Research Paper Review: Deep Blue by the IBM Watson Team

By: Miriam Farrington

The goal of this paper is to give an overview of IBM's Chess Machine Deep Blue II (which defeated Garry Kasparov in 1997) and how it compared to the previous version, Deep Blue I (which lost to Garry Kasparov in 1996) and earlier versions of chess playing machines (Chip Test, Deep Thought, Deep Thought 2) leading up to Deep Blue.

First, the paper looks at an overview of the hardware components that made up early chess machines developed in the 1980's (Chip Test, Deep Thought), and later improvements in the form of Deep Thought 2. Some of the advantages are covered, such as enhanced hardware and software capabilities as well as multiprocessing functionality not present in the earlier iterations, some of which were used in developing Deep Blue I.

Next, the paper looks more closely at Deep Blue, which utilizes a massive amount of parallel processing capacity to perform game tree searches with certain processor designated as master and the remainder as workers to search the tree and report back to the main node. As noted in the paper, the chess chips in Deep Blue are able to search about 2 to 2.5 million chess positions per second using this approach, although the speed varied depending on the complexity of the move being evaluated. Some challenges to the Deep Blue design are covered, including the depth to which the searches were performed (orders of magnitude deeper than prior attempts), along with a hybrid hardware/software design where much of the evaluation functionality was hard-coded into the chip itself, making this portion of the evaluation logic difficult to configure.

The remainder of the paper takes a deeper dive into the components of Deep Blue II, including the makeup of the chess chip and its sear, characteristics of the software selective search functionality (much of which was similar to alpha-beta search) and finally the evaluation function which takes a large number of feature inputs and some of the benefits and drawbacks of it (although one example of the 'rook on file' table is given, the full complexity of the evaluation function is not covered in this paper). Finally, the paper covers some notes about the 'opening book' of Deep Blue, which is basically a stored set of optimal opening moves, which were basically hand-selected with input from several chess grandmasters.

Interesting to note is when the paper mentions "There were a number of possible repertoires to choose from, and the choice would be made on the basis of the match situation and the previous experience playing with the same color." This hints at possibly some type of machine learning model implemented to keep track of training data from prior chess matches. Additionally, the paper briefly covers some of the miscellaneous functionality implemented in the extended book (a way for Deep Blue to keep track of, for example, moves frequently used by Grandmasters) and the endgame databases, which keep track of optimal moves when only a few pieces remain in the board.

This paper gives a comprehensive high-level introduction to the hardware and software components of Deep Blue in an understandable way, but it is clear that many of the single components mentioned here could be considered the subject of much longer and more extensive research for those wishing to take a deeper dive.