Investigations of Game of Life
Cellular Automata rules on Penrose Tilings:
lifetime, ash, and oscillator statistics

Nick Owens[1], Susan Stepney[2]

[1] Department of Electronics, University of York, UK
[2] Department of Computer Science, University of York, UK

Abstract. Conway's Game of Life (GoL) rules can be applied to Cellular Automata (CAs) running on aperiodic grids, namely Penrose tilings. Here we investigate the result of running such CAs from random initial conditions. We describe our experimental setup, and demonstrate that the GoL on the Penrose kite and dart tiling has significantly different statistical behaviour from that on the Penrose rhomb tiling.

1 Introduction

John Horton Conway's *Game of Life* [2][4] is a simple two-dimensional, two state cellular automaton (CA), remarkable for its complex behaviour [2][11]. That behaviour is known to be very sensitive to a change in the CA *rules*. Here we continue our investigations [6][9] into its sensitivity to changes in the *lattice*, by the use of an aperiodic Penrose tiling lattice [5][10].

Section 2 generalises the concepts of neighbourhood and Game of Life rules to Penrose lattices. Section 3 describes the experimental setup for running the Game of Life rules on aperiodic lattices; section 4 reports the statistics of lifetimes, ash densities, and growth of the region of activity; section 5 reports the oscillator statistics.

2 Game of Life rules on Penrose lattices

Classic cellular automata are defined on regular lattices. The update rule depends on the state of the surrounding cells, and the updating cell itself, and the structure of that surrounding neighbourhood is invariant: all places in the lattice look the same, and the update rule can be applied uniformly across the lattice.

In general, the update rule depends on the particular state of each separate neighbour. For *outer totalistic* CA rules such as the *Game of Life*, however, the next state of a cell depends only on its current state, and the number of neighbourhood cells[3] in certain states.

---

[3] The standard definition of CA 'neighbourhood' includes both the surrounding cells *and* the updating cell. Throughout this paper we use slightly different terminology (because we are referring to outer totalistic rules): by *neighbourhood* we mean only the surrounding cells.
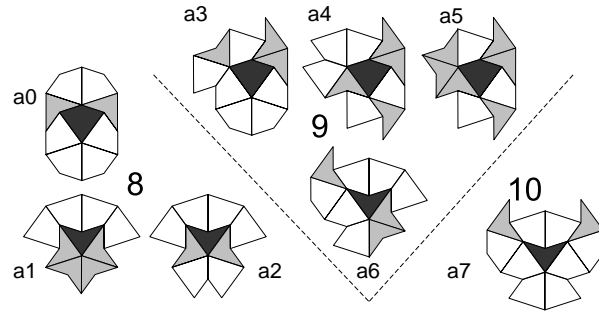
Fig. 1. The generalised Moore neighbourhoods on a kite and dart Penrose tiling, with neighbourhood sizes.

In Conway's *Game of Life* outer totalistic CA, the neighbourhood of each cell comprises the 8 nearest cells (of the regular Moore neighbourhood). Each cell has two states, 'dead' and 'alive'. If a cell is alive at time t, then it stays alive iff it has 2 or 3 live neighbours (otherwise it dies of 'loneliness' or 'overcrowding'). If a cell is dead at time t, then it becomes alive (is 'born') iff it has exactly 3 live neighbours.

For the aperiodic Penrose lattice, the detailed structure of the neighbourhood varies at different locations in the lattice. We define the generalised Moore neighbourhood of a cell to be all the cells with which it shares a vertex.

Not only do cells have irregular shaped neighbourhoods, with the generalised Moore neighbourhood not all cells have the same number of neighbours.

Figure 1 shows the eight distinct generalised Moore neighbourhoods in a kite and dart tiling: there are no other valid ways to surround a kite or a dart (this can be established by exhaustive consideration of the valid vertex configurations [9]). So there is one neighbourhood configuration of size 8 around a kite, and two around a dart; three of size 9 around a kite, and one around a dart; and one of size 10, around a dart. ([6] incorrectly states that kite and dart tilings have neighbourhoods of size 8 and 9 only.)

Similarly, figure 2 shows the 11 distinct generalised Moore neighbourhoods in a rhomb tiling. There is a larger range of distinct neighbourhood configurations for rhomb tilings.

Not all sizes of neighbourhoods appear with the same frequency. Figure 3 shows the distribution of neighbourhood sizes in a kite and dart tiling and in a rhomb tiling.

Outer totalistic rules can be given an interpretation in these aperiodic tiling neighbourhoods. Using our definition of the generalised Moore neighbourhood, the definition of the Game of Life on a regular lattice as given earlier can be used unchanged on a Penrose lattice. Some early investigations are reported in [6]; more detailed investigations are reported in [9] (including the lifetime and ash results reported here); better statistical tests, and further investigations (on oscillator numbers and periods) are additionally reported here.
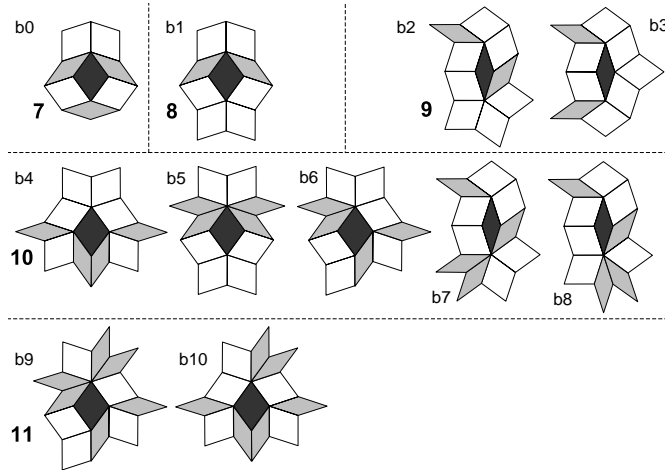
Fig. 2. The generalised Moore neighbourhoods on a rhomb Penrose tiling, with neighbourhood sizes.

3 Experimenting with Life

In our investigations, we use some typical GoL terminology, defined here. (The quoted definitions are from [12].)

soup "*A random initial pattern, often assumed to cover the whole Life universe.*" Here we consider only finite soup extents, but allow subsequent activity outside the initial soup patch.

quiescence Eventual periodic CA activity. Once the CA has entered a quiescent state, its future activity is periodic, and hence predictable.

ash "*The (stable or oscillating) debris left by a random reaction.*" Hence an ash is the quiescent state left by a soup.

In [6] we report that the Game of Life has different quantitative behaviour on a regular lattice and on a Penrose kite and dart lattice: on the Penrose lattice the lifetime to quiescence is much shorter, and the ash density is lower. This paper investigates if there are similar differences between the behaviour of the rules running on kite and dart and on rhomb lattices.

Null Hypothesis: The Game of Life run on kites and darts has identical statistical behaviour to the Game of Life run on rhombs.

To test this hypothesis, we investigate the following statistics: lifetime to quiescence, ash density, growth of the active area, number of oscillators in the ash, period of oscillators in the ash.

3.1 Experimental setup

To test the hypothesis we vary the density D of soups of similar sizes S on rhomb and kite and dart tilings, run the cellular automaton to quiescence, and record

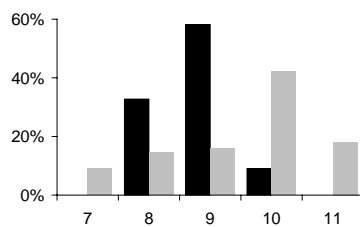| size	type	kite/dart cells	%	type	rhomb cells	%
7				b0	2831	9.1
					2831	9.1
8	a0	4994	14.7	b1	4576	14.6
	a1	4248	12.5			
	a2	1890	5.6			
		11132	32.7		4576	14.6
9	a3	6116	18.0	b2	2134	6.8
	a4	6125	18.0	b3	2842	9.1
	a5	3762	11.1			
	a6	3774	11.1			
		19777	58.2		4976	15.9
10	a7	3083	9.1	b4	2370	7.6
				b5	1735	5.6
				b6	2133	6.8
				b7	3475	11.1
				b8	3501	11.2
		3083	9.1		13214	42.3
11				b9	3522	11.3
				b10	2136	6.8
					5658	18.1

Fig. 3. Generalised Moore neighbourhood statistics, on a 33992 cell kite and dart tiling (black bars, median size = 9), and a 31255 cell rhomb tiling (grey bars, median size = 10)
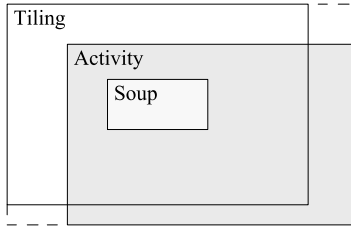
Fig. 4. The initial tiling grid G, the central soup area S, the maximum activity area during the run A, and the possibly extended tiling grid G_q (dashed box) to accommodate the activity. Our lazily expanding tiling algorithm is defined in [9].

the lifetime to quiescence t_q, ash density ρ (measured over the soup box), and soup growth g.

Lifetime t_q : The lifetime, or the time to quiescence, is defined to be the number of timesteps from the soup state ($t = 1$) until the pattern of live cells (measured over the whole tiling G_q, (figure 4) first repeats (at $t = t_q$). Each timestep, the CA's current state is stored, along with the number of live cells. To check for quiescence, the current state is compared to all previous states with the same number of live cells.

Ash period p : The number of timesteps since the first repeating state was previously seen: $p = t_q - t_{prev}$. The ash period may be 1, in which case there are no oscillators in the ash, only "still life" constructs with trivial period 1. Any oscillators in the ash have a period p_{osc} that is a factor of the ash period p.

Ash density ρ : The proportion of live cells in the ash at $t = t_q$, measured over the soup tiles S.

Soup growth g : The number of cells in the maximum active area divided by the number of cells in the soup: $g = A/S$ (figure 4).

Ash oscillator number density : The total number of oscillators of period 2 and above in the ash at $t = t_q$, divided by the number of cells in the active area A.

Tiling grid : We use a lazily expanding tiling for both kites and darts, and rhombs. We use an initial tiling of size $G = 23194$ for the kite and dart experiments, and of size $G = 23123$ for the rhomb experiments. It is difficult to produce identical sized tilings: these are close enough (within 0.3%) for fair comparison, since the statistical fluctuations in the results due to different random starting conditions are much greater than this difference in tiling sizes. These differences in tile numbers are of similar scale to the differences in tile numbers between regular and kite and dart tilings used in [6] (and the tilings are about twice the size of the largest grid explored there).

Soup area : Three initial soup areas S, covering the central 25%, 50%, 75% of the area of the tiling. See figure 5 and 6.

Soup density : 100 soup densities D, in the range $[0.01, 1.0]$ with increments of 0.01. Each cell in the soup area S is initially alive with probability D; all other cells in G are initially dead.
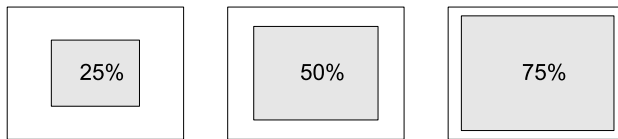
Fig. 5. The three central soup areas, to scale with the initial grid area.

	25%	50%	75%	G
kite and dart	5842	11670	17527	23194
rhomb	5815	11611	17405	23123

Fig. 6. Number of tiles involved in the experiments, soup sizes $S = 25\%$, 50% and 75%, and full initial grid size G

Runs : Each of the 100 soup densities D across the three soup sizes S is run to quiescence 1000 times from different random soup starting conditions.

3.2 Non-parametric statistical analysis

We want to test whether certain distributions are statistically the same or different.

The commonly-used statistical tests (such as Student's t-test) are *parametric* tests that assume an underlying normal distribution. Are the distributions here (sufficiently) normal[4] to justify using such parametric tests?

Figures 7 and 8 show the histograms of lifetime and ash density results over the 1000 runs for one particular soup size and soup density. The lifetime distributions, at least, do not look normal.

We investigate further the distribution of lifetimes and ash densities for these examples. We calculate the median, mean, standard deviation, skew and kurtosis of these distributions (using the MS-Excel functions MEDIAN, AVERAGE, STDEV, SKEW, and KURT respectively), for the lifetimes (figure 9) and the ash densities (figure 10).

For large samples ($N > 150$) drawn from a normal population, the skewness statistic is approximately normally distributed with mean 0 and standard deviation $s_s = \sqrt{6/N}$ [13, §5.13]; for very large samples ($N > 1000$) drawn from a normal population, the kurtosis statistic is approximately normally distributed with mean 0 and standard deviation $s_k = \sqrt{24/N}$ [13, §5.14]. Hence skew values beyond two standard errors of skewness, or kurtosis values beyond two standard errors of kurtosis, indicate that the distribution is not normal at the 95% confidence level.

For $N = 1000$ (just valid for the kurtosis test), $2s_s = 0.15$ and $2s_k = 0.31$. Both these values are considerably lower than those calculated for the lifetimes

---

[4] Use of a parametric test on a non-normal distribution can give misleading results. A good discussion of the importance of not using an inappropriate parametric statistical test can be found in [7].
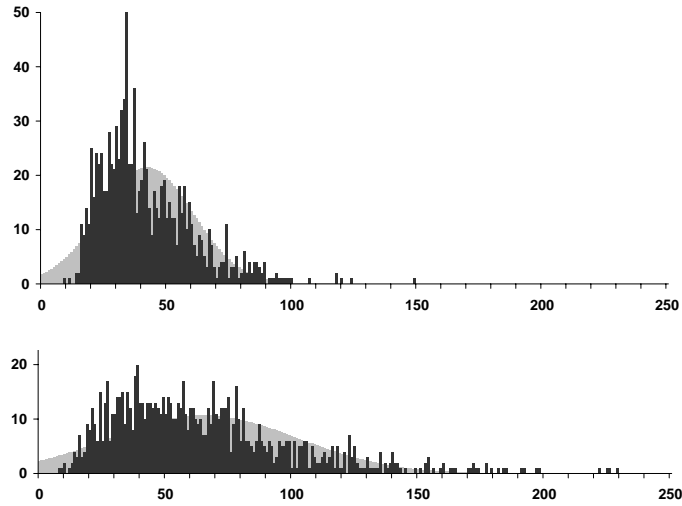
Fig. 7. The distribution of lifetimes to quiescence on the kite and dart tiling (top) and rhomb tiling (bottom), for 1000 runs with soup size $S = 25\%$ and soup density $D = 0.8$; with comparison normal distributions of the same mean and standard deviation.

Fig. 8. The distribution of ash densities on the kite and dart tiling (top) and rhomb tiling (bottom), for 1000 runs with soup size $S = 25\%$ and soup density $D = 0.8$; with comparison normal distributions of the same mean and standard deviation.
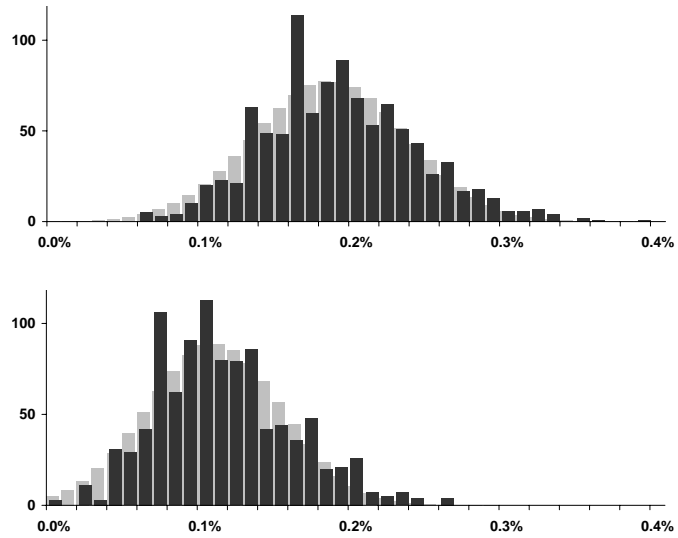
soup		D = 0.4		D = 0.8	
		k&d	rhomb	k&d	rhomb
25%	m	96	158	37	57.5
	μ	99.4	163.0	41.9	65.1
	σ	19.6	37.0	18.6	37.3
	s	1.1	0.9	1.2	1.5
	k	2.0	1.2	2.1	3.9
50%	m	108	179	40	60
	μ	111	185.1	44.7	66.6
	σ	19.6	37.2	18.5	33.7
	s	0.7	0.8	1.2	1.2
	k	0.9	0.5	2.0	2.1
75%	m	116	190	44	67
	μ	118.6	198.1	47.1	74.1
	σ	20.0	40.4	17.9	35.9
	s	0.9	1.2	1.3	1.2
	k	1.1	2.8	4.3	4.3

Fig. 9. Statistics for the lifetime distributions (median m, mean μ, standard deviation σ, skew s, kurtosis k) for soup densities $D = 0.4$ and 0.8; soup sizes $S = 25\%$, 50% and 75%. All the skewness and kurtosis values rule out normality at the 95% confidence level.

soup		D = 0.4		D = 0.8	
		k&d	rhomb	k&d	rhomb
25%	m	0.00440	0.00337	0.00181	0.00104
	μ	0.00442	0.00343	0.00184	0.00107
	σ	0.00083	0.00078	0.00051	0.00045
	s	**0.19**	**0.26**	**0.35**	**0.49**
	k	0.19	−0.02	0.19	0.10
50%	m	0.00845	0.00627	0.00220	0.00151
	μ	0.00842	0.00629	0.00224	0.00154
	σ	0.00112	0.00103	0.00058	0.00051
	s	−0.07	0.09	**0.23**	**0.30**
	k	0.14	−0.04	−0.14	0.03
75%	m	0.01233	0.00908	0.00285	0.00234
	μ	0.01234	0.00912	0.00288	0.00234
	σ	0.00136	0.00126	0.00067	0.00056
	s	−0.08	**0.19**	**0.20**	0.13
	k	−0.12	−0.10	0.04	−0.14

Fig. 10. Statistics for the ash densities (median m, mean μ, standard deviation σ, skew s, kurtosis k) for soup densities $D = 0.4$ and 0.8; soup sizes $S = 25\%$, 50% and 75%. Skewness values in bold rule out normality at the 95% confidence level.
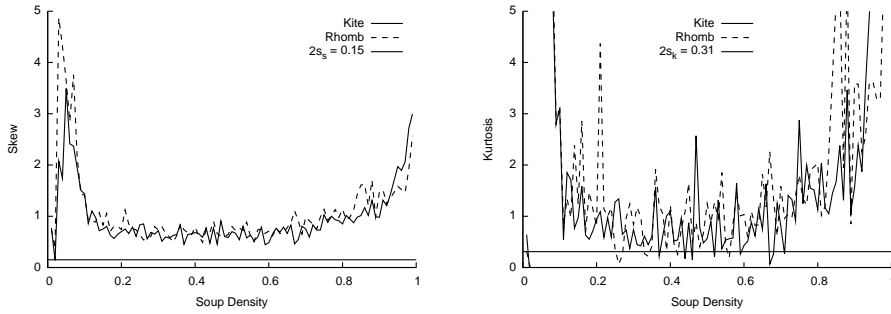
Fig. 11. Soup growth g: skew (left) and kurtosis (right) of the distributions for soup sizes 25%. The distributions clearly differ with soup density: they have a much higher skew and kurtosis at low and high soup densities than at intermediate ones. (Skew and kurtosis 95% confidence lines $2s_s = 0.15$ and $2s_k = 0.31$ are also shown.)

(figure 9), so the lifetime distributions are not normal at the 95% confidence level. Normality of the ash densities for some soup size and density parameters, but not all, is ruled out by this test (figure 10).

So, in most cases, we cannot justify using tests that assume normality of the distributions. Not only do we not know what the underlying distributions are, in some cases they can vary markedly across the parameter ranges explored (see figure 11). Since we wish to apply a uniform set of tests, we therefore use the same non-parametric (or distribution-free) tests in all cases.

The distributions are, in at least some cases, highly skewed, so the non-parametric median and quartile statistics give a better measure of the data than means and standard deviations. So we calculate the median and quartiles of each set of runs, for the range of soup densities (figures 13 and 14). These results are in qualitative agreement with those in [6]: low lifetimes and ash densities at extreme soup densities with a 'plateau' in the behaviours for soup densities $\sim 0.2 - 0.6$; lifetimes $\sim 100 - 200$; ash densities $\sim 1 - 2\%$. We now, however, have better statistics for kite and dart tilings, and new results for rhomb tilings.

3.3 Statistical significance

We use the Wilcoxon rank-sum test (also known as the Mann-Whitney U test) to test whether the medians of two samples are significantly different, using Matlab's `ranksum` function.

Null Hypothesis T: for soup size $S = 25\%$, density $D = 0.8$, there is no difference between the median lifetimes on kites and darts, and on rhombs.

If the p-value given by the rank-sum test is less than 0.05, we can say that the null hypothesis is rejected at the 95% confidence level because the medians differ. Running the test on our data for this case gives $p = 10^{-56}$, or essentially zero. Therefore we can reject *Null Hypothesis T*, with an *extremely* high degree of statistical confidence.

3.4 Effect size

The difference in the medians in the above test case is statistically significant to an almost ludicrous degree. This extreme level of statistical confidence is due mostly to the large number of samples, $N = 1000$. (Such large samples are much more typical in computer science than, say, medicine, because computer experiments are relatively cheap, and have no ethical considerations.) As Bakan says ([1, ch.1, p.7], as quoted in [8]): "there is really no good reason to expect the null hypothesis to be true in any population". A sufficiently large sample size will always be able to reject a null hypothesis: the smaller the effect, the larger the sample required to detect it. For normally-distributed populations with means and standard deviations similar to those of figure 10, sample sizes in the low tens would be sufficient to establish a statistically significant difference of their means at the 99% confidence level.

Because of this, we also perform a test of the *effect size*. The best know effect size measure is Cohen's d-test [3, §2.5]. It (roughly) measures $(\mu_1 - \mu_2)/\sigma$, the difference in the means, divided by the standard deviation (with corrections if the populations have different σs). This essentially captures how "far apart" the peaks of the two distribution are, compared to how much they spread (a physical analogy from optics might be how easily the two peaks could be "resolved").

However, Cohen's d is a parametric measure, assuming a normal distribution. Here we use the Vargha-Delaney non-parametric "measure of stochastic superiority", A [14]. A_{XY} measures the probability that a value selected at random from population X is greater than a value selected at random from population Y (with a correction to take into account the chance that they are equal). These probabilities could be estimated directly from the data samples, but [14, eqn(14)] shows that an unbiased estimator of A_{XY} is

$$\hat{A}_{XY} = \left(R_X/n_X - (n_X + 1)/2 \right)/n_Y \qquad (1)$$

where R_X is the rank sum value of X (also calculated by the Matlab `ranksum` function), and n_X, n_Y are the sizes of the samples drawn from populations X and Y.

$A = 0.5$ indicates that the populations have the same medians (the rank-sum test, above, can be used to measure the statistical significance of this). The further A is away from 0.5, the bigger is the *effect size*, analogous to Cohen's d. For the one tailed test (the probability that X is greater than Y) the effect size criterion is that $A = 0.56$ indicates a small effect, $A = 0.64$ a medium effect, and $A = 0.71$ a large effect ([14, table 1]; these values are only indicative, since they are derived directly from Cohen's criteria for the d values of different effect sizes, and those values are themselves only indicative). For the two tailed test, we test both A_{XY} and A_{YX} (and use the fact that $A_{YX} = 1 - A_{XY}$). So, for example, for a large effect in the two-tailed case, we test whether $A > 0.71$ or $A < 0.29$.

For our test soup sizes and densities, the effect size A values are shown in figure 12. These indicate a large effect from the change in the tiling.

	lifetime		ash density	
soup	$D = 0.4$	$D = 0.8$	$D = 0.4$	$D = 0.8$
25%	0.955	0.704	0.807	0.865
50%	0.975	0.711	0.918	0.809
75%	0.980	0.746	0.957	0.724

Fig. 12. Effect size A, for the difference between kite & dart and rhomb medians, for soup densities $D = 0.4$ and 0.8; soup sizes $S = 25\%$, 50% and 75%. All these values indicate large effects.

3.5 Null hypotheses

We cast our various null hypotheses in the following form:

Null Hypothesis X: for the Game of Life run on kites & darts, and on rhombs, there is no meaningful difference in property Y.

For a "meaningful difference" we require both a statistically significant difference, and a large effect size.

For all the results that follow, we do not present p value for the statistical significance: the p values are all essentially zero, and the differences are all extremely statistically significant. We present the value of the effect size A, demonstrating that all the statistics chosen exhibit a large effect with the change in the tiling.

4 Lifetime, ash, growth results

4.1 Lifetimes

Null Hypothesis 1: for the Game of Life run on kites & darts, and on rhombs, there is no meaningful difference in the lifetimes.

The lifetime distributions for the two tilings are different, with a large effect size for soup densities $0.1 < D < 0.8$ (figure 13), refuting Null Hypothesis 1. The Game of Life on the rhomb tiling has significantly longer lifetimes than it does on the kite and dart tiling. From [6], we can say that they both have shorter lifetimes than Life on a regular lattice.

4.2 Ash densities

Null Hypothesis 2: for the Game of Life run on kites & darts, and on rhombs, there is no meaningful difference in the ash densities.

The ash density distributions for the two tilings are different, with a large effect size for soup densities $0.1 < D < 0.9$ (figure 14), refuting Null Hypothesis 2. The Game of Life on the rhomb tiling has significantly lower ash densities than it does on the kite and dart tiling. From [6], we can say that they both have lower ash densities than Life on a regular lattice.
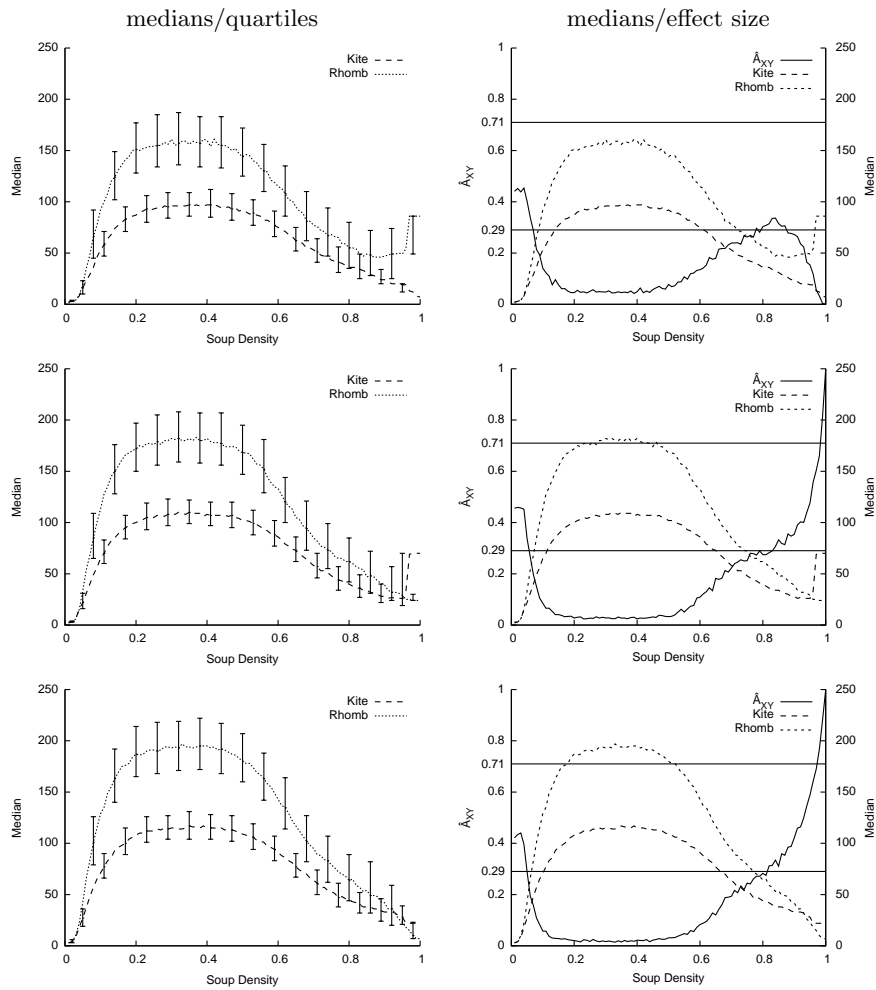
Fig. 13. Lifetime to quiescence t_q: medians and quartiles (left), and effect size (right); for soup sizes 25% top, 50% middle, 75% bottom. Effect sizes > 0.71 or < 0.29 are considered to be large.
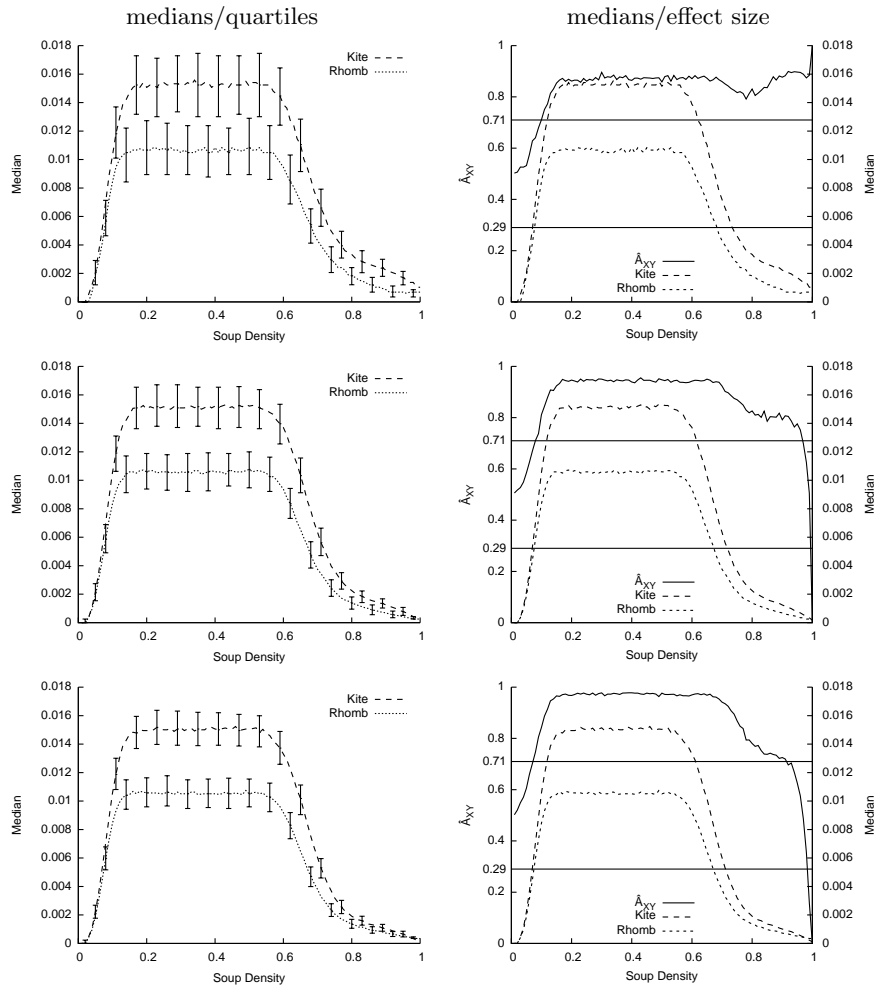
Fig. 14. Ash density ρ: medians and quartiles (left), and effect size (right); for soup sizes 25% top, 50% middle, 75% bottom.

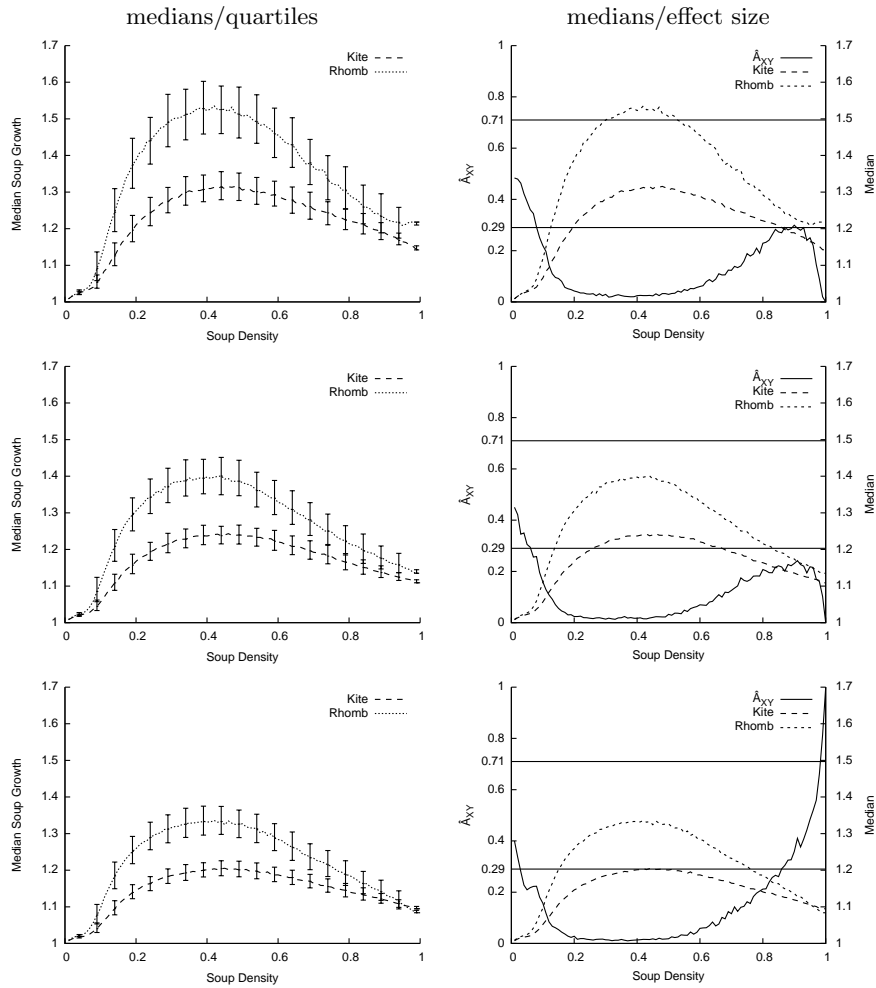**Fig. 15.** Soup growth g: medians and quartiles (left), and effect size (right); for soup sizes 25% top, 50% middle, 75% bottom.

4.3 Soup growth

Null Hypothesis 3: for the Game of Life run on kites & darts, and on rhombs, there is no meaningful difference in the soup growth.

The growths of the two tilings are different, with a large effect size for soup densities $0.1 < D < 0.8$ (figure 15), refuting Null Hypothesis 3. The Game of Life on the rhomb tiling has significantly more growth from soup than it does on the kite and dart tiling.

5 Oscillator results

5.1 Ash oscillator number density

Null Hypothesis 4: for the Game of Life run on kites & darts, and on rhombs, there is no meaningful difference in the ash oscillator number density.

The ash oscillator number densities for the two tilings are different, with a large effect size for soup densities $0.1 < D < 0.75$ (figure 16) refuting Null Hypothesis 4. The Game of Life on the rhomb tiling has significantly lower ash oscillator number density than it does on the kite and dart tiling.

5.2 Ash period statistics

Null Hypothesis 5: for the Game of Life run on kites & darts, and on rhombs, there is no meaningful difference in the ash periods.

The ash periods for the two tilings are different, with a large effect size for soup densities $0.1 < D < 0.7$ (figure 17) refuting Null Hypothesis 5. The Game of Life on the rhomb tiling has significantly lower ash period than it does on the kite and dart tiling.

Since the normality tests show that the data is highly skewed, with a fat tail, we also present "box and whisker" plots of the of the ash period distributions (figure 18), which show the outliers: high period ash very far from the median period. Since this is the overall period of the ash, large spikes can be caused when a few high period oscillators appear out of phase. The preponderance of kite/dart ash period outliers with periods 15, 30, and 60 are mostly caused by the common $p15$ *dancer* oscillator identified in [6] (figure 19). The total number of these ash period outliers (across all soup densities) is given in figure 20.

6 Conclusions

We have used a Penrose lazy tiling algorithm to perform experiments with Game of Life rules. The Game of Life on the rhomb tiling is significantly different from that on the kite and dart tiling: it has longer lifetimes, lower ash densities, higher soup growth, significantly fewer oscillators in the ash, and lower ash period.

Even the simplest oscillators seem quite difficult to define without explicit diagrams. Work is underway to develop a classification scheme for the oscillators.
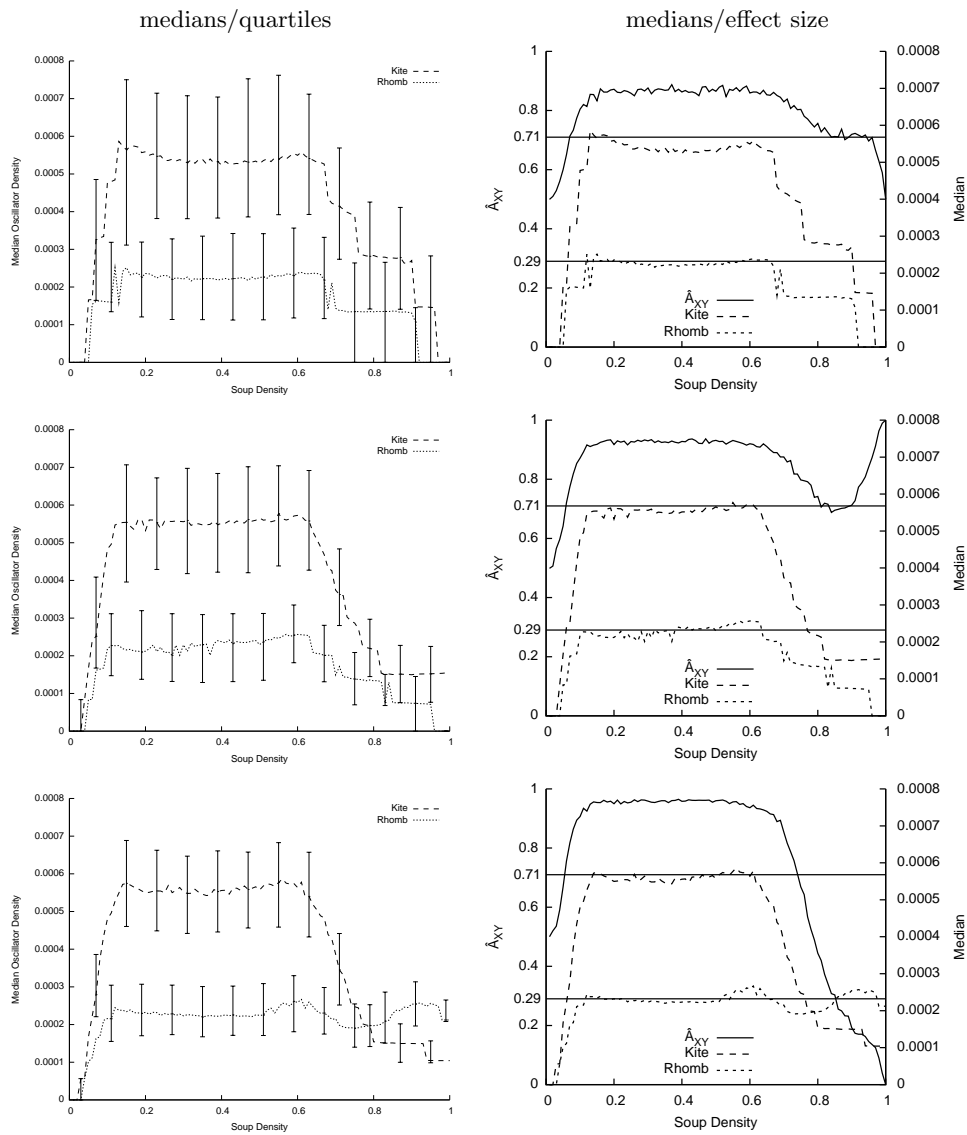
Fig. 16. Ash oscillator number density: medians and quartiles (left), and effect size (right); for soup sizes 25% top, 50% middle, 75% bottom

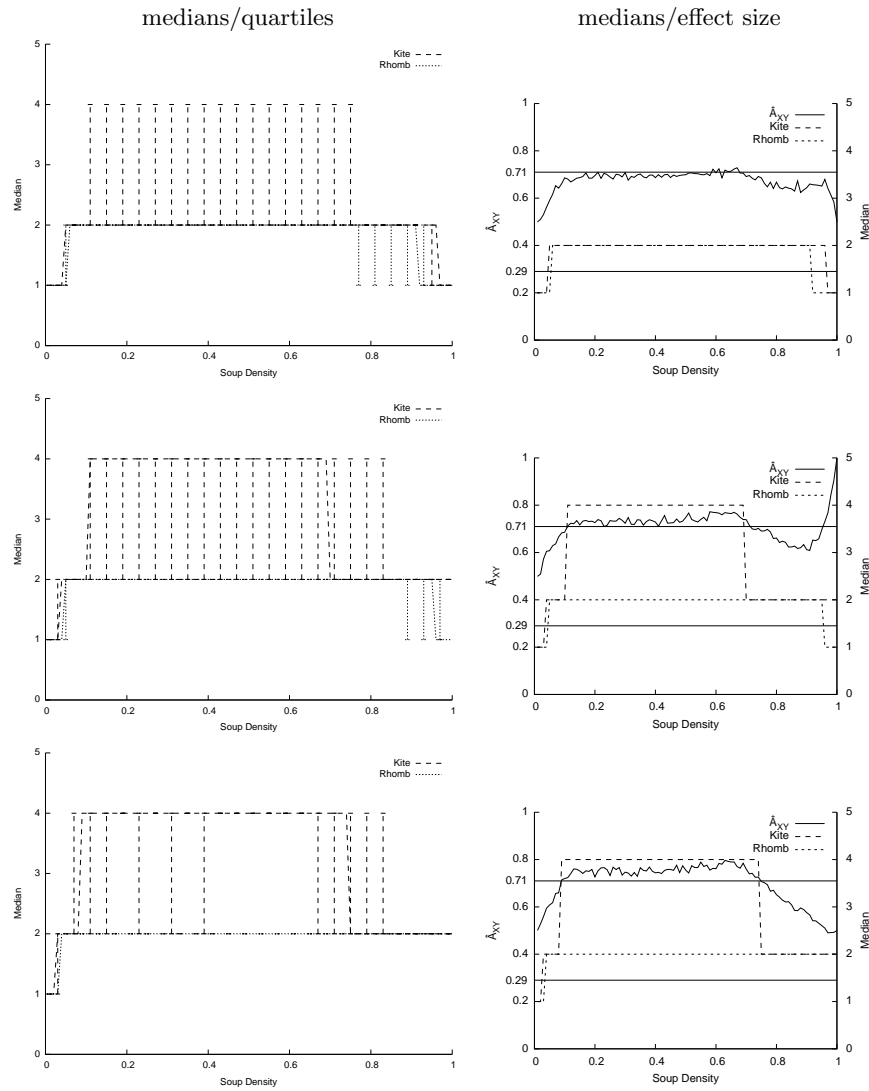medians/quartiles   medians/effect size

Fig. 17. Ash period: medians and quartiles (left), and effect size (right); for soup sizes 25% top, 50% middle, 75% bottom. (Note that the quartiles and median sometimes have the same value, in which case no quartile bar is drawn beyond a dash.)
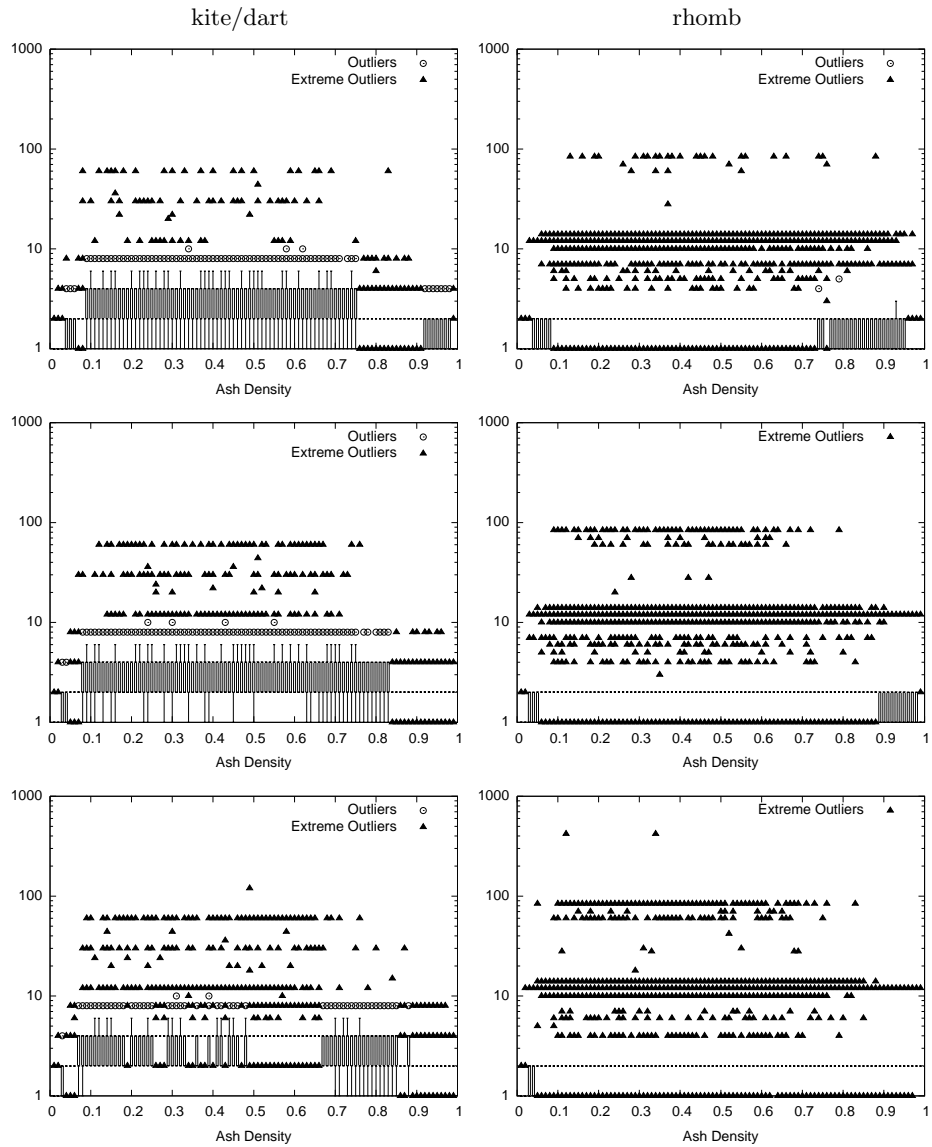
Fig. 18. Box and whisker plots of the ash period, showing median, quartiles, outliers (beyond 1.5 times interquartile range) and extreme outliers (beyond 3 times interquartile range); for soup sizes 25% top, 50% middle, 75% bottom. Note the logarithmic scale on the period axis.
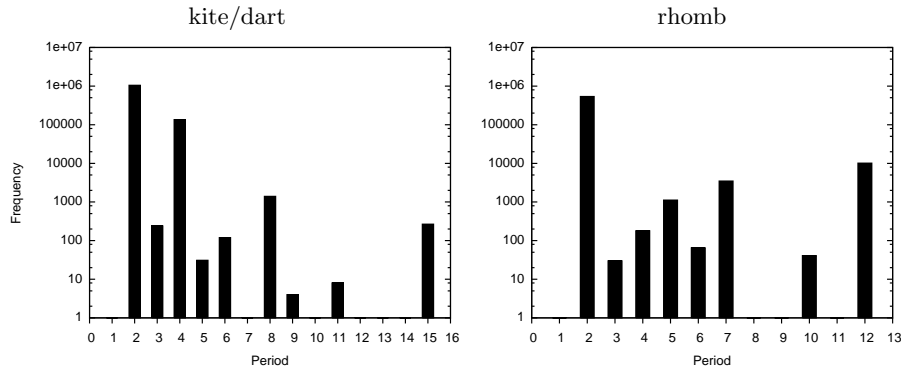
Fig. 19. Histograms of the periods of the individual oscillators in the ash. Note the relatively high number of $p15$ oscillators in the kite/dart ash.

	25%		50%		75%	
	outlier	extreme	outlier	extreme	outlier	extreme
kite and dart	844	6133	600	4362	591	10707
rhomb	2	14416	0	12035	0	10239

Fig. 20. Number of outliers and extreme outliers in the ash period distribution, soup sizes $S = 25\%$, 50% and 75%

Acknowledgements

References

1. D. Bakan. *On Method: Toward a Reconstruction of Psychological Investigation.* Josey-Bass, 1967.
2. Elwyn R. Berlekamp, John Horton Conway, and Richard K. Guy. *Winning Ways for Your Mathematical Plays Volume 2: games in particular.* Academic Press, 1982.
3. Jacob Cohen. *Statistical power analysis for the behavioral sciences.* Lawrence Erlbaum Associates, 2nd edition, 1988.
4. Martin Gardner. Mathematical games: The fantastic combinations of John Conway's new solitaire game "life". *Scientific American*, 223(4):120–123, October 1970.
5. Martin Gardner. Mathematical games: extraordinary non-periodic tiling that enriches the theory of tiles. *Scientific American*, 236(1):110–121, January 1977.
6. Margaret Hill, Susan Stepney, and Francis Wan. Penrose Life: ash and oscillators. In Mathieu S. Capcarrere, Alex A. Freitas, Peter J. Bentley, Colin G. Johnson, and Jonathan Timmis, editors, *Advances in Artificial Life: ECAL 2005, Canterbury, UK, September 2005*, volume 3630 of *LNAI*, pages 471–480. Springer, 2005.
7. Nancy L. Leech and Anthony J. Onwuegbuzie. A call for greater use of non-parametric statistics. In *Annual Meeting of the Mid-South Educational Research Association, Chatanooga, TN, USA*, November 2002.

8. Marks R. Nester. An applied statistician's creed. *Applied Statistics*, 45(4):401–410, 1996.

9. Nick Owens and Susan Stepney. Investigations of Game of Life cellular automata rules on Penrose tilings: lifetime and ash statistics. In *Automata 2008, Bristol, UK, June 2008*, pages 1–35. Luniver Press, 2008.

10. Roger Penrose. Pentaplexity. *Eureka*, 39:16–32, 1978.

11. Paul Rendell. Turing Universality of the Game of Life. In Andrew Adamatzky, editor, *Collision-Based Computing*. Springer, 2002.

12. Steven Silver. Life lexicon, release 25.
    `http://www.argentum.freeserve.co.uk/lex.htm`, February 2006.

13. George W. Snedecor and William G. Cochran. *Statistical Methods*. Iowa State University Press, 7th edition, 1980.

14. Andra Vargha and Harold D. Delaney. A critique and improvement of the "*CL*" common language effect size statistics of McGraw and Wong. *Journal of Educational and Behavioral Statistics*, 25(2):101–132, 2000.