SELF-PLAY: STATISTICAL SIGNIFICANCE

G.M^{c}C. Haworth[1]

UK

ABSTRACT

Heinz recently completed a comprehensive experiment in self-play using the FRITZ chess engine to establish the 'decreasing returns' hypothesis with specific levels of statistical confidence. This note revisits the results and recalculates the confidence levels of this and other hypotheses. These appear to be better than Heinz' initial analysis suggests.

1. INTRODUCTION

Heinz (2000a/b, 2001a/b) reported an extensive self-play experiment with FRITZ to investigate the widely-held belief that an increment of search-depth gives decreasing benefit as depth increases. After a 'White v Black' calibration match between identical players, some seven 'alternating colours' matches were staged in which one player had one ply advantage in *fixed iteration depth p*, although both sides had search extension capability. The experiments included 3,000 games per match, taking the depth *p* beyond 9 plies to a record 12 plies.

This note offers an analysis of the statistical significance of the results which extends Heinz' own analysis, differs somewhat from it, and provides confidence levels for:

- Heinz' *null hypothesis*, given the approximating binary win/loss model he adopted,
- Heinz' null hypothesis based on a ternary win/draw/loss model, and
- two subsidiary null hypotheses, on win- and draw-rates, suggested by the data.

Section 2 tables Heinz' results, section 3 establishes the relevant mathematics, and section 4 analyses the experimental results. The summary mentions some open questions that remain.

2. THE SELF-PLAY EXPERIMENT

Let P_p be a player whose search strategy has a fixed iteration depth parameter of p plies. Let M_p be a match between P_p and P_{p-1}. The one match M_p of n games in Heinz' experiment results in an observed score S_{po} for P_p and an observed number - W_{po}, D_{po} and L_{po} – of, respectively, P_p wins, draws and losses: $n = 3,000$. Table 1 gives this data, and the derived observed P_p score, win, draw and loss frequencies, namely s_{po}, w_{po}, d_{po} and l_{po}.

p	M_p	W_{po}	D_{po}	L_{po}	S_{po}	n	w_{po}	d_{po}	l_{po}	s_{po}	ELO	p
Cal	M_{cal}: P_{8W}-P_{8B}	924	1,288	788	1,568.0	3,000	0.3080	0.4293	0.2627	0.5227		Cal
6	M_6: P_6-P_5	1,686	915	399	2,143.5	3,000	0.5620	0.3050	0.1330	0.7145	+159	6
7	M_7: P_7-P_6	1,643	1,066	291	2,176.0	3,000	0.5477	0.3553	0.0970	0.7253	+169	7
8	M_8: P_8-P_7	1,457	1,212	331	2,063.0	3,000	0.4857	0.4040	0.1103	0.6877	+137	8
9	M_9: P_9-P_8	1,433	1,235	332	2,050.5	3,000	0.4777	0.4117	0.1107	0.6835	+134	9
10	M_{10}: P_{10}-P_9	1,252	1,451	297	1,977.5	3,000	0.4173	0.4837	0.0990	0.6592	+115	10
11	M_{11}: P_{11}-P_{10}	1,124	1,525	351	1,886.5	3,000	0.3747	0.5083	0.1170	0.6288	+92	11
12	M_{12}: P_{12}-P_{11}	1,059	1,592	349	1,855.0	3,000	0.3530	0.5307	0.1163	0.6183	+84	12

Table 1. Results of the Heinz self-play experiment.

---

[1] 33, Alexandra Rd., Reading, Berkshire, RG1 5PG, UK. email: guy.haworth@laposte.net

3.   UNDERLYING MATHEMATICS

Let the corresponding underlying, intrinsic P_p score, win, draw and loss frequencies in M_p be s_p, w_p, d_p and l_p respectively. These are unknown, and treated as if they were random variables S_p, W_p, D_p and L_p whose distributions are to be estimated as well as possible from the observed data. w_p, d_p and l_p are binary variables, while s_p is a ternary variable which can be approximated by a binary variable. The experiment was carefully designed to ensure that the matches M_p were a sequence of independent Bernoulli trials of these random variables.

Let B be a sequence of n Bernoulli trials of a binary, random variable X with two values 0/1: $X = 1$ with probability x, $X = 0$ with probability 1-x. The following notation and results are standard (Feller, 1968):

$E(X) \equiv$ mean $\mu \equiv \Sigma P(v).v = P(1).1 = x.1 = x$; $E(X^2) \equiv \Sigma P(v).v^2 = P(1).1^2 = x.1^2 = x$

$Var(X) \equiv E((X - \mu)^2) = E(X^2) - 2\mu.E(X) + \mu^2 = E(X^2) - \mu^2 = x - x^2 = x(1 - x)$

$Var(nX) \equiv n.Var(X) = n.x(1 - x)$ for $n \geq 0$ $Var(-X) = Var(X)$

If Y is a second, independent random variable in the trial, $Var(X \pm Y) = Var(X) + Var(Y)$

$\sigma(X) \equiv$ standard deviation$(X) \equiv$ sqrt $Var(X) =$ sqrt$[x(1 - x)]$

$\sigma(nX) \equiv$ standard deviation$(nX) \equiv$ sqrt $Var(nX) =$ sqrt$[n.x(1 - x)]$

As $n \to \infty$, the distribution of $nX \to$ a normal distribution φ, mean $n.x$, standard dev. $\sigma =$ sqrt$[n.x(1 - x)]$

Let $x_o = x \pm \varepsilon.\sigma(x) \approx x$: then Prob[rate x_o is observed | mean x] $\approx \varphi(\varepsilon)$ and $\to \varphi(\varepsilon)$ as $n \to \infty$

But $x = x_o \pm \varepsilon.\sigma(x) \approx x_o \pm \varepsilon.\sigma(x_o)$ as $\sigma(x_o) \approx \sigma(x)$. Therefore, Prob[rate x is observed | mean x_o] $\approx \varphi(\varepsilon)$

3.1   Principles of Statistical Inference

If the initial, neutral, 'know nothing' assumption is made, namely, that any value in the interval (0, 1) is equally likely, then the following Bayesian relation holds between the underlying mean x and the observed rate x_o:

Prob[mean of X is ~x | rate x_o is observed] $\propto$ Prob[rate x_o is observed | mean of X is ~x]

But, if $x_o \approx x$, Prob[rate x_o is observed | mean of X is x] $\approx$ Prob[rate x is observed | mean is x_o]. Therefore,

Prob[mean of X is ~x | rate x_o is observed] $\to$ Prob[rate x observed | mean is ~x_o] as $n \to \infty$.

Therefore, in the limit, the distribution of the random variable $E(X)$, the mean of the random variable X, is the same as the distribution of X, assuming the observed rate x_o *is* the mean.

Note that, as $n = 3,000$, generously large, $n.x(1-x) > 290$ for all frequencies that occurred in the experiment. This is certainly large enough to ensure that:

- observed frequency rates are close to the intrinsic success rates for the random variables,
- the normal distribution $\varphi((X-\mu)/\sigma)$ is a suitable approximation to the actual probability density functions,
- the theory above may be applied to the experiment, knowing that approximation errors are ignorable.

3.2   Comparison of Binary and Ternary Models

Rather than analyse match M results with a ternary win/draw/loss model with random variables W and D, Heinz used a binary model with a two-valued random variable S: Pr(win, $S = 1$) = s. Thus, as above:

$E(nS) = n.s$ and $Var(nS) = n.s(1 - s)$

However, if S is a 3-valued random variable with Pr($S = 1$) = w and Pr(draw) = d, then:

$E(S)$ $= s = w + d/2$; $E(S^2) = w + d/4$

$Var(nS) = n.Var(S) = n.[E(S^2) - (w + d/2)^2] = n.[(w + d/4) - (w + d/2)^2] =$

       $= n.[(s - d/4) - s^2] = n.[s(1-s) - d/4] \leq n.s(1-s)$, the variance of the equivalent binary model.

Thus, the variance in the ternary model is always less than in the binary model[2]. Heinz' substitution of the binary model was therefore indeed a conservative one, increasing the apparent variances in his experiment.

---

[2] In the extreme, for example, it is 0 rather than 0.25 for $d = 1$ and one trial.

4. ANALYSIS OF THE SELF-PLAY RESULTS

Heinz' null hypothesis, S, is effectively that "s_p and therefore S_p reduce as p increases". The confidence level with which this statement may be made is the estimated probability of this being demonstrated again by a further match M_p of the same length.

Crucially, Heinz did not use the independence of S_i and S_j: $\text{Var}(S_i - S_j) = \text{Var}(S_i) + \text{Var}(S_j) = \sigma^2$ and $S_i - S_j = k_{ij}\sigma$. The value of k determines the confidence level with which we might say that $S_i > S_j$.

i \ j	$(S_i - S_j)/\sigma(S_i - S_j)$: binary model							Confidence Levels for the null hypothesis S on scores						
	6	7	8	9	10	11	12	6	7	8	9	10	11	12
6		-0.934	2.271	2.619	4.629	7.095	7.941		17.5%	98.8%	99.6%	100.0%	100.0%	100.0%
7			3.206	3.554	5.566	8.036	8.884			99.9%	100.0%	100.0%	100.0%	100.0%
8				0.348	2.355	4.813	5.656				63.6%	99.1%	100.0%	100.0%
9					2.007	4.473	5.316					97.8%	100.0%	100.0%
10						2.455	3.295						99.3%	100.0%
11							0.839							79.9%

Table 2a. $(S_i - S_j)$ in units of $\sigma(S_i - S_j)$ and confidence levels for the null hypothesis: binary model.

The distribution of the random variable S_p in the ternary model also converges on a normal distribution by the Central Limit Theorem. As proved above, $\text{Var}(S_i)$ and $\text{Var}(S_i - S_j) = \text{Var}(S_i) + \text{Var}(S_j)$ are lower in this model and Heinz' null hypothesis S may be stated with somewhat more confidence.

i \ j	$(S_i - S_j)/\sigma(S_i - S_j)$: ternary model							Confidence Levels for the null hypothesis S on scores						
	6	7	8	9	10	11	12	6	7	8	9	10	11	12
6		-1.216	2.991	3.458	6.299	9.699	10.958		11.2%	99.9%	100.0%	100.0%	100.0%	100.0%
7			4.357	4.843	7.832	11.354	12.677			100.0%	100.0%	100.0%	100.0%	100.0%
8				0.479	3.347	6.869	8.150				68.4%	100.0%	100.0%	100.0%
9					2.860	6.388	7.667					99.8%	100.0%	100.0%
10						3.625	4.915						100.0%	100.0%
11							1.256							89.5%

Table 2b. $(S_i - S_j)$ in units of $\sigma(S_i - S_j)$ and confidence levels for the null hypothesis: ternary model.

Clearly, we can state with 99.9%+ confidence that the winner's score decreases in the two match sequences $\{M_6, M_8, M_{10}, M_{12}\}$ and $\{M_7, M_9, M_{11}\}$.

The null hypotheses W, "W_p reduces as p increases", and D, "D_p increases as p increases", are suggested by the data: confidence levels have also been calculated for these statements. Both W_p and D_p are binary variables. Curiously, there is no obvious trend in the corresponding binary variable L_p.

i \ j	$(W_i - W_j)/\sigma(W_i - W_j)$							Confidence Levels for the null hypothesis W on wins						
	6	7	8	9	10	11	12	6	7	8	9	10	11	12
6		1.117	5.937	6.561	11.327	14.803	16.618		86.8%	100.0%	100.0%	100.0%	100.0%	100.0%
7			4.814	5.437	10.189	13.648	15.452			100.0%	100.0%	100.0%	100.0%	100.0%
8				0.620	5.331	8.738	10.508				73.2%	100.0%	100.0%	100.0%
9					4.708	8.111	9.877					100.0%	100.0%	100.0%
10						3.382	5.131						100.0%	100.0%
11							1.745							95.9%

Table 3. $(W_i - W_j)$ in units of $\sigma(W_i - W_j)$ and confidence levels for the null hypothesis.

Again, the two match sequences $\{M_6, M_8, M_{10}, M_{12}\}$ and $\{M_7, M_9, M_{11}\}$ demonstrate the null hypothesis with high levels of confidence, as do $\{M_7, M_8\}$ and $\{M_9, M_{10}, M_{11}, M_{12}\}$.

	(D$_i$ – D$_j$)/σ(D$_i$ – D$_j$)							Confidence Levels for the null hypothesis D on draws						
i \ j	6	7	8	9	10	11	12	6	7	8	9	10	11	12
6		-4.151	-8.059	-8.669	-14.402	-16.387	-18.204		100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
7			-3.889	-4.495	-10.158	-12.108	-13.888			100.0%	100.0%	100.0%	100.0%	100.0%
8				-0.604	-6.230	-8.158	-9.913				72.7%	100.0%	100.0%	100.0%
9					-5.623	-7.547	-9.299					100.0%	100.0%	100.0%
10						-1.911	-3.645						97.2%	100.0%
11							-1.732							95.8%

Table 4. (D$_i$ – D$_j$) in units of σ(D$_i$ – D$_j$) and confidence levels for the null hypothesis.

For a third time, the two match sequences {M_6, M_8, M_{10}, M_{12}} and {M_7, M_9, M_{11}} demonstrate with high levels of confidence, the null hypothesis that the rate of draws increases with p. The sequences {M_7, M_8} and {M_9, M_{10}, M_{11}, M_{12}} are equally encouraging.

5.   SUMMARY

This note has revisited Heinz' comprehensive self-play experiment in chess and re-assessed the significance of his results for the null hypothesis "an increment of search-depth yields decreasing returns". It appears that the confidence levels with which this may be stated are better than as suggested in his original paper.

Ancillary null hypotheses, on the decreasing frequency of wins and the increasing frequency of draws, were also suggested by the data and hold with good levels of confidence.

Some open questions that arose during this investigation:

- to what extent are the characteristics of the chosen engine, FRITZ, affecting the experimental results?
- why is there no comparable trend in l_p, the underlying mean of L_p?
- given a possible stm/sntm position evaluation bias, should M_s be compared with M_{s-2} rather than M_{s-1}?
- does the one ply search-depth advantage initially deliver increasing benefits with depth, c.f. M_6 and M_7?
- are there other perhaps increasing-return trends in shallow-search matches, e.g., M_2-M_7?
- as $p \to \infty$, does s_p decrease to a limit greater, equal to or less than 0.5?

The apparently anomalous M_6/M_7 results suggest further investigation at the conveniently shallow end of the experiment. Were Heinz' experimental results here a 'one off' or does the one-ply advantage increasingly overcome random play at small depths of search?

Thanks goes to Ernst Heinz for his comprehensive self-play experiment and results.

6.   REFERENCES

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. 3rd Edition. J. Wiley and Sons: Series in Probability and Mathematical Statistics. ISBN 0-4712-5708-7.

Heinz, E.A. (2000a). *Scalable Search in Computer Chess*. Vieweg Verlag. ISBN 3-5280-5732-7, esp. p. 178.

Heinz, E.A. (2000b). *A New Self-Play Experiment in Computer Chess*. Technical Memo No. 608 (MIT-LCS-TM-608). Laboratory or Computer Science, Massachussetts Institute of Technology.

Heinz, E.A. (2001a). New Self-play Results in Computer Chess. *Proceedings of the 2nd International Conference on Computers and Games* (eds. I. Frank and T.A. Marsland), pp. 267-280. Lecture Notes in Computer Science (eds. G. Goos, J. Harmanis and J. van Leeuwen), Springer Verlag. ISBN 3-540-43080-6.

Heinz, E.A. (2001b). Self-play, Deep Search and Diminishing Returns. *ICGA Journal*, Vol. 24, No. 2, pp. 75-79. ISSN 1389-6911.