

## LIFE WITH CHESSBASE

*John Nunn*

London, England

One of the most frequently discussed questions concerning chess is whether it is an art, a sport or a science. Different types of chess activity emphasise each of these three facets of the game. Postal play emphasises the scientific aspect, chess composition the artistic and over-the-board play the sporting. Just as with any other sport, most of the time expended by the professional player is in training rather than actually competing, an extreme case being the 100-metre runner. Every grandmaster has his personal training methods, but in general these fall into two main categories.

### 1. PREPARATION FOR TOURNAMENTS

Firstly there is preparation for specific opponents. A certain amount can be done in advance of a tournament, but some of the work is best done on the morning before the game, when you know your colour, the tournament situation and how you are feeling.

However, the growth of double-round events has tilted the balance in favour of pre-tournament preparation, since you know in advance that you will have to play both White and Black against all the other competitors.

The second type of preparation, which is far more important, is general opening work. No matter how carefully you prepare for individual opponents, there is always a chance that they will do something totally unexpected, and then you have to fall back on your general opening knowledge. In these days of Informator, Tournament Chess and yes, dare I say it, ChessBase, the dissemination of new knowledge occurs at a frightening speed, so it is essential to be up-to-date and prepared for each new idea which arises in grandmaster play.

### 2. CAN COMPUTERS BE A SOLUTION?

For many years players have wondered how to make use of computers in their training. There is no substitute for the grandmasterly brain (yet) when it comes to analytical work, but in practice one finds that a good proportion of one's time is spent searching through books and magazines for games which one can vaguely remember, but alas not well enough to find quickly. Since computers are excellent information manipulators, there has always been the theoretical possibility of using them to do the boring and time-consuming job of maintaining a database of games. The obstacle to putting this into practice has been the lack of suitable software. Matthias Wullenweber, a chess-playing German programmer, decided to do the job and in time ChessBase was born. Later he teamed up with Frederic Friedel, who added both his experience with chess and computers and the essential business know-how to make ChessBase commercially viable.

I should point out straightaway that I am not writing a review of ChessBase; this would in any case be impossible because I am a contributor to *ChessBase Magazine* and so have financial links with ChessBase. I am also no computer expert, my only claim to fame being a remarkable ability to make the most well-behaved pieces of software crash without warning.

If anything can survive the first hour when I am at the keyboard, then it is doing well!

I have noticed that software crashes quite often when one first uses it, but gradually becomes more reliable. I am convinced this is due to an almost subconscious learning process which steers one away from dangerous actions. When a crash occurs on the Atari, you are treated to the appearance of little pictures of time bombs on the screen. These are apparently due to what are called "exceptions"; I have a more colourful name for them.

### 3. EXPERIENCE WITH CHESSBASE

My aim is simply to describe how a GM uses ChessBase and my experiences with it. I use ChessBase on an Atari 1040 ST computer. For those who are not familiar with this machine, it has a 68000 processor, 1 MB of RAM and a monochrome screen with a resolution of 640x400 pixels (it will also drive a colour monitor, but ChessBase only works with the monochrome screen). I have added a 20 MB hard disc, necessary now that my main games file is 2 MB long. ChessBase also comes in an MS-DOS version, but my comments apply only to the Atari. ChessBase 2, an improved and upgraded version with many extra facilities, has just been released. I will indicate where my comments do not apply to the new version.

One of the pleasures of running ChessBase on the Atari is that the screen is sufficiently sharp and clear that I can use it for many hours without fatigue. I have experienced eyestrain and headaches when using some computers, perhaps because of slight screen flickering, but I have not had any trouble with the Atari. The main ChessBase screen has a large chess diagram showing the current position with the last couple of moves appearing to the right of the diagram. Commands are entered via drop-down menus, or, if the user prefers, the most common commands have keyboard alternatives. Moves are most easily entered by moving the pieces on the diagram with the mouse, but it is also possible to type them in using algebraic notation.

### 4. IGNORANT OF CHESS

ChessBase does not play chess. It understands what a legal move is but that is the sum of its chess knowledge (ChessBase 2 has limited chess intelligence, but this is solely to make entering moves easier, e.g., you may click on a square several times, and ChessBase 2 will try to execute the most sensible series of captures on that square). All its functions relate to the manipulation of games databases and their associated opening classification systems. A games database consists of two files, distinguished by the suffixes .GMF and .GMI. The .GMF file contains the games (stored on a 1-byte-per-ply basis), the names of player, tournament, year and result. The .GMI file consists of a series of 32-bit integers, one for each game, pointing to the start of the game in the .GMF file. An openings key also consists of two files, a .KEY file and a .POS file. The key file contains the tree-like structure of the openings key and the list of the games associated with each branch of the tree; the .POS file contains the chessboard positions which are linked to each branch of the key.

### 5. THE INS AND OUTS OF OPENINGS KEYS

To take a trivial example, one may devise a simple openings key with four categories, French defence, other King's pawn openings, Queen's pawn openings, and everything else. The position after 1. e4 e6 2. d4 d5 would be linked to the first branch, that after 1. e4 to the second, that after 1. d4 to the third and the starting position would be linked to the fourth branch. ChessBase classifies games by playing them over backwards from their final position until it encounters a position which matches one in the .POS file. Then that game is added to the list associated with the branch of the tree to which that position is linked. This method offers the best chance of catching transpositions because ChessBase always takes the match farthest from the starting position. In the above trivial example, a game starting 1. d4 e6 2. e4 d5 would be given the French defence classification, because this match would occur before the faulty match to the 1. d4 position.

Even here one can see that constructing a detailed openings key is a massive task, requiring not only knowledge of the positions involved but also experience of the move-orders which most often lead to those positions in practice. Individual users would hardly want to make one themselves, even though ChessBase gives the user full facilities to do this if anyone has the time! More practically, one may modify supplied keys or construct specialised keys for openings in which one has particular interest. I see this last facility as being of great importance for serious users. The ECO openings key supplied with the original ChessBase

had a number of flaws, but improved versions are on the way which will have a higher success rate at automatic openings classification. In any case ChessBase allows the user to manually override the automatic classification, but if you are adding a large number of new games this is very time-consuming.

## 6. EXTENDING CHESSBASE

Almost all users will want to enter games themselves sometime. Even if one relies almost entirely on the update material supplied in the bi-monthly *ChessBase magazine* disc, which provides about 6000 games extra per year, then users will want to append their own games. I have found that I can average about four minutes per game entered. However there is a trap one can fall into. Because ChessBase will only accept legal moves, there is a tendency to become careless. For example, suppose there is a rook on a8 and you play ... Rc8 and then ... Rd8 a few moves later. If you erroneously enter ... Rb8 instead of ... Rc8 the sequence of moves will still be perfectly legal, so ChessBase cannot pick up the error. Every time I receive a bunch of new games I notice a few which I know are wrong and I am sure that a more detailed examination would turn up further errors. Here the problem is that a manual system would automatically incorporate proof-reading, but with the computer one tends to think that this is no longer necessary.

## 7. SIMPLE SEARCHES ...

ChessBase offers the facility to search for players, year, opening or any combination of these (ChessBase 2 adds a tournament search too). Those games matching the search criteria may be "exported" to a kind of mini-database which may then be perused at leisure. Thus one may search for all games played by Timman in the French Defence in the years 1986-88, and play them over one by one on the screen. This kind of facility is extremely useful for individual preparation, and the openings key is useful for general openings work. When I go to a tournament I take a print-out of all the games in my main database played by my opponents. It is much less tiring to look through a "Ljubojevic-White" print-out than to plough through several issues of *Tournament Chess* searching out these games from amongst thousands of others. This is an important factor if you are preparing on the morning of the game. Taking a 300+ page print-out to a tournament might seem a chore, but if it is printed on lightweight paper it only comes to a couple of pounds, less than one volume of ECO. Even Jon Speelman, a noted computerphobe, admitted that the print-out I made for the Brussels World Cup was useful. A paper print-out is more useful than a portable computer would be; it is reliable, cannot easily be destroyed and is available at an instant's notice.

The search facilities in ChessBase are unfortunately rather simplistic. You may successfully search for Pflieger, but to search for Pflieger with Black is another matter. The reason is that ChessBase only finds an exact string match, with no support for wildcards. So if you search for "-Pflieger", you would find Nunn-Pflieger, but not Nunn-Dr.Pflieger or Nunn-H.Pflieger. ChessBase 2 allows you to search for -\*Pflieger, a vast improvement. I also had a problem when I met the Swiss player Züger. I couldn't generate a ü from my English keyboard and since I couldn't search for Z\*ger I was stuck with either no games at all or all games by players starting with Z. However even with ChessBase 2 you cannot search for more than 2 players at the same time. Since I often prepare for whole tournaments, I wrote my own program to search for up to 18 players in one pass through the database, with wildcards and with all White and Black games neatly separated out. Curiously enough, this program is over 9 times faster than the ChessBase search, even when you are searching for just one player.

I have noticed a peculiar psychological attitude which comes over computer users. They undertake operations which would take hours or days if performed manually, and then they become furious if it takes their computer more than a couple of minutes to do the same thing. Perhaps this is why I have taken the major decision to stop using opening keys with my databases. The reason is that the presence of an openings key slows down every operation enormously. For example, suppose I decide to delete game number 2 in my main database, which has 16,000 games in it. Without the openings key this just means moving a chunk of 2 MB along and re-adjusting 16,000 pointers. With the openings key the program must go down every branch of a convoluted tree and change all the numbers it finds at the end. This makes deleting a game a

very slow operation. It is strange that while the computer database is so much better than a manual card index for many operations, the electronic equivalent of throwing away a card is incredibly slow.

## 8. ON NEW ACQUISITIONS

There are two other problems which often arise. They are both associated with the situation in which you have just acquired a brand new disc with, say, 1000 games on it. You want to add it to your main database, but you realise that you probably already have a lot of these games in your database. How can you identify duplicates? The answer is that you can't, not within ChessBase. Again I wrote a special program to delete duplicate games, but it won't modify an associated openings key. A friend of mine had her 22,000-game database reduced by 3000 games when she used this program, which shows how easy it is to build up unwanted copies.

The other problem is that each branch of the openings key tree can only hold a maximum of 127 games. When you add 1000 games some of these branches overflow. You can note these down and go back to break the problem keys into more detailed subkeys, but then you have the awful problem that some of the games you have added are unclassified, because they were put in after the overflow. You can reclassify a key, but these games are not in a key, so that won't work, or you can reclassify the whole database, taking  $3.5 \times 16000 = 56000$  seconds. All these points led me to abandon the openings key for my main database, excellent though the system is for small collections of games or for specialised openings. The problem here is just one of limited computer power; if the operations were 100 times faster I would still use the key, but it just slows all database manipulations down too much when you have a lot of games spread over the whole range of openings. Instead I make do with my own utility which searches the database for all games starting with any one of a list of specified move-orders. This works well for the openings I play, which tend to have few transpositional possibilities, but would be hopeless for many openings (e.g., the Catalan), which can arise via dozens of move-orders. Despite these reservations I find ChessBase a very valuable tool. The real test of a piece of software is how often you find yourself using it, and I am turning to ChessBase more and more often as my games collection builds up.

## 9. WHERE CHESSWARE MEETS SOFTWARE

On a more general note, it is quite interesting to see how non-computer literate GMs react to ChessBase. Computer manual writers never seem to be able to get the level right. They either adopt a patronising attitude, as if the reader is eight years old, or they assume too much knowledge. They forget that the reader is likely to be an intelligent well-educated adult who simply doesn't know much about computers. The ChessBase manual is very much better than most, but nevertheless assumes that the reader is familiar with the GEM operating system from having used other pieces of Atari software. This may be justified for a C compiler, but many chessplayers have bought their Atari just to run ChessBase. Since you need to be able to use the operating system with ChessBase (e.g., to change filenames), you can't get away with just running the program.

Let me give two examples of the problems software designers might encounter. When one well-known 2600+ GM asked me what the mouse did, I explained that when you moved it the little arrow on the screen followed the movement. After waving the mouse about in the air for a few seconds, he complained that the arrow wasn't moving. The second, rather more serious example, concerns exiting from the program. You have to do this to close the files properly; simply switching the machine off will probably fatally corrupt files. This is a totally alien concept for 99% of the population. You enter a game, you save it, the disc whirrs. Clearly the game is saved on disc. When you have finished entering games you switch the machine off. Result: you lose all your games. I did this a couple of times when I got ChessBase, and I have yet to meet a user who hasn't done it at least once. I suppose this is another case in which the card-index system is superior, since forgetting to close the drawer of a filing cabinet usually doesn't result in the spontaneous disappearance of the files therein. In my view page 1 of any software manual should have a list of all the things you shouldn't do with the software!

I remember when I first got a printer, I spent the best part of a day fiddling with the little DIP-switches inside trying to change the page length. Finally I found a tiny note at the bottom of one page in the middle of a thick manual, which told me that changing the switches had no effect unless you switched the printer off and then on again. I was amazed that anything so completely counter-intuitive would be hidden away like this. Afterwards I tried switching my television off and on when I wanted to change channels, but luckily enough I always got BBC1 ...

In conclusion, I believe that computers have now entered the professional chess scene permanently, and we will see more computer-based aids in the years to come. ChessBase has carved a niche for itself in this market and I would certainly not be without mine now. But I had better sign off now to do some serious chess work; where's that disc with the K+Q v K+R database on it ... ?

## REPORT ON THE 1988 FREDKIN MASTERS OPEN

*Murray Campbell, Feng-hsiung Hsu and Gordon Goetsch*

Carnegie-Mellon University

### 1. INTRODUCTION

At various times in the past decade, the Fredkin Foundation, under the directorship of Dr. Hans Berliner of Carnegie-Mellon University, has sponsored events intended to estimate the strength of the current top chess programs. Along with the support for such tournaments, the Fredkin Foundation is offering a \$10,000 prize for the first program to produce a consistent 2500 performance, and a \$100,000 prize for the first program to surpass the human World Champion. The prize of \$5000 for the first master-level program was awarded to the creators of Belle, Ken Thompson and Joe Condon, in 1983.

The last Fredkin event, the 1985 Fredkin Masters Invitational, had clearly established that Hitech was capable of holding its own against well prepared master-level competition. The 1988 Fredkin Masters Open, held over the weekend of May 28-30, 1988, was organized to pit a number of the current top programs against a similar level of competition. Unfortunately only a few programs were able to obtain the necessary machine time.

### 2. THE PROGRAMS

Four programs participated in the tournament:

- Hitech: Rated 2376 USCF going into this tournament (and having since surpassed 2400); Hitech has the highest established computer rating. Hitech has special purpose hardware that allows search speeds of up to 150,000 nodes/second, and has a number of programmable feature recognizers for evaluating positions. A number of people have worked on Hitech, including Hans Berliner, Carl Ebeling, Gordon Goetsch, Murray Campbell, Larry Slomer, Andy Palay and Andy Gruss. Hitech was developed at Carnegie-Mellon University.
- ChipTest: The current North American computer champion, ChipTest also has special purpose hardware that, running on a Sun 4, can reach about 500,000 nodes/second. ChipTest has a very simple evaluation function, but a rather complex searching algorithm, described in part in Anantharaman *et al.*, (1988). ChipTest was unrated going into the Fredkin tournament. The main authors of ChipTest are Thomas Anantharaman, Feng-hsiung Hsu and Murray Campbell, all at Carnegie-Mellon University.