Perturbations of Binary de Bruijn sequences

Martianus Frederic Ezerman, Adamas Aqsa Fahreza
NTU, Singapore
Janusz Szmidt, MCI, Poland

The 3rd International Workshop on Boolean Functions and their Applications
BFA 2018

20 June 2018

The Fryers Formula - 1

- Let $S(n)$ be the set of functions $\mathbb{F}_2^{n-1} \to \mathbb{F}_2$, which generate de Bruijn sequences of order n. For a function $F : \mathbb{F}_2^{n-1} \to \mathbb{F}_2$, we define the set $S(F; k)$ of functions $g \in S(n)$ such that the weight of the function $F + g$ equals k. It means that the number of inputs for which the functions F and g are different equals k.
- We introduce the notation $N(F; k) = |S(F; k)|$ and

$$G(F; y) = \sum_k N(F; k) y^k \tag{1}$$

- Let a linear function $\ell : \mathbb{F}_2^{n-1} \to \mathbb{F}_2$ generate the m-sequence of the period $2^n - 1$.

The Fryers Formula - 2

- In the work: *Counting de Bruijn sequences as perturbations of linear recursions* (arXiv e-prints, May 2017) by D. Coppersmith, R. C. Rhoades, J. M. Vanderkam it was proved the formula due to Michael Fryer:

$$G(\ell; y) = \frac{1}{2^n}\left((1+y)^{2^{n-1}} - (1-y)^{2^{n-1}}\right) \tag{2}$$

- The coefficients $N(\ell; k)$ can be calculated from the MacLaurin expansion for $G(\ell; y)$:

$$G(\ell; y) = \sum_{k=1}^{\infty} a_k y^k, \qquad a_k = \frac{1}{k!} G^{(k)}(\ell; y) \mid_{(y=0)} \tag{3}$$

- Let's consider some consequences.

The Fryers Formula - Corrolaries - 1

- $N(\ell; 1) = 1$. Hence, from an m-sequence we get one de Bruijn sequence by adding the cycle of the zero state, corresponding to one change in the truth table of the function ℓ.
- $N(\ell; 2) = 0$. The truth table of ℓ is changed in two places. One change adds the zero cycle and the second cuts the full cycle into two cycles. No new de Bruijn sequence obtained.
- In general, $N(\ell; k) = 0$ for all even k since an even number of changes in the truth table of ℓ always lead to disjoint cycles.
- There is in fact an interesting combinatorial view on the non-vanishing (> 0) coefficients of the polynomial $G(\ell, y)$.

The Fryers Formula - A Combinatorial View

- Recall that ℓ generates an m-sequence of period $2^n - 1$.
- Sequence A281123 in OEIS gives the formula for the positive coefficients of the polynomial

$$G(\ell; y) = q(n-1, y) = \frac{(1+y)^{2^{n-1}} - (1-y)^{2^{n-1}}}{2^n}.$$

- Hence, for odd $1 \leqslant k \leqslant 2^{n-1} - 1$, the formula for $N(\ell, k)$ is

$$N(\ell, k) = \frac{1}{2^{n-1}} \binom{2^{n-1}}{k} \text{ for } n \geqslant 2. \tag{4}$$

- More details and related integer sequences (such as Pascal Triangle) can be found in `https://oeis.org/A281123`

The Fryers Formula - Corrolaries - 2

- The Helleseth and Kløve formula (1991) follows from (4)

$$N(\ell; 3) = \frac{(2^{n-1} - 1)(2^{n-1} - 2)}{3!}$$

for the number of cross-join pairs for an m-sequence.

- The higher Helleseth and Kløve formula

$$N(\ell; k = 2j + 1 \geqslant 5) = \frac{1}{k!} \prod_{i=1}^{k-1} (2^{n-1} - i)$$

gives the number of new de Bruijn sequences obtained after the j-th application of cross-join method: (a) start from an m-sequence, (b) add 0 to obtain a de Bruijn sequence, (c) find all of its cross-join pairs, (d) use them to construct new de Bruijn sequences, (e) for each resulting sequence, repeat (c) and (d) $j - 1$ times.

The Fryers Formula - Corrolaries - 3

- Using (4) we easily obtain the number of all cyclically non-equivalent de Bruijn sequences of order n:

$$G(\ell,1)=\sum_{k=1}^{2^{n-1}-1}N(\ell,k)=\frac{1}{2^{n-1}}\underbrace{\sum_{k=1}^{2^{n-1}-1}\binom{2^{n-1}}{k}}_{:=\alpha}=2^{2^{n-1}-n} \tag{5}$$

since $\alpha=2^{2^{n-1}-1}$ is the sum of the odd entries in row 2^{n-1} of the Pascal Triangle.

- For $n=4$

$$\sum_{k=1}^{7}N(\ell,k)=1+7+7+1=2^4.$$

For $n = 5$ and $n = 6$

- For $n = 5$

$$\sum_{k=1}^{15} N(\ell, k) = 1 + 35 + 273 + 715 + 715 + 273 + 35 + 1 = 2^{11}.$$

- For $n = 6$

$$\sum_{k=1}^{31} N(\ell, k) = 1 + 155 + 6293 + 105183 + 876525 + 4032015$$
$$+ 10855425 + 17678835 + 17678835 + 10855425 + 4032015$$
$$+ 876525 + 105183 + 6293 + 155 + 1 = 67108864 = 2^{26}$$

- The coefficients are symmetric.

Definitions and Notations

- Two binary de Bruijn sequences $\mathbf{v}$ and $\mathbf{u}$ of order n can be obtained from each other by applying the cross join method, possibly repeatedly [1].
- The *truth table distance* is the smallest number of assignments in the *truth table* of the feedback function of $\mathbf{v}$ that must be changed to get the truth table of $\mathbf{u}$.
- If $\mathbf{v}$ and $\mathbf{u}$ have distance $2j$, then there are j cross join pairs between them.
- Let $\mathbf{v}$ be a de Bruijn sequence constructed by adding 0 to the longest string of zeros in an m-sequence $\mathbf{v}'$ of length $2^n - 1$.
- The Fryer's formula gives the number of de Bruijn sequences of distance $2i + 1$ for all $0 \leqslant i \leqslant 2^{n-2} - 1$.

---

[1] J. Mykkeltveit and J. Szmidt, *On cross joining de Bruijn sequences*, Contemporary Mathematics, **632**, pp. 333-344 (2015).

General Perturbation Patterns

- The situation when $\mathbf{v}'$ is **not** an m-sequence is less clear.
- We study also the perturbation of such sequences and provide a complete classification for some small orders.
- The list of non-ovelapping cross join pairs between de Bruijn sequences $\mathbf{v}$ and $\mathbf{u}$ can be determined.

The Setup

- Objective: Given n, generate all de Bruijn Sequences of order n as perturbations of any one of them, say $\mathbf{v}$.
- Let $g := \mathbb{F}_2^{n-1} \mapsto \mathbb{F}_2$.
- If, after altering the truth table of $\mathbf{v}$ by pairs, there are two runs of n zeroes, then the resulting sequence is **not** a de Bruijn sequence.
- The TEST is passed when there are no two runs of n zeros.

An Algorithm (Not Very Efficient Yet)

Input: Any de Bruijn sequence **v** of order n.
Output: All $2^{2^{n-1}-n}$ de Bruijn sequences of order n.

1: $count \leftarrow 0$
2: Construct the truth table of **v**
3: Evaluate any $n-1$ string based on $f := x_0 + g(x_1, \ldots, x_{n-1})$
4: **for** i from 1 to $2^{n-2}-1$ **do** ▷ There are $2^{n-2}-1$ possible pairings
5: Order the $n-1$ strings lexicographically
6: Pair any $2i$ strings of length $n-1$ systematically
7: Exchange their truth table values to form the sequence **w**
8: **if w** passes the TEST **then**
9: output **w** and $count \leftarrow count + 1$
10: **end if**
11: **end for**
12: **if** $count = 2^{2^{n-1}-n}$ **then**
13: output $count$
14: **else**
15: print an error message
16: **end if**

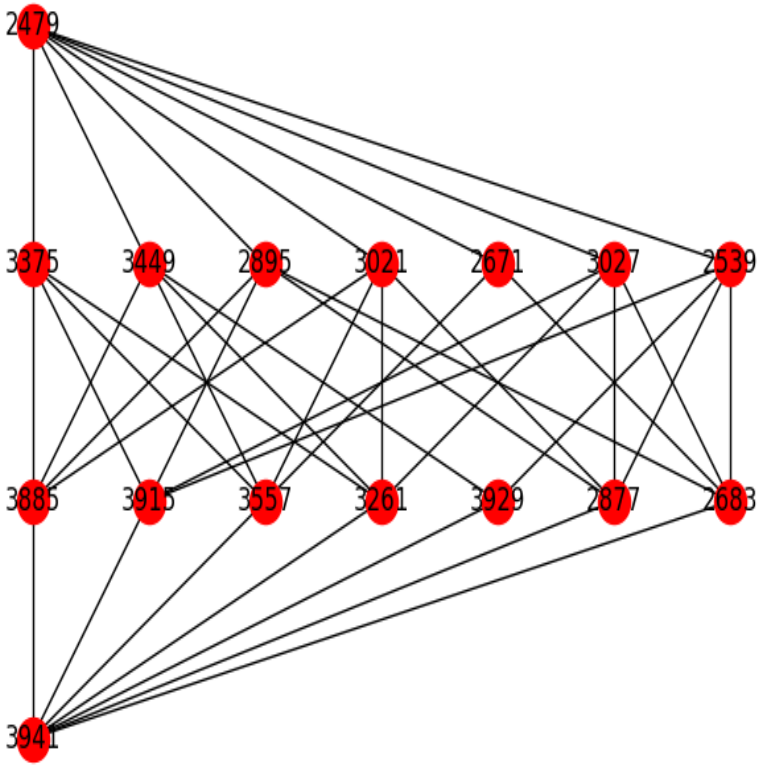
An Example

Let $\mathbf{v} = (0000\ 1011\ 1101\ 0011)$ with $\mathbf{v}'$ not an m-sequence.

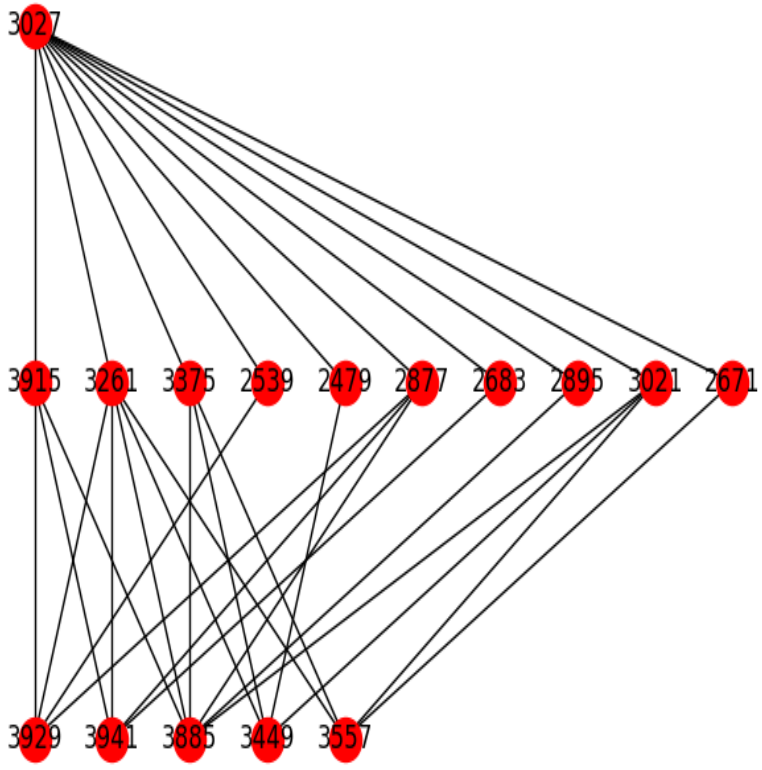
No.	Resulting Sequence	Cross-Join Pairs
1	(0000 1111 0100 1011)	$(001, 011)$
2	(0000 1100 1011 1101)	$(001, 100)$
3	(0000 1101 0010 1111)	$(001, 110)$
4	(0000 1001 1110 1011)	$(010, 011)$
5	(0000 1001 1010 1111)	$(010, 110)$
6	(0000 1011 0011 1101)	$(011, 100)$
7	(0000 1010 0111 1011)	$(011, 101)$
8	(0000 1011 0100 1111)	$(011, 110)$
9	(0000 1011 1100 1101)	$(100, 110)$
10	(0000 1010 0110 1111)	$(101, 110)$
1	(0000 1111 0101 1001)	$(001, 010, 011, 100)$
2	(0000 1101 0111 1001)	$(001, 010, 100, 110)$
3	(0000 1111 0110 0101)	$(001, 011, 100, 101)$
4	(0000 1111 0010 1101)	$(001, 011, 100, 110)$
5	(0000 1101 1110 0101)	$(001, 100, 101, 110)$

Two Patterns for $n = 4$

Symmetric $(1, 7, 7, 1)$



Asymmetric $(1, 10, 5, 0)$



Generating Starting Sequences

- The starting de Bruijn sequence is generated by the cycle-joining method from LFSR with the specified characteristic polynomial.
- The construction is discussed By Z. Chang et al. in `arxiv.org/pdf/1611.10088v3.pdf`.
- `python` implementation of the construction is in `https://github.com/adamasstokhorst/debruijn`
- The remaing slides present results on the general patterns for $n \in \{4, 5\}$.

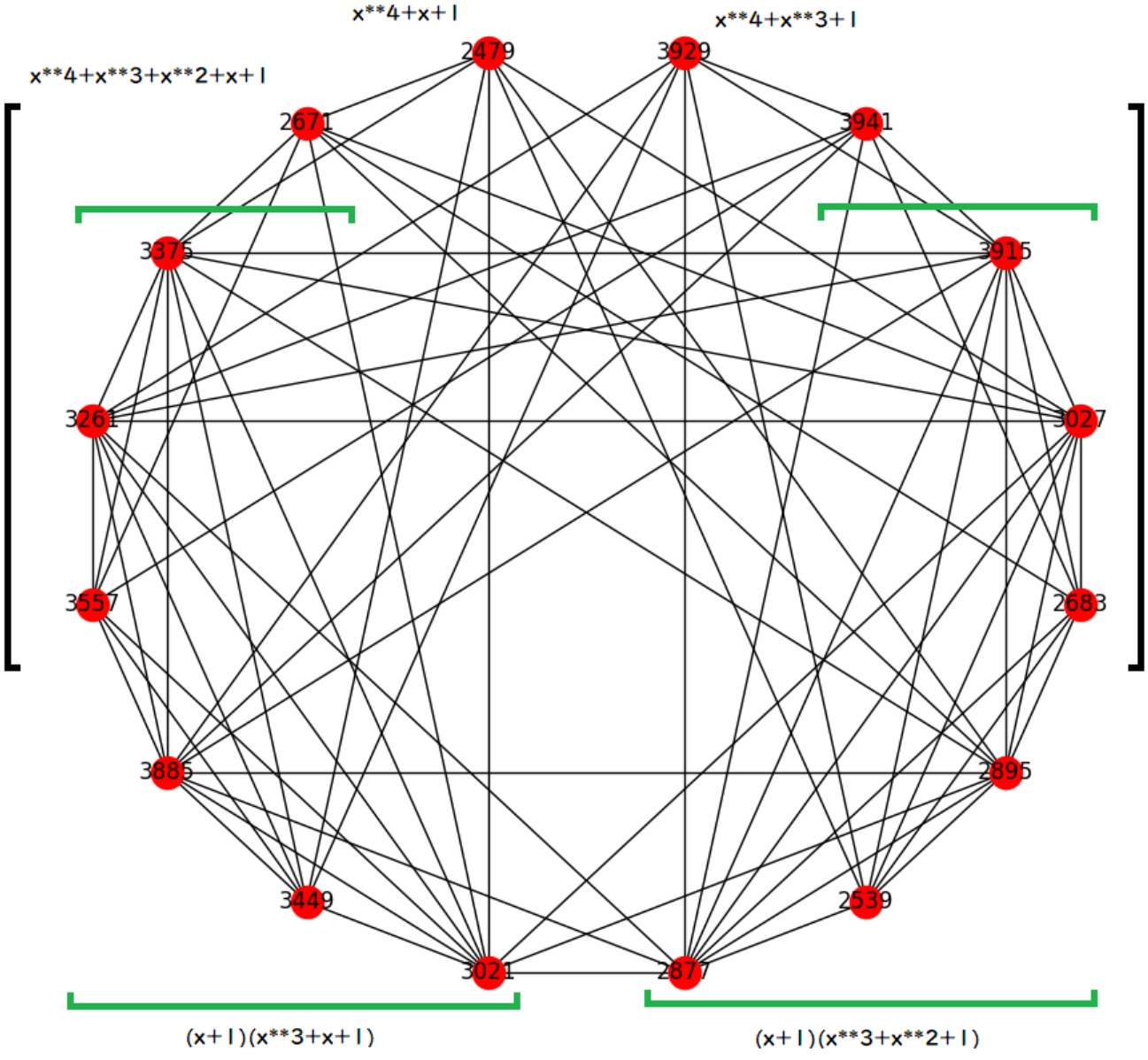
Sequences of order $n = 4$ with Coefficients $(1, 7, 7, 1)$

No.	de Bruijn Sequence	Decimal	Char. Poly. of LFSR
1	0000100110101111	2479	$x^4 + x + 1$
2	0000111101011001	3929	$x^4 + x^3 + 1$
3	0000101001101111	2671	$x^4 + x^3 + x^2 + x + 1$
4	0000111101100101	3941	
5	0000110111100101	3557	$x^4 + x^3 + x^2 + x + 1$ or $(x + 1)(x^3 + x + 1)$
6	0000101001111011	2683	$x^4 + x^3 + x^2 + x + 1$ or $(x + 1)(x^3 + x^2 + 1)$
7	0000110101111001	3449	$(x + 1)(x^3 + x + 1)$
8	0000100111101011	2539	$(x + 1)(x^3 + x^2 + 1)$

Sequences of order $n = 4$ with Coefficients $(1, 10, 5, 0)$

No.	de Bruijn Sequence	Decimal	Char. Poly. of LFSR
1	0000110010111101	3261	$x^4 + x^3 + x^2 + x + 1$
2	0000110100101111	3375	or $(x + 1)(x^3 + x + 1)$
3	0000101111010011	3027	$x^4 + x^3 + x^2 + x + 1$
4	0000111101001011	3915	or $(x + 1)(x^3 + x^2 + 1)$
5	0000111100101101	3885	$(x + 1)(x^3 + x + 1)$
6	0000101111001101	3021	
7	0000101100111101	2877	$(x + 1)(x^3 + x^2 + 1)$
8	0000101101001111	2895	

The "Cross-Join Connectivity Graphs" for $n = 4$



Patterns for $n = 5$

- For $n = 5$ there are exactly 60 perturbation patterns.
- The pattern in blue matches the Fryer's coefficients. Note that only 6 out of the 96 sequences correspond to the m-sequences.
- Other symmetric patterns are in bold.
- Each of the two pairs of patterns in red consists of reversals.
- Our investigation into various interesting patterns has only just begun.
- There are many open questions and directions.

First 30 of Exactly 60 Distribution Patterns for Order 5

Coefficients	#	Coefficients	#
(**1**, **34**, **276**, **713**, **713**, **276**, **34**, **1**)	192	(1, 34, 297, 804, 699, 202, 11, 0)	32
(1, 39, 310, 790, 677, 211, 20, 0)	112	(1, 40, 317, 768, 691, 216, 15, 0)	32
(1, 37, 322, 770, 685, 217, 16, 0)	96	(1, 35, 294, 806, 701, 199, 12, 0)	32
(**1**, **35**, **273**, **715**, **715**, **273**, **35**, **1**)	**96**	(1, 32, 278, 717, 709, 274, 36, 1)	32
(1, 32, 237, 640, 739, 352, 47, 0)	80	(1, 36, 315, 792, 671, 212, 21, 0)	32
(1, 45, 351, 743, 639, 235, 33, 1)	64	(1, 34, 235, 636, 743, 354, 45, 0)	32
(**1**, **36**, **270**, **717**, **717**, **270**, **36**, **1**)	64	(1, 31, 260, 726, 737, 267, 26, 0)	32
(1, 37, 274, 706, 717, 281, 32, 0)	64	(1, 32, 261, 720, 739, 272, 23, 0)	32
(1, 33, 235, 639, 743, 351, 45, 1)	64	(**1**, **33**, **275**, **719**, **711**, **271**, **37**, **1**)	32
(1, 33, 262, 714, 741, 277, 20, 0)	64	(**1**, **37**, **271**, **711**, **719**, **275**, **33**, **1**)	32
(1, 33, 278, 714, 709, 277, 36, 0)	64	(1, 39, 316, 774, 689, 211, 18, 0)	32
(**1**, **47**, **349**, **739**, **643**, **237**, **31**, **1**)	48	(1, 48, 301, 672, 675, 304, 47, 0)	32
(**1**, **31**, **237**, **643**, **739**, **349**, **47**, **1**)	48	(1, 49, 364, 834, 633, 157, 10, 0)	32
(1, 36, 274, 709, 717, 278, 32, 1)	32	(1, 47, 370, 830, 629, 163, 8, 0)	32
(1, 40, 341, 752, 659, 232, 23, 0)	32	(**1**, **47**, **301**, **675**, **675**, **301**, **47**, **1**)	32

Last 30 of Exactly 60 Distribution Patterns for Order 5

Coefficients	#	Coefficients	#
(1, 35, 276, 710, 713, 279, 34, 0)	32	(1, 41, 352, 858, 649, 141, 6, 0)	16
(1, 35, 318, 790, 669, 215, 20, 0)	24	(1, 45, 372, 834, 625, 161, 10, 0)	16
(1, 43, 374, 838, 621, 159, 12, 0)	24	(1, 43, 352, 846, 665, 135, 6, 0)	16
(1, 45, 366, 850, 613, 161, 12, 0)	16	(1, 41, 314, 770, 693, 213, 16, 0)	16
(1, 42, 315, 764, 695, 218, 13, 0)	16	(1, 38, 319, 772, 687, 214, 17, 0)	16
(1, 39, 352, 870, 633, 147, 6, 0)	16	(1, 41, 382, 826, 629, 157, 12, 0)	16
(1, 45, 344, 858, 657, 137, 6, 0)	16	(1, 48, 349, 736, 643, 240, 31, 0)	16
(1, 47, 346, 846, 661, 147, 0, 0)	16	(1, 32, 269, 704, 739, 288, 15, 0)	8
(1, 34, 267, 700, 743, 290, 13, 0)	16	(1, 47, 366, 838, 629, 155, 12, 0)	8
(1, 37, 306, 770, 717, 217, 0, 0)	16	(1, 36, 265, 696, 747, 292, 11, 0)	8
(1, 51, 362, 830, 637, 159, 8, 0)	16	(1, 39, 382, 838, 613, 163, 12, 0)	8
(1, 35, 324, 774, 681, 215, 18, 0)	16	(1, 39, 330, 878, 677, 123, 0, 0)	8
(1, 33, 296, 810, 697, 197, 14, 0)	16	(1, 60, 401, 776, 603, 188, 19, 0)	8
(1, 62, 399, 772, 607, 190, 17, 0)	16	(1, 31, 282, 814, 725, 195, 0, 0)	8
(1, 35, 304, 782, 713, 207, 6, 0)	16	(1, 64, 397, 768, 611, 192, 15, 0)	8