Modelling and Simulation of Systems
Exercise 3: Pseudorandom number generators

Agata Radys, Paweł Cejrowski, Łukasz Myśliński

November 16, 2016

1 Generator 1

GFSR(generalized feedback shift register) with parameters: $j = 1$, $k = 2$, $m = 2^{32}$.

1.1 Equation

Numbers are generated according to equation 1.

$$x_n = x_{n-1} \oplus x_{n-2} \quad (mod\ 2^{32}) \tag{1}$$

where x_0 and x_1 are given.

1.2 Questions

- **What is the minimum(maximum) possible value of the period? Give an example of initial values for which the period is small(large).**
  Minimum value of period is equal 1, when both x_0 and x_1 are equal 0. Based on lecture, we know that maximum value of period is $(2^k - 1) \cdot k = (2^2 - 1) \cdot 2 = 6$, but this value of period was not found during tests. Found 3 when x_0 or x_1 was different from 0(e.g. $x_0 = 0$ and $x_1 = 16$).
- **What is the minimum(maximum) possible mean value? Give an example of initial values for which the average value is small(large).**
  Minimum possible mean value is equal 0, when both x_0 and x_1 are equal 0. Maximum found mean value within period is equal $\frac{2}{3} \cdot (2^{32} - 1) = 2863311530$, when $x_0 = x_1 = 2^{32} - 1$
- **What is the minimum(maximum) possible variance? Give an example of initial values for which the variance is small(large).**
  Minimum possible variance value is equal 0, when both x_0 and x_1 are equal 0. Maximum found variance value within period is equal $4.0992764589155E + 18$, when $x_0 = x_1 = 2^{32} - 1$
- **Does the generator meet the requirements that good generators should satisfy? If not, which of the requirements are not satisfied and why?**
  The generator does **not** meet all requirements for good generators.
  Characteristics of good generators:
  1. **generated numbers distributions are as close as possible to the desired one**
     Not satisfied: 3 point distribution.
  2. **subsequences of the produced sequence are mutually independent**
     Not satisfied: each subsequence of length 3 contains the same 3 numbers.
  3. **long period, with length at least $\sqrt{l}$, where l is the length of the used subsequence**
     Not satisfied: short period with maximum length of 3.
  4. **the ability to make jumps, i.e. to compute x_j from x_i for every $j > i$**
     Not satisfied: x_j cannot be computed from single number x_i, at least one other number x_k, where $k(mod\ 3) \neq i(mod\ 3)$ needed.
  5. **repeatable, portable and efficient**
     Satisfied.
- **Is the generator suitable for use in the cryptography? If not, why?**
  The generator is **not** suitable for use in the cryptography, because it does not satisfy any of the desired conditions.
  Characteristics of generators suitable for use in the cryptography:

1. **it must be impossible to predict its seed and internal state even if we have a large sample of the numbers it produced**
   Not satisfied: it is very easy to predict the number when the period is at most 3 numbers long.
2. **it must have a long period for every possible value of its seed**
   Not satisfied: do not have long period for any seed.
3. **it should be unpredictable to the public, i.e. the probability of predicting the subsequent numbers should be low even if have a large sample of the numbers it produced**
   Not satisfied: having 3 consecutive numbers one can predict any other.

2 Generator 2

LCG(linear congruential generator) with parameters $a = 3$, $c = -1$, $m = 2^{32}$.

2.1 Equation

Numbers are generated according to equation 2.

$$x_n = 3 \cdot x_{n-1} - 1 \pmod{2^{32}} \tag{2}$$

where x_0 is given.

2.2 Questions

Below corner cases were found by 10000 generator executions with randomized n, m and *seed* by shell `$RANDOM`, which generates pseudorandom number from range 0-32767.

- **What is the minimum(maximum) possible value of the period? Give an example of initial values for which the period is small(large).**
  Minimum value of period is equal $m - n + 1$ for $m < 100000$. Based on lecture, we know that maximum value of period is 2^30 and is reached when x_0 is odd.

- **What is the minimum(maximum) possible mean value? Give an example of initial values for which the average value is small(large).**
  Minimum mean value for longer ranges($m - n > 50$) is bigger than $1e9$. Maximum found mean value is 2373236302.0435, but still it is only empirical.

- **What is the minimum(maximum) possible variance? Give an example of initial values for which the variance is small(large).**
  Minimum and maximum variance value for longer ranges($m - n > 50$) is bigger than $1e18$.

- **Does the generator meet the requirements that good generators should satisfy? If not, which of the requirements are not satisfied and why?**
  The generator meets all requirements for good generators.
  Characteristics of good generators:

  1. **generated numbers distributions are as close as possible to the desired one**
     Satisfied: K^+ and K^- satisfied in 99.9% cases with the $\alpha = 0.05$. All satisfied with $\alpha = 0.15$.
  2. **subsequences of the produced sequence are mutually independent**
     Satisfied: chi-square always satisfied on the $\alpha = 0.05$.
  3. **long period, with length at least $\sqrt{l}$, where l is the length of the used subsequence**
     Satisfied: proportional to l.
  4. **the ability to make jumps, i.e. to compute x_j from x_i for every $j > i$**
     Satisfied.
  5. **repeatable, portable and efficient**
     Satisfied.

- **Is the generator suitable for use in the cryptography? If not, why?**
  The generator is **not** suitable for use in the cryptography, because it does not satisfy two of the desired conditions.
  Characteristics of generators suitable for use in the cryptography:

1. **it must be impossible to predict its seed and internal state even if we have a large sample of the numbers it produced**
   Not satisfied: it is possible to predict seed internal state based on subsequence.
2. **it must have a long period for every possible value of its seed**
   Satisfied.
3. **it should be unpredictable to the public, i.e. the probability of predicting the subsequent numbers should be low even if have a large sample of the numbers it produced**
   Not satisfied: the formula is easy to guess based on consecutive numbers.

3 Generator 3

3.1 Equation

Numbers are generated according to equation 3.

$$\begin{cases} x'_n = x'_{n-1} \oplus x'_{n-2} \pmod{2^{32}} \\ x''_n = 3 \cdot x''_{n-1} - 1 \pmod{2^{32}} \\ x_n = x'_n \cdot x''_n \pmod{2^{32}} \end{cases} \tag{3}$$

where x'_0, x'_1 and x''_0 are given.

3.2 Questions

Below corner cases were found by 1000 generator executions with randomized n, m and seeds x'_0, x'_1 and x''_0 by shell `$RANDOM`, which generates pseudorandom number from range 0-32767.

- **What is the minimum(maximum) possible value of the period? Give an example of initial values for which the period is small(large).**
  Value of period is proportional to $m - n + 1$ for $m - n < 100000$.

- **What is the minimum(maximum) possible mean value? Give an example of initial values for which the average value is small(large).**
  Minimum and maximum values were over 2000000 in tests.

- **What is the minimum(maximum) possible variance? Give an example of initial values for which the variance is small(large).**
  Minimum and maximum values were over $1.7e18$ in tests.

- **Does the generator meet the requirements that good generators should satisfy? If not, which of the requirements are not satisfied and why?**
  The generator does **not** meet all requirements for good generators.
  Characteristics of good generators:

  1. **generated numbers distributions are as close as possible to the desired one**
     Satisfied: K^+ and K^- satisfied in 99.9% cases with the $\alpha = 0.011$. All satisfied with $\alpha = 0.10$.
  2. **subsequences of the produced sequence are mutually independent**
     Satisfied: chi-square always satisfied on the $\alpha = 0.05$.
  3. **long period, with length at least $\sqrt{l}$, where l is the length of the used subsequence**
     Satisfied: proportional to l.
  4. **the ability to make jumps, i.e. to compute x_j from x_i for every $j > i$**
     Not satisfied: x_{i-1} also needed.
  5. **repeatable, portable and efficient**
     Satisfied.

- **Is the generator suitable for use in the cryptography? If not, why?**
  The generator is suitable for use in the cryptography.
  Characteristics of generators suitable for use in the cryptography:

  1. **it must be impossible to predict its seed and internal state even if we have a large sample of the numbers it produced**
     Satisfied: strength is based on modulo factorization.

2. **it must have a long period for every possible value of its seed**
   Satisfied.
3. **it should be unpredictable to the public, i.e. the probability of predicting the subsequent numbers should be low even if have a large sample of the numbers it produced**
   Satisfied: strength is based on modulo factorization.

A Statistics comparisons

Table 1: Comparisons of minimum and maximum values of period, mean and variance

		Generator 1		Generator 2		Generator 3	
Equation		$x_n = x_{n-1} \oplus x_{n-2} \pmod{2^{32}}$		$y_n = 3 \cdot y_{n-1} - 1 \pmod{2^{32}}$		$z_n = x_n \cdot y_n \pmod{2^{32}}$	
		value	condition	value	condition	value	condition
period	min	1	$x_0 = x_1 = 0$	proportional to $m-n$(tested up to 100000)		proportional to $m-n$(tested up to 100000)	
	max	3	elsewhere	2^{30}	x_0 is odd	proportional to $m-n$(tested up to 100000)	
mean	min	0	$x_0 = x_1 = 0$	✗		✗	
	max	2863311530	$x_0 = x_1 = 2^{32} - 1$	✗		✗	
variance	min	0	$x_0 = x_1 = 0$	✗		✗	
	max	*$4.0992764589155E+18$	$x_0 = x_1 = 2^{32} - 1$	✗		✗	

* - found during tests(not analytically checked)

B Good generator requirements

Table 2: Comparison of meeting requirements for good generators

	Generator 1	Generator 2	Generator 3
generated numbers distributions are as close as possible to the desired one	✗	✓	✓
subsequences of the produced sequence are mutually independent	✗	✓	✓
long period, with length at least $\sqrt{n}$, where n is the length of the used subsequence	✗	✓	✓
the ability to make jumps, i.e. to compute x_j from x_i for every $j > i$	✗	✓	✗
repeatable, portable and efficient	✓	✓	✓

C Cryptography suitability

Table 3: Comparison of cryptography suitability

	Generator 1	Generator 2	Generator 3
it must be impossible to predict its seed and internal state even if we have a large sample of the numbers it produced	✗	✗	✓
it must have a long period for every possible value of its seed	✗	✓	✓
it should be unpredictable to the public, i.e. the probability of predicting the subsequent numbers should be low even if have a large sample of the numbers it produced	✗	✗	✓