Aplikace matematiky

Jaroslav Král

Experimental properties of some additive pseudorandom number generators with random shuffling

Aplikace matematiky, Vol. 16 (1971), No. 6, 395–401

Persistent URL: http://dml.cz/dmlcz/103374

SVAZEK 16 (1971) APLIKACE MATEMATIKY ČISLO 6

EXPERIMENTAL PROPERTIES OF SOME ADDITIVE PSEUDORANDOM NUMBER GENERATORS WITH RANDOM SHUFFLING

JAROSLAV KRÁL

(Received April 29, 1970)

The work reported in this paper was induced by an urgent practical request to construct a pseudorandom number generator for the LINK 8 computer. Doubts sometimes occur whether such a generator can be constructed. We show that it is possible. We shall not give any theoretical treatment – the found generators are too complex from the theoretical point of view. This disadvantage was compensated with a complex testing. The best generator found is now used on the LINK 8 computer with very good results.

As mentioned above the generators were designed for the LINK 8 computer. The properties of the computer (a short word and a very slow multiplication) excluded multiplicative generators. It was therefore decided to design an additive generator or a generator based on shift and addition. The further study has shown that shifting does not improve substantially the quality of the generators. The interest was therefore focused on additive generators in which the computation mod 2^k is performed.

The first generator of such type studied was a generator based on the Fibonacci sequence, i.e. a generator using a sequence

$$a_i, a_{i+1}, \dots \quad \text{where} \quad a_{i+1} = a_i = a_{i-1} \pmod{2^{22}} \quad \text{for} \quad i = 2, 3, \dots$$

It is known that such sequences have bad statistical properties (see [3]). It was therefore decided to use the random shuffling technique, which can be described as follows. Let us have a sequence of $n = 2^m$ numbers $A_0, A_1, \dots, A_{n-1}$ and two numbers a and b. The generator can be given by the following Algol procedure (instead of Fibonacci sequence $a_1, a_2, a_3, \dots, a_n$ the sequence $b_i = a_i \times 2^{-22}$ is used):

```
real procedure random;
    begin integer i; real d;
        real procedure Fib;
            begin real c; c := if (a + b) ≧ 1 then a + b else a + b − 1;
            a := b; Fib := b := c end This procedure generates a Fibonacci
            sequence in the scale $2^{-22}$, the addition being performed in the fixed
            point arithmetics;
            i := entier (16 × Fib); comment on a binary computer realisable by
            a disjunction and a shift;
            d := A[i] ;  A[i] := Fib ;  random := d;
    end;
```

a, b, $A[0], \ldots, A[15]$ are nonlocal variables of the values less than 1 and not less than 2^{-22}. At least one number $a \times 2^{22}$ or $b \times 2^{22}$ must be odd. Although the above given program seems to be a little complicated it can be realized on a binary machine by few machine code instructions.Note that the procedure *Fib* is called twice. At first it is called when an "address" i in the array A is computed. The value of $A[i]$ is a new value of *random*. To the $A[i]$ is then assigned the result of the second calling of *Fib*. The generator just described will be called FRS-generator (Fibonacci with Random Shuffling).

It is very difficult to study the properties of the FRS-generators from the theoretical point of view. It can be exptected that the period of the FRS-generator is very large because FRS-generator produces the same sequence of pseudorandom numbers only after all the values of numbeis a, b, $A[o], \ldots, A[15]$ are the same. It follows that the length of the period of the FRS-generator is not great than 2^a, where $a = 18 \times 22$. The question how close is the length of the period of the FRS-generator to this very great value is open. Tests on the length of the period were performed for FRS-generators with different choice of the values a and b (the starting values of $A[o], \ldots \ldots, A[15]$ do not affect the properties of the FRS-generator substantially). It was found that the period in all the tested cases was greater than $4 \,.\, 10^6$. Tests of statistical properties (see [1,3]).

The following statistical tests were performed for the generators (see [1]).

(A) Uniformity tests. Using χ^2-test for the sequencies of the pseudorandom numbers of the length 2^{13}, the uniformity of the distribution of the numbers produced by the generators was tested, 256 classes.

(B) Uniformity of the maximum. It is the test (A) for the pseudorandom numbers of the form $c_j = (\max(b_{3j}, b_{3j+1}, b_{3j+2}))^3$, where b_k is the k-th number produced by the generator.

(C) Test of uniformity of pairs. It was tested for various $k \geqq 1$ whether the pairs (x_i, x_{i+k}) are uniformly distributed on the unit square. Number of classes 256.

(D) Runs above and below the median for the sequences of the length 2^{13}, number of classes (for the χ^2 test) seven.

(E) Runs up and down, number of classes for the χ^2-test seven, the length of sequences 24 000.

(F) For each test mentioned above, a "global" test (see [1]) was performed, i.e. the values $p = P(\chi_f^2 > \chi_c^2)$, where χ^2 is the obtained value of the test statistic, was computed for each test.

For the sequence of p (for the given test), it was tested whether p is uniformly distributed on $(0, 1)$. 5 classes.

Autocorrelation coefficients for produced sequences of *random* numbers were also computed. The generator (1) was tested for the following starting values of a and b (the values are given in the octal form)

	$a \,.\, 2^{22}$	$b \,.\, 2^{22}$
1	1453631	77 2517
2	1453631	77 0117
3	1453631	77 2617
4	1453631	77 2516
5	1453631	77 2511
6	1453631	77 2514
7	1453631	77 2515

It was discovered that the properties of FRS-generators essentially depend on the starting values of a and b. The properties of generators with the values of a, b equal to 1, 2, ..., 6 were bad. The choice 7 yielded almost satisfactory results. The greatest influence is exerted by changes in the last bits. It was confirmed that the test of runs up and down is the most sensitive one, but this fact is not too distinct. The global test is very sensitive. It discovers deviations from randomness, which cannot be discovered by separate tests (e.g. "too good" results of separate tests). It is likely that by a proper choice of the values of a and b even better properties of the FRS-generator can be achieved.

The strong dependence of FRS-generators on the starting values of a, b seems to be a little surprising, because in all cases the corresponding Fibonacci sequence has the period 2^{20} (see [3]). This is probably caused by the fact that there are $(2^{22} - 1) \,.\, (2^{21} - 2)\, 2 \geqq 2^{41}$ possibilities of the choice of the starting values of a and b and during one period with the given starting values only 2^{20} pairs of values of a, b appear. Therefore it exists about (for the word length 22 bits) 2^{20} Fibonacci generators, which can have different properties. The above given values of a and b can be used as the starting values for a FRS-generator using a sequence given by the recurrence $a'_{i+1} = a'_i + a'_{i-1} \pmod{2^{19}}$. It is interesting that this generator has for the same starting values very similar properties as the 22 bit generator discussed above. A pseudorandom number generator was also tested, the structure of which can be described in Algol 60 as follows.

```
real procedure random;
    begin integer i; real d;
    real procedure Perron;
    begin $d := a + b$; if $d \geqq 1$ then $d := d - 1$;
    $a := b$; $b := c$; $Perron := c := d$ end;
    $i := entier(16 \times Perron)$;
    $a := A[i]$; $random := A[i] := Perron$;
    end;
```

This generator yields better results than FRS-generator for all the above discussed tests excluding the run test, which yields a very bad result. The generator has a tendency to form long increasing and decreasing sequences. It is quite likely that this property could be avoided by other choice of the starting values.

Since for FRS-generator a satisfactory choice has not been found, another additive generator was studied. We shall call this generator PRS-generator. It can be described as follws (The design of the generator was influenced by [3]. However, we use the random shuffling technique and, moreover, entries from the shuffling table are substituted into the basic additive generator).

```
real procedure random;
    begin integer i; real c;
    $c := a + b$, if $c \geqq 1$ then $c := c - 1$; $i = entier(16 \times c)$;
    $b := c + A[i]$; if $b \geqq 1$ then $b := b - 1$;
    $random := A[i] := b$;
    $a := c$; comment $A$ is a nonlocal array declared as real array $A[0 : 15]$;
    end;
```

This generator is realizable by the following program ($\oplus$ is addition mod 1): $c := b$; $a := a \oplus b$; $i := entier\ (16 \times a)$; $b := b + A[i]$; $random\ := A[i]\ :=$ b; . On many computers it is faster as usuall multiplicative generator.

The PRS-generator was tested for the starting values a = 145 3631, b = 1, c = = 77 2515 (in the octal representation). This generator has very good statistical properties – see Table I.

In the above mentioned generators the array A contains 16 elements. PRS generators in which A contains 32, 64, 128 elements were also tested. The statistical properties of them were not better than those of the above described generator.

Worse properties were obtained only for the table length eight, especially for autocorrelation coefficients. Again it was observed that for the word length not less than 18 the properties of the generator do not vary substantially when the word length is increased.

The generator can be easily implemented on all computers, for example even on the byte computer of today without decimal arithmetics, and can be also easily implemented in firmware. It is convenient that the generator works for small word lengths. On computers with multiplication in hardware it can be slower than multiplicative generators.

The testing of the above discussed generators has shown that the most sensitive test is that on runs down and up and the global test. It seems that especially the global test which is not often used in literature is very useful, because it discovers deviations from "randomness" which cannot be discovered by other tests; see for example results for the PRS-generator in the appendix, viz. the case of pairs with $k = 31$, or the results of the up and down test for the FRS generator. In Tables 1 and II results are given not for all the studied generators. A great experimental evidence shows that in all cases when some individual test fails, the same holds also for the test (although not necessarily for the given test – see Table II). When the global test gave good results, then for example the autocorrelation coefficients of the produced sequence T as well as those of pseudorandom chains produced from Thad the expected properties. There were cases, when individual tests gave "good" results, but the global test failed. On the other hand, the global test is not sensitive to exceptional (extreme) values of individual tests. For example the case occured when one individual test twice produced (in a run of 50 tests) a value of statistics rejecting the randomness hypothesis on the level 10^{-5}.

APPENDIX

The results obtained for the above mentioned generators are given below.

For each test we give: the number of tests performed, number of cases when the hypothesis is rejected on the level $1^0/_{00}$, on the level 1% but not $1^0/_{00}$, on the level 3% but not 1% and on the level 5% but not on the level 3%. In the last column the probability (for the global test) of obtaining "worse" values of the statistic under the randomness hypothesis is given.

References

[1] *S. Gorenstein:* Testing a Random Numbers Generator, Comm. of ACM 10, (Feb. 1967), 111–118.

[2] *G. Marsaglia, M. D. McLaren:* Uniform Random Numbers Generators, Journ. of ACM 12, No 1 (1965).

[3] *J. C. P. Miller, M. J. Prentice:* Additive congruential pseudorandom number generators, Comp. J. 11, No 3, (1968), 341–346.

Table I — PRS-generator

	Number of tests	$1^0/_{00}$	1%	3%	5%	Global test %
Uniformity						
	53	—	2	—	4	23
Pairs (x_i, x_{i+k})						
$k = 1$	53	—	1	2	—	93
2	53	—	—	3	2	21
3	45	—	—	1	—	11
7	51	—	2	—	—	36
31	51	—	—	—	—	4
127	53	—	—	—	1	97,5
255	50	—	—	—	—	62
1023	55	—	—	—	2	29
8191	53	—	—	—	—	95
65535	50	—	—	2	—	7
Runs up and down						
	71	—	1	1	2	94
Runs above and below the median						
	82	1[1]	—	3	1	60
Triplets						
	52	—	—	—	—	69
$(\max(x_{3j}, x_{3j+1}, x_{3j+2}))^3$						
	56	1[2]	—	—	—	95

[1] $0{\cdot}8^0/_{00}$ [2] $0{\cdot}7^0/_{00}$

Table II — FRS generator

Uniformity						
	53	—	1	2	2	7
Pairs						
$k = 1$	52	—	1	1	—	10
2	56	—	2	—	—	30
3	51	—	2	—	2	6,6
7	52	—	—	1	—	17
19	63	—	1	2	1	10
31	58	1	—	1	1	34
127	56	—	1	—	1	65
1023	53	—	—	—	2	25
8191	51	—	—	2	1	26
6535	52	2	—	—	1	23[1]
Runs up and down						
	94	—	2	4	7	0·02
Runs above and below the median						
	227	—	3	—	—	75
$(\max(_{3j}, x_{3j+1}, x_{3j+2}))^3$						
	54	—	1	3	3	8·5
Triplets not performed						

[1] A value of statistics was obtained which rejects the randomness hypothesis on the level 10^{-}

Souhrn

EXPERIMENTÁLNÍ VLASTNOSTI JEDNOHO ADITIVNÍHO GENERÁTORU PSEUDONÁHODNÝCH ČÍSEL S NÁHODNÝM POSUVEM

Jaroslav Král

V článku je uveden algoritmus pro vytvoření generátoru pseudonáhodných čísel hlavně s použitím operací sčítání a posuvu, který je možno naprogramovat na libovolný počítač. Jsou uvedeny některé statistické testy a je doporučen optimální způsob implementace.

Author's address: Dr. *Jaroslav Král,* Ústav výpočtové techniky ČVUT, Horská 3, Praha 2.