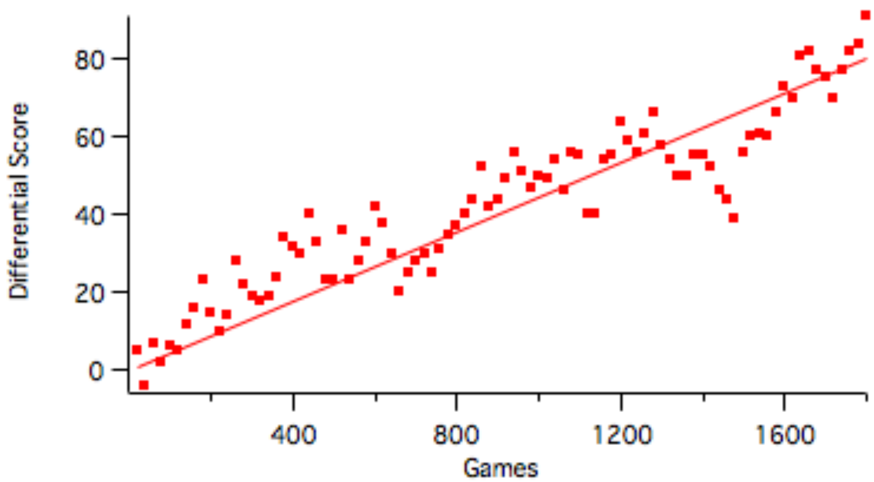
Experiment: 1800 games were played in unlimited mode between Gnu_bg and a human player. The human player (myself) played the first 20 games then all the remaining games were played using the “end game” option (basically the algorithm played against itself).



	Gnu_bg	Human	t-test. 1; 2 tails; n
Total wins	923	877	
Total score	1187	1085	
1 pt. wins	678	669	
gammon wins	254	208	
1 pt. wins/100 games	37.17±3.94	36.5±5.83	0.34; 0.069; n=18
gammon wins/100 games	14.11±3.01	11.56±3.84	0.016; 0.033; n=18

Legend

Differential score = Points of Gnu_bg – Points of human

1 pt. wins/100 games = mean number of 1 points wins per 100 games (mean ± standard deviation). This is equal to the probability to win a gammon that is gammon wins/1800
gammon wins/100 games = see above

t-test = t-test results between 1 pt. wins and gammon wins with 1-tail (appropriate in this case since we are looking for bias) and 2-tail (not necessary but indicative anyway).

Data fitted to y = ax where the coefficient a and one standard deviation are = 4.48±0.09 points per 100 games.

There were 10 backgammons each. These have not included in the calculations.

The difference between the probabilities to win a gammon is 2.55%. With this error, after 1800 games the level of confidence is 98%.