue's large search capacity is leveraged by a highly non-uniform search, which was implemented after realizing that human experts can clearly search beyond the depth reachable by uniform search.
On top of quiescent search and negascout, Deep Blue introduced a "dual credit with delayed extensions" technique in its software search: Forcing/forced pairs of moves (ffp) play a critical role in chess and the search should ideally be extended by 2 plies when encountering a ffp. However, that is not achievable since the number of searches would be too high. Deep Blue smartly identifies which moves are worth extending. It keeps track of an "extension credit" for each player and path. For each ffp along the path, a player accumulates a credit (the credit depends on factors such as how forcing the move is). If, at the end of the path, a player has accumulated more credit than the threshold, the search is extended by a number of plys that depends on the accumulated credit.

Evaluation function

The evaluation function sums the values of over 8,000 features. Some features can be as simple as assigning a value to a piece on a square and others are very complex. Static features are set once at the beginning of a search. Dynamic values are initialized at the beginning of a search and they are adjusted at evaluation time based on the board situation. A lot of time was spent on creating features and tuning weights.

Other improvements

On top of an opening book created manually by experts that references about 4,000 positions, Deep Blue has access to an “extended” opening book that helps in the absence of opening book information: Having access to a 700,000 game database, Deep Blue calculates the expert consensus move using a non-linear combination of features such as “number of times the move was played” and “strength of the player who made the move”. It then does its regular search but offsets the alpha-beta window by the calculated expert score.

Deep Blue also had access to endgame databases for all positions containing less than 5 pieces and some selected 6-pieces positions.

Deep Blue’s victory cannot be attributed to a single feature. All the techniques described above played a significant role.

Reference:

Murray Campbell, A.Joseph Hoane Jr., Feng-hsiung Hsu, Deep Blue, *Artificial Intelligence 134 (2002) 57-83*