A Gross Miscarriage of Justice in Computer Chess

by Dr. Søren Riis

Introduction

In June 2011 it was widely reported in the global media that the International Computer Games Association (ICGA) had found chess programmer International Master Vasik Rajlich in breach of the ICGA’s annual World Computer Chess Championship (WCCC) tournament rule related to program originality. In the ICGA’s accompanying report it was asserted that Rajlich’s chess program Rybka contained “plagiarized” code from Fruit, a program authored by Fabien Letouzey of France.

Some of the headlines reporting the charges and ruling in the media were “Computer Chess Champion Caught Injecting Performance-Enhancing Code”, “Computer Chess Reels from Biggest Sporting Scandal Since Ben Johnson” and “Czech Mate, Mr. Cheat”, accompanied by a photo of Rajlich and his wife at their wedding.

In June 2011 it was widely reported in the global media that the International Computer Games Association (ICGA) had found chess programmer International Master Vasik Rajlich in breach of the ICGA’s annual World Computer Chess Championship (WCCC) tournament rule related to program originality. In the ICGA’s accompanying report it was asserted that Rajlich’s chess program Rybka contained “plagiarized” code from Fruit, a program authored by Fabien Letouzey of France.

In response, Rajlich claimed complete innocence and made it clear that he found the ICGA’s investigatory process and conclusions to be biased and unprofessional, and the charges baseless and unworthy. He refused to be drawn into a protracted dispute with his accusers or mount a comprehensive defense.

This article re-examines the case. With the support of an extensive technical report by Ed Schröder, author of chess program Rebel (World Computer Chess champion in 1991 and 1992) as well as support in the form of unpublished notes from chess programmer Sven Schülle, I argue that the ICGA’s findings were misleading and its ruling lacked any sense of proportion. The purpose of this paper is to defend the reputation of Vasik Rajlich, whose innovative and influential program Rybka was in the vanguard of a mid-decade paradigm change within the computer chess community.
**History of Rybka**

On December 4, 2005, the computer chess community was astonished by the initial release of a free, downloadable chess program named Rybka 1.0 Beta, which within days took a sizable lead on all then-existing chess program rankings, surpassing all commercial programs, including renowned engines Shredder, HIARCS, Fritz and Junior. Starting from this top perch, Rybka then proceeded to rapidly widen its lead with subsequent versions. Rybka went on to become a commercial engine in 2006.

Rybka maintained unbroken supremacy on the chess engine rating lists for five years. However its performance in dozens of competitive tournaments held all over the world was, if anything, even more spectacular. Rybka did not merely win nearly every tournament it entered; it won them with a near-90% success rate. It is difficult to overstate the degree of superiority that the Rybka team exhibited in these years in chess software, mastery of hardware, and even in opening theory.

Going from one innovation to another, the Rybka team during this period developed the strongest chess-playing entity in the history of mankind up to the present day: the Rybka Cluster, which marshals the power of 300 late-model computer processors and plays chess at a hitherto unmatched and as-yet undefeated level. Rybka was the first chess program to routinely produce highly artistic masterpieces of chess while avoiding a great many pointless “computer” moves that for many years had been a source of ridicule among strong human players.

The latest public edition of Rybka (Rybka 4.1) is more than 300 Elo points stronger than the top competitors that existed in late 2005 on comparable hardware (who, at the time, were themselves approximately as strong as the world’s top grandmasters). Rajlich has indicated that his next edition of Rybka (Rybka 5) will be issued in the first half of 2012.

**The ICGA’s investigation and outcome**

In early 2011 sixteen chess programmers, many of whose programs were in direct competition with Rybka, signed a letter wherein they asserted that Rajlich copied programming code from another engine, Fruit, authored by Fabien Letouzey and released to the public in June 2005, about six months before Rybka 1.0 Beta. They requested that the ICGA investigate their charges and, implicitly, take punitive action on the grounds that Rajlich had violated WCCC tournament rules. At this point over five years had elapsed since the alleged offense, and four consecutive world computer chess championships had been decisively won in head-to-head competition by (newer and uncontested versions of) Rybka.

In response to the accusing letter the ICGA formed a committee consisting of 34 experts, some with genuinely distinguished CVs, for their investigation. Approximately seven of these experts actively participated in the discussions, and three of them wrote a report wherein they argued that Rybka 1.0 Beta had plagiarized large parts of Fruit. None of the actual Rybka versions that participated in the four WCCC tournaments were investigated.

The ICGA committee found that Rybka 1.0 Beta had violated Rule 2 of the ICGA-organized WCCC, and published extensive findings in support of its action. Ed Schröder was one of the original sixteen signatories calling for an investigation. However, he later began to doubt the ICGA’s investigatory process and, ultimately, its findings. Eventually he had a complete change
of heart and published his own investigation with the help of five other dissenting chess programmers.

A paradigm shift in computer chess

The last seven or eight years saw some striking changes in computer chess, the most striking being the breakthrough program Fruit, authored by Fabien Letouzey, released in 2004 and steadily improved as an open source engine through June 2005. Until then most chess programs, from the early 1990s through 2004, were “slow climbers,” progressing relatively slowly. Then something happened: the release of of Fruit and the open-source engine Stockfish, and especially the release of reverse-engineered Rybka derivatives, highly detailed recipes for building strong, modern chess engines went into the public domain. Existing and new programs receive a tremendous performance boost and become “fast climbers”.

These facts do not merely suggest that everyone in the top tier of chess programming learned from Fruit. Retrospectively, what now seems clear is that Fruit also unwittingly triggered a revolution in the whole ethos of chess programming. From the emergence of Fruit the premise within the programming community was that it was perfectly fine to re-use and share ideas and algorithms from leading programs whether they were open source or not. This was a new attitude which led us to the current landscape in computer chess. The ICGA did not recognize this transformation in computer chess in 2004-2005, for which they can be forgiven (institutions almost never immediately see when long-established traditions and norms have quietly slipped into obsolescence). But for them to still not recognize in 2011 how the chess software development panorama had changed, and to fail to adapt accordingly, risked putting the integrity of the WCCC enterprise in jeopardy.

How to succeed in programming without really trying

What made Fruit such an epochal breakthrough was that it expressed old ideas far more cleanly and efficiently than anyone else had done before – and the program was open source! Rybka was released six months later, incorporating every useful Fruit idea that could be incorporated, and adding many more improvements. It assumed a position of huge dominance. Other program authors likewise learned from Fruit and went after Rajlich like a pack of hounds chasing a fox.

Then something unprecedented happened: in May 2007 a strong closed source program called Strelka, which was clearly a very close derivative of Rybka, was released by a hitherto unknown individual. With this a very public precedent was established: someone had reverse-engineered a closed-source program with impunity. Rybka 1.0 Beta’s secrets were revealed for any programmer to use. And sometime thereafter another key Rybka secret, namely how Rajlich tuned his evaluation, was stolen by dishonest individuals whose apparent goal was to forcibly strip every last bit of proprietary information from Rajlich. In mid-2009 Rybka 3 was reverse-engineered by a group of chess programmers using fake names and published on the Internet. Soon thereafter chess programmers were freely using this source code to create strong Rybka-flavored derivatives.

All of this is well-known history: it is clear is that Rajlich’s original ideas have been lifted from various reverse-engineered editions of Rybka again and again and again – his work has been pilfered as comprehensively as anyone’s in all of computer chess history. Yet it is Rajlich who was investigated and found guilty of plagiarism in absentia, banned for life, stripped of titles, and
vilified in the international press over a five-year-old tournament rule violation. Note the irony: the chess program that has been the most influential to the others over the past several years (for very good reasons) is the target of plagiarism allegations!

**On the matter of “plagiarism”**

Rajlich conceded from the very beginning that he had studied Fruit’s open source code very closely and learned a great deal from it. In an interview from 2005, right after Rybka 1.0 Beta was released, he acknowledged his own debt to Fruit: “The publication of Fruit 2.1 was huge. Look at how many engines took a massive jump in its wake: Rybka, HIARCS, Fritz, Zappa, Spike, List, and so on. I went through the Fruit 2.1 source code forwards and backwards and took many things.” He specified that the ideas he found really useful were “Very specific tricks, mostly related to search, and the philosophy of the engine (and in particular of the search). Fabien is a very good engineer, and also has a very clear and simple conception of how his search should behave.” Finally, Rybka 1.0 Beta’s Readme file gave credit to Fruit in a “Special Thanks” section. By definition, plagiarism only happens when credit to sources is not given. These acknowledgements were widely known at the time Rybka entered WCCC tournaments and, rather obviously, Rybka’s debt to Fruit steadily diminished over the years as program improvements superseded the original program code in Rybka 1.0 Beta.

**Playing similarity testing of computer chess programs**

When comparing chess engines in a code-copying investigation what really matters is how similarly they play. If two chess programs played exactly the same moves, this would be prima facie evidence of a direct translation of program A into program B, rather than an original program. A computer chess theorist with statistical expertise conducted an analysis of “ponder-hits” produced statistical “fingerprints” of a number of chess engines. His results: the ponder-hit “fingerprints” show that the similarity between Fruit and Rybka 1.0 is actually not large, certainly not when compared to other programs. Moreover, the difference between Fruit and later versions of Rybka 3 and Rybka 4.1 is so large that it’s fair to say there is no similarity at all in practical terms. Another expert came up with similar findings based a larger set of data: the distance between Rybka 1.0 Beta and Fruit 2.1 is not less than other unrelated engines.

**What defines an original program?**

The widely understood ‘originality’ standard among programmers is that they may adopt the ideas of others but not duplicate their actual source code. This was articulated by Fruit programmer Fabien Letouzey in 2008: “Whether an engine is someone's ‘own work’ makes little sense to me. The reason is that all engines, whether amateur or commercial, share most of the techniques. Alpha-beta, iterative deepening, check extensions, null move, etc ... are shared by most and have been published, mostly by researchers, some of them more than 30 years ago!” About his program he explained: “I can't think of a search feature in it that was not described before.”

**Evaluation: a tale of two programs**

The core essence of the ICGA’s case against Rajlich is that Rybka and Fruit have very similar positional evaluations. The problem is that the ICGA’s findings are based on evidence that is simply untrue, evidence that is true with conclusions that are false, or evidence and conclusions that are true but irrelevant or immaterial.
It is important to itemize the significant ways Rybka’s evaluation differs from Fruit’s. In our paper we have enumerated ten substantive evaluation differences which, combined with Rybka’s entirely different search and board representation, signify that Rybka and Fruit must be considered two different chess engines by any reasonable person. These differences go a considerable distance to explain why Rybka 1.0 Beta played some 150+ Elo points stronger than Fruit 2.1. We can also see, in a series of notes kept by Rajlich in 2004 and 2005, that he was intellectually engaged in evaluation in the same period that his accusers claim he was feverishly copying Fruit’s evaluation. It is exceedingly hard to see the point of developing a slew of original ideas for Rybka while he was copying Fruit’s evaluation.

Given these points it is hard to make anything of the following categorical statement by Zach Wegner in his ICGA report findings: “Simply put, Rybka's evaluation is virtually identical to Fruit’s.” It is demonstrably incorrect and tendentious, and extremely misleading to someone who lacked the requisite technical expertise or was not prepared to invest the necessary time to study the full contents of his paper.

**Dots amazing: the case of the errant ‘0.0’**

Fruit used a system of floating point numbers for managing its time; Rybka 1.0 Beta had a faster and simpler approach using integers for checking time. However the ICCA experts found floating point in the Rybka code, which they assumed could only have come from a copy process. Ironically, the basis of the ICGA’s argument boils down to an interpretation of one line of source code in Rybka 1.0 Beta, which contains ‘0.0’ (no joke, ‘0.0’ appearing in a program written in 2005 has been a major issue for the ICGA investigators). The line is:

```
If (movetime >= 0.0)
```

Instead of what one would expect in an integer-based program, which should look like this:

```
If (movetime >= 0)
```

Hundreds of forum postings discussed this “smoking gun” proof of plagiarism: the fact that Rybka uses a floating point convention just like Fruit, is undeniable proof that code-copying occurred.

When asked Rajlich was baffled how the ‘0.0’ might have entered Rybka code (“It’s definitely weird/wrong”). But other programmers provided the answer. When you type “movetime >= 0.” i.e. if you add a period behind the zero, it compiles to exactly the same as if ‘0.0’ had been coded. Any programmer will tell you how easily this happens and how often it has happened to them in their programming.

So ultimately the smoking gun proof of copying comes down to a dispute about one extra keystroke, one single dot, on one line of code, that incidentally has zero impact on how the program actually plays. On what reasonable basis can a person conclude from this one superfluous dot that Rybka is non-original and Rajlich deserves to have all his titles stripped and be banned for life? How could this literally nugatory piece of evidence tip the scales in favour of the prosecution? How many devils can dance on a dot of code?
One of the central arguments in the ICGA report has been that Rybka and Fruit had very similar “Piece-Square Tables” (PSTs). In actual fact, every one of Rybka’s PST values is different from those in Fruit, as Dr. Miguel Ballicora persuasively shows here and also here. In addition, Ed Schröder has summarized Ballicora’s and Chris Whittington’s analysis on his website here. The arguments of these gentlemen simply demolish the PST case.

The ICGA report contends that Rybka used Fruit’s PSTs in its evaluation function. To support this the report provides page after page of near-identical source code side by side. However, it turns out on closer inspection that the most damning portion of the ICGA report is in fact a work of fiction and therefore profoundly misleading. Ed Schröder’s reports [1] [2] make the case powerfully. And this is what Vas Rajlich said to Dr. Levy in a terse email when the ICGA investigation was in progress: “This is horribly bogus. All that ‘Rybka code’ isn’t Rybka code, it’s just someone’s imagination.” You can only come to one conclusion: either Rajlich is flat-out lying or the ICGA report is wrong, or possibly even fraudulent. There is no middle ground.

Critics of the ICGA soon realized that no one had actual Rybka source code from before 2010, not even Rajlich himself, who sheepishly admitted that he had never maintained any form of version control for Rybka source code until Rybka 4. This is entirely believable because Rajlich asked on the Rybka forum for copies of his own program long before the ICGA controversy started.

This leads to the next insight. One of the premises of the ICGA’s report is that original Rybka source code can be reconstructed from the reverse-engineered binary of Rybka 1.0 Beta. That is absolutely not possible because of the one-to-many mapping involved in the process. There are many ways that Rajlich could have written Rybka and it is impossible to say exactly which path he took. This was confirmed by Dr. Hyatt himself: “We are not trying to take a binary executable and turn it into C. That is a one to MANY (MANY = INFINITE) mapping.” In spite of this the ICGA report shows us five pages of near-identical code, with actual Fruit code in the left column and “Rybka code” on the right. Here’s an actual example of falsification in the ICGA’s report:

```
Fruit
static const int KingCentreEndgame = 12;
static const int KingFileOpening = 10;
static const int KingRankOpening = 10;
...
for (sq = 0; sq < 64; sq++) {
    if (piece == sq.Endgame) {
        KingCentreEndgame = KingLine[square_file[sq]] * KingCentreEndgame;
        KingFileOpening = KingLine[square_rank[sq]] * KingFileOpening;
    }
}

Rybka
static const int KingCentreEndgame = 12;
static const int KingFileOpening = 10;
static const int KingRankOpening = 10;
...
for (sq = 0; sq < 64; sq++) {
    if (piece == sq.Endgame) {
        KingCentreEndgame = KingLine[square_file[sq]] * KingCentreEndgame;
        KingFileOpening = KingLine[square_rank[sq]] * KingFileOpening;
    }
}
```

Note the "0" used ???? FAKE CODE. NONE of this code is in Rybka
Had the ICGA titled the right-hand column "functionally-equivalent code possibly used by Rybka" that would still be misleading, as all that would be compared would be schematic PSTs with low information content. But at least there would be truth in the labeling.

As an aside we should note that the third line of the manufactured Rybka code does not make sense at all. Are we supposed to believe that the weight is multiplied by zero in the static declarations? If this were actually so it would not appear in the executables because any compiler would optimize and remove the unnecessary step. The only conclusion someone familiar with programming can make is that this code is fictitious.

It is reasonable to conclude that the ICGA members who drafted the report knew exactly the desired effect that labeling pages of speculative material "Rybka" would have. They could not have failed to intuit that most people lack the technical expertise and the time to comprehensively audit and assess technical documents. And on some level Dr. Hyatt in particular must have known that the mild-mannered Rajlich would not fight the charges, knowing that he would be assailed by an unending hail of accusations, insults and sophistries. I can say this because I and others have publicly defended Rajlich, and that is exactly what has happened over the course of thousands of Dr. Hyatt’s posts. In my capacity as Rybka forum moderator I have access to posting statistics. The chart below speaks for itself. Four months of relentless attacks on Rajlich’s own website!

These observations are not personal; they are simply factual evidence of the singular intensity and motivation of Rajlich’s chief accuser. Imagine how long it would take you to write 40 lucid forum posts in one day. Dr. Hyatt achieved this stupendous level of vitriol no fewer than 26 times in a four month span, peaking at 71 posts. Yet, Dr. Hyatt believes this is perfectly normal behavior for an associate professor of computer science and is not a relevant datum. I mention it because I think the reading public may have justifiable concern about Dr. Hyatt’s excessive devotion to the Rajlich-is-Guilty crusade.
Dr. Hyatt’s explanation

For a protracted period of time following the release of the ICGA report, Dr. Hyatt repeatedly stated that Rajlich made a direct copy of Fruit, and referred readers to the report to prove his case, citing the side-by-side comparisons shown there as functionally equivalent to DNA evidence. He did not even concede the “functional equivalency” cited in the report until this point was brought to his attention. Over the course of the forum debates Dr. Hyatt made a series of three remarkable statements which tell us what actually happened.

1. Statement 1 - 26 July 2011: The evidence is not based on "conjecture". It is based on specific analysis of Rybka and Crafty or Rybka and Fruit. There is no "interpretation" required. Have you actually read Zach's and Mark's report? People keep saying "show me side by side comparisons." First page of Zach's report has exactly that. Two columns. The comparison goes on for pages and pages. Side by side. Piece by piece...

2. Statement 2 - 29 July 2011: you realize that the code on the right is imaginary? It is the code on the left, with the weights modified, so that you get the same PST values that Rybka actually uses.

3. Statement 3 - 7 October 2011: The easiest way to show a layperson that the Fruit source matches the Rybka binary is to make our "pseudo-Rybka source" match Fruit as closely as possible.

This may be a good moment to take two aspirin pills. Let's summarize these statements:

1. There is Fruit and Rybka code side by side in the report. Pages and pages of it.
2. OK, we admit the Rybka code was imaginary, with weights modified.
3. OK, we admit the Rybka code we imagined with weights modified was deliberately manipulated to look identical to Fruit code. (!)

Caught in this web of his own making, at one point Dr. Hyatt even admitted that the PSTs in the Rybka column were not copied code, boldly asserting: “There was NO CODE COPYING for the PST issue. NONE. NADA. ZILCH. ZIPPO.” An emphatic statement! But two days later, apparently realizing that such a statement would mean that Rajlich was innocent, Hyatt changed his mind and wrote that Rajlich had copied PST code after all.

The ICGA’s problematic handling of the case

Apart from the substantive claims made by the ICGA a dispassionate observer ought to reflect on whether the structure and process of the investigation as well as its conclusions were reasonable and proportional to the alleged rule violation.

The ICGA decided to mount an investigation of Rybka after sixteen programmers submitted an open letter wherein they claimed Rybka contained illegal Fruit code. A panel was formed. Dr. Hyatt served as panel gatekeeper and determined who was and was not allowed to participate. Rybka competitors, individuals with obvious conflicts of interest, and individuals who had publicly expressed their predetermined conclusion of guilt, were allowed to join the investigation. While this jury-stacking was going on the president of ICGA, Dr. David Levy, made a preemptive declaration of Rajlich’s guilt in a ChessVibes article, before his own panel had had sufficient time to investigate and fully deliberate the facts.
Not even half of the original committee of 34 voted for a guilty verdict. Was it even clear in advance how many guilty votes were needed to convict? In addition members on the panel were only asked to decide the issue of guilt or innocence. They had no influence on the kind of penalty that would be handed down were they to find Rajlich guilty. The matter of sentencing was in the hands of the ICGA’s board, headed by Dr. Levy, which he duly exercised.

While no one questions the fact that the ICGA gave Rajlich ample opportunity to respond to their charges and he did not, there is much more to the matter than “we queried him and he did not respond.” Rajlich was not merely queried. He was publicly accused by the head of the ICGA and publicly excoriated by a group of individuals who stirred themselves up into a crusading lynch mob. A pile of “evidence” was jubilantly thrown together, based on a passionately-held predetermined conclusion of code-copying which happened to be wholly at variance with actual reality. Only then was Rajlich offered the opportunity to formally respond. The whole process was an unprofessional disgrace.

As for the nature of the punishment meted out by the ICGA, we might observe that justice can be defined as every man getting his due and letting the punishment fit the crime. There is no evidence that justice was done in this case in either sense, which is why I wrote this article: to publicly address an injustice and, perhaps, remedy it. We know that rule violations, if they occur, do not merit the equivalent of capital punishment rulings five years after the fact!

A subjective view of what really went down

We need to return to situation in computer chess just prior to the emergence of the Rybka allegations, to the matter of Rybka’s near-monopoly over computer chess competitions and chess engine commerce for a number of years. Not only did Rybka have a massive lead in tournament play, but it had access to massive hardware and its latest Rybka Cluster developments were locked up, beyond the reach of reverse-engineers. Rybka’s opening book was (and is) among the world’s best. Leading the team was Rajlich himself, a hypercompetitive genius with an insatiable desire to win and win again, and a business model that methodically froze out everyone else. He had no friends in his peer group to watch his back because he had no peers.

It is reasonable to conclude that this dominance was so pronounced and seemed so insurmountable to Rajlich’s rivals that they seized the only available opportunity to banish Rajlich and Rybka forever, not merely from ICGA-sponsored tournaments, but all tournaments anywhere in the world. It is also reasonable to conclude that other programmers found it unacceptable to attend week-long World Computer Chess Championship tournaments in far-off places like Beijing, China and Kanazawa, Japan, out of their own funds, paying entry fees, air fare, hotel, food and incidentals, only to be repeatedly blown off the board by a program whose dominance seemed to increase year after year with no end in sight. The economics of this no-win proposition understandably did not work for them. Rybka was just in a class by itself, everyone knew it, and this apparently intractable fact simply became intolerable.

Finally, it is reasonable to conclude that Rajlich’s long reign at the top of the rating lists, his monolithic dominance in public tournaments, sequence of menacing strategic actions such as his development of the Rybka Cluster, his publicly-stated intention to sequester his best development work so that it could not be reverse-engineered, his business alliances with Convexa and ChessBase and publicity juggernaut – all of these things and more marked
Rajlich as a convention-flaunting rogue programmer and hence, in the eyes of some, a public enemy.

In something of a surprise epilogue that took place as this article was in its final stages of being written, it emerged that the times they are a-changing for computer chess generally and the ICGA in particular. At the recent Rybka-less 2011 WCCC held in the Netherlands, none of the top seven ranked programs on the **CCRL 40/40** list (Rybka 4 is in second place) were able to participate, which stimulated doubts about the credibility of the “World Championship” title. During this competition the programmers met and some expressed a desire to change WCCC Rule 2. Without question, updating it to reflect contemporary reality would be a years-overdue positive step. However, without justice for Rajlich as a first step any proposed rule amendment would mean Rajlich would continue in his ICGA-imposed pariah status while other programmers would be free to use Rajlich’s ideas, algorithms and reverse-engineered source code (from existing and future editions of Rybka) with little fear of reprisal. First things must come first: the ICGA must retract a grave injustice inflicted upon a great chess programmer, world champion and innocent man.

**Acknowledgements**

Thanks to Ed Schröder for encouraging me to write this article as well as his insights on the computer chess scene going back decades. A special thanks to Nelson Hernandez, Nick Carlin, Chris Whittington, Sven Schüle and Alan Sassler for their first class editing as well as their many valuable suggestions. Without the lively collaboration of these individuals spanning several weeks this paper could not have been written. Finally, let me thank Vasik Rajlich for his clarification of various technical points and contemporaneous notes.

Thanks also to Dann Corbit, Miguel Ballicora, Rasmus Lerchedahl Petersen, Cock de Gorter, Jiri Dufek for their excellent suggestions and eagle-eyed proof reading.

---

*Søren Riis* is a Computer Scientist at Queen Mary University of London. He has a PhD in Maths from University of Oxford. He used to play competitive chess (Elo 2300).