GLI®

GLI Europe BV
Diakenhuisweg 29-35
2033 AP Haarlem
The Netherlands

Tel +31 (0)88 220 6600
www.gaminglabs.com

Chamber of Commerce
Leiden nr. 28117769
VAT Identification number
NL 8184.73.393.B.01

RvA Registration Number of Accreditation applicable to this Report	
Testing	L372

Issuing Laboratory	GLI Europe B.V.	**Evaluating Laboratory**	GLI Europe B.V.
Recipient	AI Factory Limited 23 The Avenue Hatch End, Middlesex HA5 4EN United Kingdom		
Tested against Requirements	N/A		
Jurisdiction	Non-jurisdictional		
Manufacturer	AI Factory Limited 23 The Avenue Hatch End, Middlesex HA5 4EN United Kingdom		
Submitter	AI Factory Limited 23 The Avenue Hatch End, Middlesex HA5 4EN United Kingdom		
Product Name	**Backgammon RNG v2.241**		
Description of the Product Tested	**uk.co.aifactory.backgammonfree.apk** As requested per manufacturer's letter received on 29th August 2017.		
Evaluation Period	22nd September 2017 / 23rd January 2018		
Internal Reference:	RN-246-AFL-17-01-000		
Result	Pass		
Internal methods used reference	Random Number Generator (RNG) Analysis WI-MA-006		

Technical Evaluation authorized by:

Martin Britton
Managing Director

Worldwide Locations

World Headquarters
Lakewood, New Jersey

International Offices
GLI Africa
GLI Asia
GLI Australia Pty Ltd
GLI Austria GmbH
GLI Europe BV
GLI Italy
GLI South America

U.S. Regional Offices
Colorado
Nevada

This Revised Report replaces Evaluation Report 2018AFL001RNG246 dated 25th January 2018. This Revised Report was issued in order to meet the client's updates requests regarding the phrasing in Sections I, II, and IV.

RANDOMNESS REPORT FOR THE AI FACTORY LIMITED BACKGAMMON RNG

The intent of this report is to indicate that **Gaming Laboratories International**, **LLC** (GLI) has completed its evaluation of the Backgammon random number generator (RNG), version 2.241, provided by AI Factory Limited. This version was published on October 5, 2017 and in service as of January 2018.

SECTION I – SCOPE OF TESTING

AI Factory Limited submitted the required materials to GLI in order to conduct a random number generator analysis on the Backgammon RNG. An extensive source code review and analysis was done to inspect for possible manipulation of the RNG state and results in such a way that would give the AI an unfair advantage over the player (i.e. 'cheating'). The scope of this analysis was limited to software verification, source code review, and data analysis. The RNG was tested for its ability to randomly produce outcomes for the Backgammon games.

SECTION II – SOFTWARE VERIFICATION

Verify+ by Kobetron™ signatures for the Backgammon RNG are as follows:

File	Version	Type	Signature
uk.co.aifactory.backgammonfree.apk	2.241	MD5	F2051C6D379E1F3AE3E629597FF9B6D2
		SHA-1	B3DCDB7C37F22D1EE57328B8EC7000668628CB93
		KOBE4	1PC9

Table 1. Digital Signatures

GLI conducted an independent build of the source code that functionally matched the release version provided by AI Factory Ltd.

SECTION III – SOURCE CODE REVIEW

AI Factory Limited submitted appropriate documentation and full source code which pertains to the generation of random numbers. GLI reviewed the source code provided by tracing the path of the RNG application from the initiation of the draw to the selected output of random numbers. GLI inspected the source code, where practicable, in an attempt to find any undisclosed switches or parameters having a possible influence on randomness and fair play. GLI assessed the ability of the RNG to produce all numbers within the desired range.

SECTION IV – DATA ANALYSIS

The game configuration and parameters for the data obtained and tested are listed in Table 2. GLI performed a data format check on each data set listed in order to confirm that the game parameters were correctly represented in the data analyzed.

A set of numbers is said to be drawn *with replacement* if a number can be selected multiple times within the same draw. A set of numbers is said to be drawn *without replacement* if a number can only be selected once within the same draw.

Data Set	Range	Positions	Replacement	Draws
Data Set 1	0 - 5	2	Yes	5,000,000
Data Set 2	0 - 1	1	N/A	1,000,000
Data Set 3	0 - 99	1	N/A	100,000,000
Data Set 4	0 - 100	1	N/A	100,000,000
Data Set 5	0 - 9999	1	N/A	2,000,000,000

Table 2. Game Parameters

For a summary of the statistical tests applied to each data set, see *Appendix A*. For a description of the overall test methodology and a description of each test used, see *Appendix B*.

Overall, the RNG passed the battery of tests for each configuration at the 95%, 98%, and 99% confidence levels (within 1, 2 and 3 standard deviations).

SECTION V — GAME USAGE

Additional steps were taken to ensure the RNG is not called inappropriately when evaluating moves to be made within the Backgammon game by the AI (Artificial Intelligence) to gain an unfair advantage over the opponent. Detailed analysis and inspection during emulation of various games did not reveal the AI to gain access to knowledge of previous or future dice rolls. Furthermore, no attempts were made to predict, anticipate, or modify the behavior of the RNG which could be considered cheating.

SECTION VI - SUMMARY

Overall Evaluation of the Random Number Generator

GLI's conclusion based upon the tests applied to the AI Factory Limited Backgammon RNG data is that this random number generator has exhibited random behavior and is suitable for the applications as described herein. No player is deliberately given a more or less likely advantageous position over the other in any particular game by the RNG or its usage. If a game utilizes a different range or a different number of selections from the included ranges, the RNG should be resubmitted to test that set of parameters.

APPENDIX A: Statistical Test Summary

Data Set	Range	Positions	Replacement	Draws	Runs	Serial Corr.	Interplay Corr.	Adj. Max-Min	Coupon	Duplicates	Overlaps	Tot. Dist.	Tot. Dist. by Pos.
					Test Names								
Data Set 1	0 - 5	2	Yes	5,000,000	X	X	X	X	X	X	X	X	X
Data Set 2	0 - 1	1	N/A	1,000,000	X	X			X	X	X	X	
Data Set 3	0 - 99	1	N/A	100,000,000	X	X			X	X	X	X	
Data Set 4	0 - 100	1	N/A	100,000,000	X	X			X	X	X	X	
Data Set 5	0 - 9999	1	N/A	2,000,000,000	X	X			X	X	X	X	

Table A 1. Tests Applied

APPENDIX B: Test Descriptions

B.1 **Definitions.** The following terms apply to the below test descriptions. Randomness Device or Random Number Generator (RNG) output may be collected multiple numbers at a time. Each set of numbers is called a draw. Each individual number has a particular order within the *draw*. This is referred to as the number *position*.

B.2 **Distribution Comparisons.** Many of the tests compare an observed numerical distribution with an expected distribution. Unless otherwise specified, this is done by means of a statistical chi-square goodness-of-fit test. The value chi-square is computed in the standard way. If k is a possible value, o_k is the observed count of that value, and e_k is the expected count:

$$\chi^2 = \sum_k \frac{(o_k - e_k)^2}{e_k}$$

In the case where expected counts are too small for accurate use of the above formula, values are 'binned' together to ensure an appropriate minimum expected count. The resultant value for chi-square is compared against the distribution for the appropriate number of degrees of freedom. Unusually high (distribution mismatch) or unusually low (insufficient randomness) chi-square values can be causes for data failure.

B.3 **Meta-testing**. Evaluation of groups of p-values may include a meta-test for extremity of high or low p-values, a meta-test for frequency of high or low p-values, and a meta-test for uniformity of p-values, as appropriate.

B.4 **Confidence Level**. The statistical tests conducted by GLI are done at a particular *confidence level*. Common confidence levels used include 95%, 98%, and 99%, depending on jurisdictional requirements, and intended use of the RNG. High confidence level testing has low risk of mistakenly failing a good RNG, but higher risk of passing a bad RNG. Lower confidence level testing has increased power of detecting bad RNGs, while also increasing the risk of false failures of good RNGs. Specifically, the confidence level represents the probability that an ideal source of randomness would pass the testing. If an RNG passes statistical tests at a given confidence level, passage at all *higher* confidence levels is implied.

B.5 **Tests**. Some tests are only applicable to certain types of data. Some tests may be applied only to a portion of the data. Some tests may require that the data be parsed, binned, or otherwise transformed, as necessitated by data format.

Adjacency Max-Min:

For each draw, the difference between the maximum and minimum values is calculated and recorded. This is compared with the expected theoretical distribution. For example, if a draw consists of the numbers

2, 3, 6, 7, 4

the resulting statistic would be 5, the difference between the maximum value (7) and the minimum value (2).

Coupon Collector's:

The Coupon Collector's Test is applied positionally. The data is parsed until all possible values have been observed, then the number of values checked is recorded and the count is restarted. This is compared with the expected distribution. For example, if the set of all possible values is {0, 1, 2} and the first position of each draw is

1, 0, 1, 0, 2, 0, 1, 2, …

then all values are observed in the first position by the fifth draw. All values are then observed within the next 3 draws, so the first two statistics for the first position would be 5 and 3.

Duplicates:

The Duplicates Test counts the number of times a draw is exactly duplicated in the data. In the case that a particular draw is repeated more than twice, every possible way to generate a duplicate is counted. This is compared against the theoretical distribution to verify that the number of duplicate draws falls within expected bounds. For example, consider the dataset consisting of the following draws of two numbers each.

a) 1, 3
b) 4, 1
c) 1, 3
d) 1, 3
e) 4, 1
f) 3, 1

The duplicate pairs are (*a*, *c*), (*a*, *d*), (*c*, *d*), and (*b*, *e*), for a total of 4 duplicates. *(f)* is not counted as a duplicate since the draw must match in order as well as values.

Interplay Correlation:

The Interplay Correlation Test measures statistical correlation between different positions of the same draw. For each pair of positions, statistical correlation is calculated as in the Serial Correlation Test. In the case of without replacement data, an adjustment is made to account for the expected resulting negative correlation.

Overlaps:

The Overlaps Test compares consecutive draws for overlapping values. The number of overlapping values is recorded for each pair of draws. This observed distribution of overlaps is then compared against the expected distribution. For example, if the following draws are observed consecutively,

a) 1, 4, 5, 6
b) 4, 1, 7, 6

the number of overlaps would be 3, representing the values 1, 4, and 6.

Runs:

The Wald-Wolfowitz Runs Test is applied to each position within the draw. A center is established, typically the data median, and the number of ‘runs’ above and below the center are tallied. Values exactly equal to the center are discarded. This is compared to the expected distribution, which depends on the number of values above and below the center. For example, if the numbers drawn at a particular position were

$$2, 3, 1, 5, 4, 7, 3, 2, 3, 2, 3, 2, 6, 7, 3, 5$$

and the established center were the data median of 3, the data would be parsed for runs above 3 and runs below 3.

$$\underbrace{2, \mathbf{3}, 1}, \overbrace{5, 4, 7}, \mathbf{3}, \underbrace{2, \mathbf{3}, 2, \mathbf{3}, 2}, \overbrace{6, 7, \mathbf{3}, 5}$$

This would be counted as 4 runs.

Serial Correlation:

The Serial Correlation Test measures statistical correlation between consecutive draws of the same position. For each position, the sample Pearson correlation coefficient is calculated. If X represents the first number, and Y the number that follows, then the coefficient is

$$r = \frac{cov(X, Y)}{s_X s_Y}$$

where s denotes the sample standard deviation. The coefficients are used to generate a p-value for each position.

Total Distribution:

The Total Distribution Test is a simple tally of all observed values throughout the data. This is compared with the expected distribution. Typically the expected distribution is a uniform distribution. In the case of unequal weighting of values, an appropriate discrete distribution is used.

Total Distribution by Position:

The Total Distribution by Position Test tallies the observed distribution of values for each position within the draw. Each of these distributions is then compared with the expected.