

Naum 1.91



Erst im September letzten Jahres qualifizierte sich Naum mit der Version 1.8 erstmalig für die CSS-Rangliste. Vor kurzem präsentierte der Autor A. Naumov aus Serbien bereits eine aktualisierte Version. Die Version 1.8 des heute in Kanada lebenden Autors erzielte seinerzeit ein Einstiegsrating von etwa 2610 Punkten. Da Naumov in letzter Zeit enorme Entwicklungssprünge verzeichnete, war ich natürlich gespannt, wohin die Reise mit der aktuellen Version geht.

Am eigentlichen Spielstil konnte ich auf den ersten Blick keine großen Veränderungen erkennen. Naum pflegt nach wie vor einen grundsoliden Stil und vermeidet möglichst taktische Verwicklungen. Was die Knotenzahlen anbetrifft, ist Naum trotz größer gewordener Programmdatei schneller geworden. Dies wirkt sich insbesondere im Mittelspiel positiv auf die Rechentiefe aus, Naum 1.8 kommt nach einer Minute Bedenkzeit etwa einen Halbzug tiefer als die Vorgängerversion. Im Endspiel konnte ich keine großartigen Veränderungen feststellen.



Aleksander Naumov

Engine-Recheninformationen

	Naum 1.91	Naum 1.8	+/-
Engineversion	30.01.2006	04.08.2005	
Dateigröße (kB)	348	292	56
Knotenzahl Mittelspiel (KN/s)	779	621	158
Rechentiefe Mittelspiel (Halbzüge)	14	13	1
Rechentiefe Endspiel (Halbzüge)	19	19	0

Das allein muss noch nichts bedeuten, daher soll ein kompletter Testdurchgang gegen sämtliche gelisteten Engines der aktuellen CSS-Rangliste die Spielstärkeentwicklung aufzeigen. Aktuell liegt die Vorgängerversion bei einem Score von 40,6% bzw. einem Elorating von 2602 Punkten auf Rang 19 der Rangliste. Ältere kommerzielle Engines liegen also nur etwa 40 Elopunkte vor Naum.

Nach den 480 Partien des Testdurchgangs verbuchte Naum 1.91 eine Scoresteigerung gegenüber der Vorgängerversion von 3,85%, dies entspricht einer Elosteigerung von knapp 30 Punkten. Auffällig dabei, dass Naum 1.91 lediglich mit den schwarzen Figuren zugelegt hat, während die Punktausbeute als Anziehender eher stagnierte.

Gesamttestergebnis

	Naum 1.91	Naum 1.8	+/-
Gesamtscore:	44,48%	40,63%	3,85%
Rating (Elo)	2629	2602	27
Remisquote:	32,71%	30,00%	2,71%
Weiß/Schwarz-Score			
	Naum 1.91	Naum 1.8	+/-
Weißscore:	47,92%	47,71%	0,21%
Schwarzscore:	41,04%	33,75%	7,29%
Erzielte Ratings gegen verschieden starke Gegnergruppen			
Gegnergruppe	Naum 1.91	Naum 1.8	+/-
"Top 8" 2744 Elo	2650	2619	31
"Mittelfeld" 2664 Elo	2610	2573	37
"unteres Tabellendrittel" 2598 Elo	2625	2612	13

Am erfolgreichsten spielt Naum mit seinem nach wie vor sehr soliden Spielstil gegen die stärkeren Gegner. Sein erzieltes Rating gegen die „Top-8“ liegt bei 2650 Elo, gegen die Gegner aus dem Tabellenmittelfeld hingegen schon 40 Punkte niedriger. Herausragend dabei ist der überraschende Matchgewinn gegen Spike 1.0a sowie das Matchremis gegen Loop List 600.

Detaillierte Matchergebnisse im Vergleich zur Vorgängerversion

Gegner	Elo	Naum 1.91			Score	+/-
		+	=	-		
Fruit 2.2.1	2816	1	4	15	15,00%	-2,50%
Hiarcs 10	2803	3	7	10	32,50%	12,50%
Fritz 9	2794	5	4	11	35,00%	7,50%
Shredder 9 UCI	2770	3	9	8	37,50%	-5,00%
Loop List 600	2696	7	6	7	50,00%	12,50%
SmarThink 1.00	2695	6	4	10	40,00%	-5,00%
Spike 1.0a Mainz	2688	7	7	6	52,50%	27,50%
Ktulu 7.5	2688	3	8	9	35,00%	-12,50%
"Top 8"	2744	35	49	76	37,19%	4,38%
Chess Tiger 15.0	2688	3	6	11	30,00%	-15,00%
Junior 9	2686	6	6	8	45,00%	-5,00%
Glaurung 1.0.1	2663	5	7	8	42,50%	25,00%
Gandalf 6.0	2663	3	6	11	30,00%	2,50%
The King 3.33	2660	6	5	9	42,50%	2,50%
Pro Deo 1.1	2660	7	6	7	50,00%	-7,50%
Ruffian 2.1.0	2649	4	11	5	47,50%	12,50%
SlowChess Blitz WV2	2646	6	9	5	52,50%	27,50%
"Mittelfeld"	2664	40	56	64	42,50%	5,31%
SOS 5 for Arena	2624	8	7	5	57,50%	25,00%
Aristarch 4.50	2621	6	8	6	50,00%	7,50%
Pharaon 3.3	2601	8	6	6	55,00%	2,50%
Deep Sjeng 1.6	2596	8	5	7	52,50%	-15,00%
Pseudo 0.7c	2591	7	9	4	57,50%	10,00%
Delfi 4.6	2590	3	9	8	37,50%	-27,50%
Jonny 2.82	2584	8	5	7	52,50%	15,00%
Thinker 4.7a	2577	12	3	5	67,50%	-2,50%
"unteres Tabellendrittel"	2598	60	52	48	53,75%	1,88%
Gesamt	2669	135	157	188	44,48%	3,85%

Spike 1.1



Die UCI-Engine Spike vom Autorenteam Ralf Schäfer und Volker Böhm gehörte im letzten Jahr zu den Senkrechtstartern. Mit Spike 0.9a gelang im Mai 2005 problemlos die Qualifikation für die CSS-Rangliste. Mit einem Startrating von 2644 Elo platzierte sich Engine nur etwa 10 Punkte hinter der damaligen Freeware-Referenz ProDeo 1.1. Doch nur vier Monate später war selbst ProDeo keine wirkliche Hürde. Mit der neuen Version Spike 1.0 Mainz gelang eine Steigerung von 60 Elo, womit Spike die 2700-Elogrenze durchbrach. Für die Freewarekrone reichte diese enorme Steigerung jedoch nicht, diese hatte sich zwischenzeitlich Shootingstar Fruit 2.1 gesichert. Dennoch kann man diese Leistungsexplosion als phänomenal bezeichnen, denn zu diesem Zeitpunkt liegt Spike bereits auf Rang 3 der CSS-Rangliste, sogar vor Fritz 8 Bilbao und nur 80 Elo hinter dem damaligen Spitzenreiter Shredder 9 UCI zurück.

Spike-Versionen in der "ewigen CSS-Rangliste"

Rg	Programm	Version	Partien	+	=	-	Remis-Quote	Score	Elo-Gegner	Rating
16	Spike 1.0a Mainz	17.08.2005	740	304	220	216	29,7%	55,9%	2653	2694
32	Spike 0.9a	15.04.2005	620	232	195	193	31,5%	53,1%	2623	2644

Der derzeit rasanten Spielstärkeentwicklung musste auch Spike Tribut zollen. In der Folgezeit verliert die Engine über 20 Elo und wird bis auf Platz 10 durchgereicht. Der Abstand zur Tabellenspitze ist bereits auf fast 140 Elo angewachsen. Da kommt die Ankündigung der neuesten Version Spike 1.1 gerade recht. Mit dem Testdurchgang innerhalb der CSS-Rangliste sollte geklärt werden, ob dem Autorenteam wieder eine spürbare Steigerung gelungen ist.

Das Besondere an der Vorgängerversion Spike 1.0 Mainz ist, dass diese Version, entgegen sonstigen Geflogenheiten, am stärksten gegen schwächere Gegner punktet. Kein Wunder, Spike 1.0a wurde laut Ralf Schäfer fast ausschließlich gegen stärkere Amateurengines getestet. Sein Rating gegen diese Gegnergruppe liegt mit 2712 Punkten etwa 55 Elo über dem erzielten Rating gegen die Top-Gegner. Dies erklärt auch, warum Spike in der CSS-Rangliste meist etwas besser platziert ist als in anderen gängigen Ranglisten, welche Spike überwiegend gegen die stärkste Konkurrenz antreten ließen. Ziel musste es also sein, insbesondere das Ergebnis gegen die Topgegner zu steigern.



Ralf Schäfer und Volker Böhm

Die starke Gegnerschaft ist es auch, gegen die Spike 1.1, ganz im Gegensatz zur Vorgängerversion, am besten punktet. Die aktuelle Version steigert den Score gegen die starke Gegnerschaft um fast 14% und kommt mit diesem Ergebnis zu einem Score von ausgezeichneten 51,5%. Dies entspricht gegen diese Gegnergruppe einem Rating von fantastischen 2756 Punkten beziehungsweise eine Elosteigerung von knapp 100 Punkten. Insbesondere die zuvor mageren Resultate gegen Fruit 2.2.1 und Hiarc 10, welche unter der 20%-Marke angesiedelt waren, gehören der Vergangenheit an. Die Frage stellt sich nun, ob gegen die etwas schwächere Gegnerschaft ähnliche Leistungssteigerungen möglich sind. Gegen die Gegner aus dem Mittelfeld wollte dies nicht gelingen. Hier kam Spike auf ein Rating von 2724 Punkten. Ganz ausgezeichnet punktet Spike 1.1 allerdings nach wie vor gegen die Kontrahenten vom unteren Tabellendrittel. Zwar kommt die aktuelle Version hier nur zu einer Scoresteigerung von etwa 5,0%, allerdings hatte hier die Vorgängerversion bereits ein sehr gutes Resultat vorgelegt, sodass am Ende auch hier ein Rating von 2753 Punkten zu verzeichnen ist.

Gesamttestergebnis Spike 1.1

	Spike 1.1	Spike 1.0a Mainz	+/-	
Gesamtscore:	60,52%	52,29%	8,23%	
Rating (Elo)	2740	2682	58	
Remisquote:	33,54%	29,17%	4,38%	
Weiß/Schwarz-Score				
	Spike 1.1	Spike 1.0a Mainz	+/-	
Weißscore:	66,88%	55,42%	11,46%	
Schwarzscore:	54,17%	49,17%	5,00%	
Erzielte Scores gegen verschieden starke Gegnergruppen				
Gegnergruppe	Spike 1.1	Spike 1.0a Mainz	+/-	
"Top 8"	2745 Elo	51,56%	37,81%	13,75%
"Mittelfeld"	2656 Elo	59,38%	53,44%	5,94%
"unteres Tabellendrittel"	2598 Elo	70,63%	65,63%	5,00%
Erzielte Ratings gegen verschieden starke Gegnergruppen				
Gegnergruppe	Spike 1.1	Spike 1.0a Mainz	+/-	
"Top 8"	2745 Elo	2756	2660	96
"Mittelfeld"	2656 Elo	2724	2682	42

Diese Ergebnisse zusammengenommen ergeben nun ein Gesamtrating von 2740 Punkten beziehungsweise eine Elosteigerung zu Spike 1.0a Mainz von hervorragenden 58 Punkten.

Matchergebnisse Spike 1.1

Gegner	Elo	Spike 1.1				+/-
		+	=	-	Score	
Fruit 2.2.1	2818	6	5	9	42,50%	30,00%
Hiarcs 10	2800	3	8	9	35,00%	17,50%
Fritz 9	2792	6	5	9	42,50%	7,50%
Shredder 9 UCI	2773	6	7	7	47,50%	7,50%
SmarThink 1.00	2697	6	8	6	50,00%	-2,50%
Loop List 600	2693	8	10	2	65,00%	25,00%
Chess Tiger 15.0	2693	10	5	5	62,50%	-7,50%
Ktulu 7.5	2693	10	7	3	67,50%	32,50%
"Top 8"	2745	55	55	50	51,56%	13,75%
Junior 9	2689	6	7	7	47,50%	-10,00%
Pro Deo 1.1	2664	8	7	5	57,50%	-5,00%
Gandalf 6.0	2663	5	7	8	42,50%	10,00%
The King 3.33	2661	11	5	4	67,50%	10,00%
Glaurung 1.0.1	2657	8	8	4	60,00%	27,50%
Ruffian 2.1.0	2647	12	5	3	72,50%	7,50%
SlowChess Blitz WV2	2639	10	5	5	62,50%	-10,00%
Naum 1.91	2629	10	6	4	65,00%	17,50%
"Mittelfeld"	2656	70	50	40	59,38%	5,94%
Aristarch 4.50	2619	12	6	2	75,00%	12,50%
SOS 5 for Arena	2618	13	5	2	77,50%	25,00%
Pharaon 3.3	2601	9	6	5	60,00%	0,00%
Deep Sjeng 1.6	2601	7	11	2	62,50%	-12,50%
Delfi 4.6	2599	11	8	1	75,00%	15,00%
Pseudo 0.7c	2589	7	10	3	60,00%	-5,00%
Jonny 2.82	2580	14	3	3	77,50%	-5,00%
Thinker 4.7a	2579	12	7	1	77,50%	10,00%
"unteres Tabellendrittel"	2598	85	56	19	70,63%	5,00%
Gesamt	2666	210	161	109	60,52%	8,23%

Der Spielstil von Spike 1.1 wirkt deutlich angriffsbetonter als noch bei der Vorgängerversion. Dabei produzierte Spike einige sehenswerte Kurzpartien. Eine davon habe ich weiter unten kommentiert. Insbesondere das Spiel mit ungleichem Material mag Spike 1.1 und entschied sehr viele solcher Partien zu seinen Gunsten. Ebenfalls verbessert wurde laut den Autoren die Königssicherheit. Das größte Verbesserungspotenzial liegt meines Erachtens noch im Endspiel. Hier ist die

Bewertung oft allzu optimistisch und nicht selten wickelt Spike dann vorschnell in ein remisliches Endspiel ab. Besonders auffällig ist diese zu positive Bewertung bei Endspielen mit ungleichen Läufern. Aber kommen wir zu einer von den vielen Kurzpartien.

Spike 1.1 - Loop List 600 [B78]

CSS-Rangliste (10+10)

Klaus Wlotzka

1.e4 c5 2.Sf3 d6 3.d4 cxd4 4.Sxd4 Sf6 5.Sc3 g6 6.Lc4 Lg7 7.f3 0-0 8.Le3 Sc6 9.Dd2 Se5 10.Lb3 Ld7 11.0-0-0 Tc8 12.g4



Bisher folgten beide bekannten Pfaden. Statt des Textzuges wird meist 12.h4 h5 13.Lg5 Tc5 14.g4 hxg4 15.f4 Sc4 16.De2 b5 17.f5 mit guten weißen Angriffschancen am Königsflügel gespielt.

12...Dc7 13.h4 Spike sucht seine Chance am Königsflügel, interessant sieht auch 13.g5 Sh5 14.Sd5 Dd8 15.f4 Sg4 16.f5 Sxe3 17.Dxe3 a5 18.a4 Te8 mit aktiver weißer Stellung aus.

13...Sc4 14.Lxc4 Dxc4 15.Lh6 Lxh6?



Mehr Widerstand leistet 15...b5 16.h5 b4 17.Sd5 Sxd5 18.Lxg7 Kxg7 19.hxg6 fxg6 20.Dh6+ Kf7 21.Dxh7+ Ke8 22.Dxg6+ Kd8 23.exd5 Dxa2 24.Dg7 Te8 mit schwieriger Stellung für Schwarz.

16.Dxh6 Da6 Auch 16...Kh8 hilft nicht wirklich, denn nach dem Verteidigungsversuch 17.h5 Sg8 18.Dg5 h6 19.hxg6 Kg7 erzielt Weiß nach dem Turmpfer 20.Txh6! Sxh6 21.gxf7+ Kxf7 22.Dxh6 Th8 23.Df4+ Ke8 24.Sd5 eine vorteilhafte Stellung.

17.Td3 b5 Schwarz versucht seinerseits einen Gegenangriff am Damenflügel. Doch hierzu ist es schon zu spät.

18.h5 b4 19.g5 Sxh5 20.Txh5!!



Die Pointe. Das Turmopfer bringt die Entscheidung.

20...gxh5 21.Sd5 Kh8 22.Sxe7 Gegen die Drohung Df6+ ist kein Kraut gewachsen. 1–0

Partie online nachspielen

Aktuelle Rangliste

Seit dem letzten Bericht machte Spike den größten Sprung und liegt nur noch 30 Elo hinter „Altmeister“ Shredder 9 UCI. Eine fantastische Entwicklung dieser noch so jungen Engine von Ralf Schäfer und Volker Böhm. Ansonsten haben sich nur kleinere Verschiebungen ergeben.

Ausblick

Immer noch warte ich gespannt auf die Veröffentlichung von Rybka. Erst vor kurzem kündigte Vasik Rajlich an, dass eine Final-Version wohl erst im Mai 2006 veröffentlicht wird. Dafür soll in Kürze eine Zwischenversion mit dem Namen Rybka 1.1 erscheinen. Diese werde ich dann sofort in die CSS-Rangliste integrieren. Weiterer Testkandidat ist WildCat 6, welche gegenüber der zuletzt hier getesteten Version WildCat 4 offensichtlich deutlich stärker geworden ist.

Testbedingungen	
Bedenkzeit/Engine	10 Min. + 10 Sek./Zug
Partien je Gegner	20
Buchfunktionen	deaktiviert
Pondern	deaktiviert
Engine-Settings	default
Eröffnungsbücher	keine
Eröffnungsvorgaben	10 SeE-Stellungen (überarbeitet)
Endspieldatenbanken	Nalimov - 3-,4- und 5-Steiner
Testplattform	Fritz 8-GUI
Eloauswertung	ELOStat 1.1.c32 von Frank Schubert
Prozessor	AMD Athlon XP2800+ (Barton)
Hauptspeicher	1024 MB
Hashtables/Engine	128 MB
Tablebases-Cache	32 MB

Informationen zum Autor:

Klaus Wlotzka
