

Cloning Chess Engines

By David Levy

Introduction

The cloning of chess engines appears to have been steadily on the rise in recent years and is a practice strongly disapproved of by the International Computer Games Association (ICGA). In the world of computer chess cloning not only damages the commercial opportunities for the original programmers, it also steals the kudos of tournament successes. Genuinely achieving a great result in a top level chess tournament requires years of painstaking effort by a highly skilled and highly motivated programmer or team of programmers, yet the creation of a clone steals the glory and public acclaim from its rightful owner. The ICGA would like to see this disgusting practice stopped and those who perpetrate the cloning publicly exposed for what they are. This article is the ICGA's opening shot in that struggle.

We start by considering two aspects of cloning, and presenting links to various Internet postings (by others) on specific allegations, as well as some additional quotations.

The Langer Case

First we consider cases where an entire chess engine has been ripped off, without any attempt being made to change its code. The first such case to come to the attention of the ICGA (which was then called the ICCA), was at the 1989 World Microcomputer Chess Championship in Portoroz, where play took place in the very same hall where, 31 years earlier, the 15-year-old Bobby Fischer qualified for the first time for the Candidates stage of the World Chess Championship. I well remember how, during the first round of the 1989 event, I was impressed with the play of the program Quickstep, entered by a German programmer, Herr Langer. I became less impressed shortly afterwards when Richard Lang, then the programmer of the Mephisto range of chess computers, revealed that the user interface of Quickstep was identical to that of his own program. The matter was investigated on the spot by interrogating Herr Langer who at first denied that he had copied the Mephisto Almeria code. But when Richard Lang demonstrated a bug in his own program, and it was found that exactly the same bug existed in Quickstep, Mr. Langer confessed and was immediately disqualified. Mr Langer's embarrassment was compounded by the fact that he and his wife were on their honeymoon in Portoroz, and his wife witnessed his unmasking and disqualification.

The Espin Case

Much more recently the ICGA experienced a 21st century attempt at something similar, when the FIDE Master Johnadry Gonzalez Espin of Habana, Cuba, applied to enter the 2010 World Computer Chess Championship in Kanazawa, Japan. After making great efforts, successfully, to help Espin obtain a visa to participate in Japan, the ICGA was informed that “his” program SquarknII is a clone of the program Robbolito 0.85g3 with only 3 values changed in the entire code. Espin was duly barred from entering the tournament and will not be permitted to take part in ICGA events in the future. For more information about the Espin case visit

http://www.grappa.univ-lille3.fr/icga/news_item.php?id=57&lang=3

or

<http://susanpolgar.blogspot.com/2010/07/cuban-chess-program-disqualified.html>

The Rybka-Fruit Case

In cases such as the antics of Langer and Espin very little proof is needed to establish the cloning. But in some cases there is a more sophisticated cloning effort, when the clone programmer(s) attempt to hide their actions by making changes to the code of “their” program, presumably hoping to obscure the original source of the algorithms, ideas and the original code itself. The most serious allegations we have come across of this type relate to Rybka, currently the world’s top rated chess program and the winner of the World Computer Chess Championship in 2007, 2008, 2009, 2010. Rybka’s programmer is Vasik Rajlich, an International Master. For more than three years we have been hearing rumours in the computer chess world that Rybka’s engine was derived from the program Fruit, programmed by Fabien Letouzey, which placed second in the 13th World Computer Chess Championship in Reykjavik in 2005. Soon after his success in Reykjavik Fabien Letouzey made his program open source, under a Gnu Public License (GPL), so its copyright is now controlled by the Free Software Foundation.

In order to consider how the published Fruit source code might have influenced the development of Rybka, it is perhaps useful to examine some of the history of both programs. First let us go back a few years, to a time before the Fruit source code was made public. The Hiarcs forum contains the results of the CCCT6 tournament, played on January 31st and February 1st 2004, in which Rybka finished in 53rd place out of 54 contestants.

<http://www.hiarcs.net/forums/viewtopic.php?p=35546&sid=fa5c074eadaf5789ab6119878ae0eddd>

On the Fruit Web site we find the following details of the open source versions of Fruit (<http://www.fruitchess.com/about.htm>)

“It made its first appearance to the public in March 2004. Fruit was then just a basic program with a very simple evaluation and basic search. However since then it made skirmish progress adding about 100 Elo to each new release (1.5, 2.0, 2.1 and Fruit 2.2). The latest version from Fabien is "Fruit Beta 05/11/07" compiled on November, the 3rd 2005. Since then no new versions were released.

Until Version 2.1, Fruit was open source. But with Fruit becoming the strongest engine, the author decided to close the source code to avoid clones which might participate in official tournaments.”

And furthermore, Fruit 2.1 was released with source code on June 17th 2005 under the GNU GPL license.

Let us now consider the point in time when it became clear that Rybka had become enormously strong. From Wikipedia we learn that: (<http://en.wikipedia.org/wiki/Rybka>)

“Vasik Rajlich started working on his chess program at the beginning of 2003. The first Rybka beta was released on December 2, 2005 . . . In December 2005, Rybka participated in the 15th International Paderborn Computer Chess Championship. Rybka won the tournament with a score of 5½ points out of 7, ahead of other engines such as Gandalf, Zappa, Spike, Shredder and Fruit.”

So Rybka’s first outstanding tournament success would seem to have been in December 2005, six months after the date of the release of the open source version of Fruit 2.1. One can understand from this coincidence of timing how many computer chess experts might have been led to think that Rybka’s development owed a considerable debt to the Fruit source code.

But as I have mentioned, at first the Rybka-Fruit case was mere rumour. More recently, however, these rumours have become firm allegations, made by expert chess programmers and supported by evidence which appears on the surface to be rather compelling, both in its nature and in its volume. At this point in time I do not intend to make any definitive statement of my own on these allegations, but will allow the reader to form their own opinion after reading the following.

First, here is a posting by Zach Wegner, who currently develops (with the full permission of Anthony Cozzie, the original Zappa programmer) an upgraded version of Zappa, the World Computer Chess Champion in 2005. Wegner participated in the 2010 World Computer Chess Championship with their program which is called Rondo.

Evaluation

Rybka's evaluation has been the subject of much speculation ever since its appearance. Various theories have been put forth about the inner workings of the evaluation, but with the publication of Strelka, it was shown just how wrong everyone was. It is perhaps

ironic that Rybka's evaluation is its most similar part to Fruit; it contains, in my opinion, the most damning evidence of all.

General Differences

Simply put, Rybka's evaluation is virtually identical to Fruit's. There are a few important changes though, that should be kept in mind when viewing this analysis.

- *Most obviously, the translation to Rybka's bitboard data structures. In some instances, such as in the pawn evaluation, the bitboard version will behave slightly differently than the original. But the high-level functionality is always equivalent in these cases; the changes are brought about because of a more natural representation in bitboards, or for a slight speed gain. In other cases the code has been reorganized a bit; this should be seen more as an optimization than as a real change, since the end result is the same.*
- *All of the endgame and draw recognition logic in Fruit has been replaced by a large material table in Rybka. This serves mostly the same purpose as the material hash table in Fruit, since it has an evaluation and a flags field.*
- *All of the weights have been tuned. Due to the unnatural values of Rybka's evaluation parameters, they were mostly likely tuned in some automated fashion. However, there are a few places where the origin of the values in Fruit is still apparent: [piece square tables](#), [passed pawn scores](#), and the flags in the [material table](#).*

Evaluation Detail

In the following pages I will go into more depth about the details of each aspect of the evaluations and their similarities and differences.

- *Pawn evaluation: [pawn_get_info\(\)](#)*
- *Piece evaluation: [eval_piece\(\)](#)*
- *King Safety/Shelter: [eval_king\(\)](#)*
- *Passed Pawns: [eval_passer\(\)](#)*
- *Patterns: [eval_pattern\(\)](#)*
- *[Material](#)*

Zach Wegner's full posting can be found at:

<https://webpace.utexas.edu/zzw57/rtc/eval/eval.html>

Now to some other comments on this matter. The main thread to look at is:

<http://www.talkchess.com/forum/viewtopic.php?t=37762>

from which the following quotations have been extracted:

"Vasik claimed that Strelka 2.0 is a clone of Rybka 1.0 (and you know what that would imply!)" - Fabien Letouzey [NOTE: Strelka is a Russian program that Yuri Osipov claims as his own, but that too seems to be a clone of Fruit. This might explain why some experts note a similarity between Strelka and Rybka. – DL]

" Zach Wegner found many Fruit ideas (and nearly identical code) in Rybka 1.0; I think someone else did, too" - Fabien Letouzey

"Implemented algorithms were "translated" to a bitboard infrastructure. The translator ends up with a working implementation of an algorithm he may never have been able to code from scratch. It is "copy and paste" with unavoidable adaptations/translations. That is how DanChess once cloned Crafty's evaluation, and was condemned in this forum as a clone." - Matthias Gemuh

"More important is

'It was however a whole re-write (copy with different words if you like, similar to a translation) of the algorithms. Not just an extraction of a couple of ideas as is common, and normal.'

That says a lot about the issue. In a short summary, strelka encapsulates the ideas and data structures in Fruit, although there is not a character by character match since strelka came from reverse-engineering of Rybka 1. He makes it pretty clear that the "source" for strelka was obviously fruit. Which is what several of us have been saying for a long time. Function names and variable names are, of course, meaningless, assuming the original Rybka beta was stripped of all symbols for reasons unknown. But as I mentioned when the debate started, there are global similarities that are way too significant to write off as random chance of two different programmers writing the same code..." - Bob Hyatt

"Of course, this is nothing new to those of us that looked and listened in the first place..." Bob Hyatt

"Bob, We never really met, so, nice to meet you! I am sorry that many people apparently did not listen to you." Fabien Letouzey.

" Me too. But if someone believes something strong enough, nothing will alter their opinion. Basic human nature. We have now reached a totally insane point in computer chess where we have dozens of strong programs, claiming to not be derivative works." - Bob Hyatt

"Proof that Rybka is a Fruit derivative has been presented since Rybka 1.0 beta, though some people will never accept any amount of proof as adequate." - Matthias Gemuh

"VR used Fruit code. Noone who looked at the facts can disagree." . . . " Takeing the code, and changeing it is a GPL violation. Thats what happened. There is no other explanation for similaries in non-chess-playing-related-parts like the UCI communication, error messages, crashes." - Alexander Schmidt

And now some more extracts from a forum dialogue between Bob Hyatt and Fabien Letouzey:

Bob Hyatt:

"Just for the record, to eliminate this specific argument, when Zach, CT, I and others looked at the fruit/rybka1 question, we did_not_involove Strelka. Strelka was the thing that exposed the issue, but we directly compared fruit to rybka, so the strelka issue could not be raised again..."

Fabien Letouzey:

" What happened then?"

Bob Hyatt:

"We found_lots_ of similarities. Zach created a web page that went into great detail with the analysis. There are some obvious differences between Fruit and Rybka, but there are a_ton_ of similarities. Too many to be pure luck."

"I would like to put events into proper time-frame / perspective.

- 1. Strelka came along.*
 - 2. Vas looked at it, claimed it was Rybka and therefore he was going to release the strelka source under his name.*
 - 3. He then discovered that even if you take a source from a questionable place, once you modify it, the modifications are_your_ copyright. Since Strelka admittedly had some changes made (improvements according to Vas) he then realized he could not release it as his code as then he would be violating the "author's" copyright. Even though Vas had violated the Fruit GPL.*
- So he was damned if he did (release the code as his own) or damned if he didn't (someone else was getting credit and his ideas were exposed)... That's all there was to that little episode." - Bob Hyatt*

"I am not sure how Rybka would be legal in light of the GPL fruit was issued under. I've not seen any source code for Rybka 1 beta, yet I have seen executables, which is a direct violation of the GPL. Strelka is a reverse engineered derivative of fruit. I am not sure what in the hell that means, however. :-)" - Bob Hyatt

"Maybe we can ask other people whether they would consider it a new engine. Apart from the different ordering of moves with equal score (due to mailbox vs. bitboard move generation order), the engine behaviour would be identical.

Same search, same eval.

By "huge undertaking", maybe you mean that one has to understand CC basics and modify a lot of code.

But the way I see it is different:

Which is the hugher undertaking:

a) convert Fruit to bitboards

b) write a world-class engine (as of 2005) from scratch" - Fabien Letouzey

"I have always said that changing the infrastucture of move generation (e.g. from mailbox to bitboards) and then adapting (or "translating") the remaining source code is just another form of "copy and paste". The result of the adaption is not a new engine. As a chess programmer, I know what I'm talking about." - Matthias Gemuh

"If you start with the complete Fruit code and change it, is it copy and paste? Yes. It's a copy of the whole code. You answered the question yourself. Thats what happend with Rybka, there is no other logical explanation for the similaries in not chess related parts." - Alexander Schmidt

Responses from Vasik Rajlich

When it was suggested in 2007 in an Internet posting that Rybka was a clone of Fruit, Vasik Rajlich strongly denied it, see:

http://rybkaforum.net/cgi-bin/rybkaforum/topic_show.pl?pid=19118#pid19118

"Osipov's speculation is not correct. Rybka is and always was completely original code, with the exception of various low-level snippets which are in the public domain.

Rybka's scores are minimax score - they are propagated up the search tree. In principle, they should be from the tip of the PV, but because Rybka takes the PV from the hash table, this may not always be the case.

Re. depth, this is simply a tool to drive the iterative search. By conventional I mean 'in the normal range'.

Vas"

Additionally, when the origins of Strelka became the subject of debate heated in the computer chess forums, Vasik pitched in with his own comments, claiming that Strelka was a clone of Rybka. Vasik posted the following on the Rybka forum.

http://rybkaforum.net/cgi-bin/rybkaforum/topic_show.pl?pid=19118#pid19118

By Vasik Rajlich Date 2008-01-11 12:26

I've taken a look this morning at the Strelka 2.0 sources. The picture is quite clear.

Vast sections of these sources started their life as a decompiled Rybka 1.0. The traces of this are everywhere. The board representation is identical, and all sorts of absolutely unique Rybka code methods, bitboard tricks and even exact data tables are used throughout. Significant portions of the search and evaluation logic are not fully disassembled - the author has left in hardcoded constants and used generic names (such as "PawnStruScore0" & "PawnStruScore1", "PassedPawnValue0" through "PassedPawnValue7", etc) which show that he hasn't yet fully understood what is happening.

In some cases, these traces do also extend beyond the inner search and evaluation kernel. For instance, Rybka and Strelka are the only engines which I know about which don't report "seldepth" and "hashfull". Rybka's UCI strings are used throughout.

The author did at first make attempts to hide the Rybka origins, for example by masking the table values in earlier Strelka versions. He also made significant attempts to improve the program. The attempts at improvement are not very original, but they are everywhere. They include PV collection, null verification (and in fact changes to the null implementation itself), some endgame drawishness heuristics, a handful of new evaluation term, a new approach to blending between opening and endgame eval terms, and so on. They also do include various structural changes, such as knight underpromotions, on-the-fly calculations of many tables, the setting of piece-square table values, etc. These changes are extensive and no doubt lead to differences in playing style and perhaps a useful engine for users to have, but they do not change the illegality of the code base.

In light of the above, I am claiming Strelka 2.0 as my own and will release it in the next few days under my own name. The name of the author with the pen name "Osipov" will be included if he comes forward with his own real name, otherwise an anonymous contribution will be noted. The contributions of Igor Korshunov will also be confirmed and noted if appropriate. All usage permissions will be granted with this release.

I do not see obvious signs of other code usage, but perhaps this deserves a closer look. Some of the transplanted ideas, such as the null verification search, are rather naive implementations of the approach in Fruit/Toga, although my first impression is that that code itself is original. The Winboard parser from Beowolf which was added to Strelka 1.0 seems to have been completely removed. If someone else does find other signs of code theft, please get in touch with me and I will give proper credit in the upcoming release.

If someone has suggestions about an appropriate license, and in particular the pros and cons of the GPL for a chess engine and for this unusual scenario, or if someone would be willing to help in preparing this code and license for release, please also get in touch with me.

As this code is two years and several hundred Elo old, I am not going to launch any major action. However, 'Osipov' has already threatened to repeat the procedure with Rybka 2.3.2a. (He did this after I declined to grant him rights to commercialize Strelka.) If this situation does repeat with a newer Rybka version, I will not just stand and watch any more. In the meantime, if someone has information about 'Osipov', please get in touch with me.

Furthermore, when I contacted Vasik a few days before writing this article, inviting him to comment on Zach Wegner's analysis, he responded as follows:

"Hi David,

I'm not really sure what to say. The Rybka source code is original. I used lots of ideas from Fruit, as I have mentioned many times. Both Fruit and Rybka also use all sorts of common computer chess ideas.

Aside from that, this document is horribly bogus. All that "Rybka code" isn't Rybka code, it's just someone's imagination.

*Best regards,
Vas"*

And when I asked for clarification as to whether this response meant that the Rybka 1 source code was original, Vasik replied:

"all of the Rybka versions are original, in the sense that I always wrote the source code myself (with the standard exceptions like various low-level snippets, magic numbers, etc)."

Fringe Problems

There is one other type of offence that I would like to mention here in connection with cloning, namely entering a cloned program created by someone other than the entrant, in a tournament, with the entrant knowing it be a clone. One might draw an analogy between the criminal law offence of theft and the crime of handling goods knowing them to be stolen. This offence in the computer chess world is similar to one that recently caused something of a scandal in the Netherlands, when a board member of the Dutch Computer Chess Association (CSVN), the body that organises the prestigious Leiden tournaments entered a pirated copy of Junior in one of the major online annual tournaments. (See <http://www.hiarcs.net/forums/viewtopic.php?t=3896> for more details.) Put simply, if someone knows that a program has been ripped off, either by cloning or through piracy, they will not be permitted to use a ripped off copy to compete in any ICGA event.

How to investigate such allegations and deal with cloning?

The ICGA intends to set up a forum for investigating prima facie claims of cloning in the world of computer strategy games. Claims that are proven to the satisfaction of the ICGA will result in sanctions being imposed by the ICGA on the offending persons, who will be named and shamed on the Internet.

Setting up such a forum for chess will require the support of leading members of the computer chess fraternity. We will need people willing to examine and compare source codes and to write reports on what they discover. The ICGA does not have a source of funds to pay for any such work, so anyone helping us will be a volunteer. Our current thinking is to make this chess forum open only to those who have already participated with their own chess program in an ICGA event. Anyone who comes into this category will be most welcome as a founder member of the group.

The first thing we need is someone willing to set up and operate a bulletin board where members of the forum can “meet” and exchange views. Will someone volunteer to do this to help the ICGA on its way to stamping out these insidious practices?

David Levy can be reached at davidlevylondon@yahoo.com