Evolving Constructions for Balanced, Highly Nonlinear Boolean Functions with GP

Claude Carlet, Marko Djurasevic, Domagoj Jakobovic,
Luca Mariot, Stjepan Picek

luca.mariot@ru.nl

GECCO 2022 – Boston, July 11, 2022

Boolean Functions - Basic Definitions

- $\mathbb{F}_2 = \{0,1\}$, $\mathbb{F}_2^n$: n-dimensional vector space over $\mathbb{F}_2$
- A *Boolean function* of n variables is a mapping $f : \mathbb{F}_2^n \to \mathbb{F}_2$, most commonly represented by its *Truth Table* (TT) Ω_f
- *Walsh Transform* (WT): represents f as *correlations* with *linear* functions $a \cdot x$, for $a \in \mathbb{F}_2^n$

$$W_f(a) = \sum_{x \in \mathbb{F}_2^n} (-1)^{f(x) \oplus a \cdot x}$$

- Example with $n = 3$ variables:

(x_1, x_2, x_3)	000	001	010	011	100	101	110	111
Ω_f	0	1	1	0	1	0	1	0
$W_f(a)$	0	-4	0	4	0	4	0	4

Boolean Functions - Cryptographic Properties

To be useful in cryptography, $f : \mathbb{F}_2^n \to \mathbb{F}_2$ should be:

- **Balanced:** the TT of f has the same number of 0s and 1s
- **Highly nonlinear:** the nonlinearity of f is given by the WT as follows [C21]:

$$nl(f) = 2^{n-1} - \frac{1}{2} \max_{a \in \mathbb{F}_2^n} \left\{ \left| W_f(a) \right| \right\}$$

(x_1, x_2, x_3)	000	001	010	011	100	101	110	111
Ω_f	0	1	1	0	1	0	1	0
$W_f(a)$	0	-4	0	4	0	4	0	4

- Example: f balanced, $nl(f) = 2^{3-1} - \frac{1}{2} \cdot 4 = 2$

Constructions of good Boolean Functions

- Number of Boolean functions of n variables: 2^{2^n}

n	3	4	5	6	7	8
2^{2^n}	256	65536	$4.3 \cdot 10^9$	$1.8 \cdot 10^{19}$	$3.4 \cdot 10^{38}$	$1.2 \cdot 10^{77}$

- $\Rightarrow$ too huge for exhaustive search when $n > 5$!

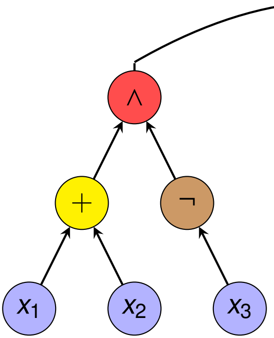
In practice, one usually resorts to:

- **Primary Constructions**, where Boolean functions with certain properties are built from scratch [M73, D74]
- **Secondary Constructions**, where new Boolean functions are obtained from existing one [R76]
- **Heuristic optimization algorithms** (GA [M98, M15b, M20], GP [P16, M19a, M19b], PSO [M15a], ...)

Direct Construction with GP

- The truth table $g(x)$ is synthesized from the tree [P16]

Example: $n = 3$, $f : \{0,1\}^3 \rightarrow \{0,1\}$



x_1	x_2	x_3	$f(x)$
0	0	0	0
0	0	1	0
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	0
1	1	0	0
1	1	1	0

- **Problem:** scales poorly when n increases

Evolving Secondary Constructions

Example of secondary construction: *Rothaus's* construction [R76]

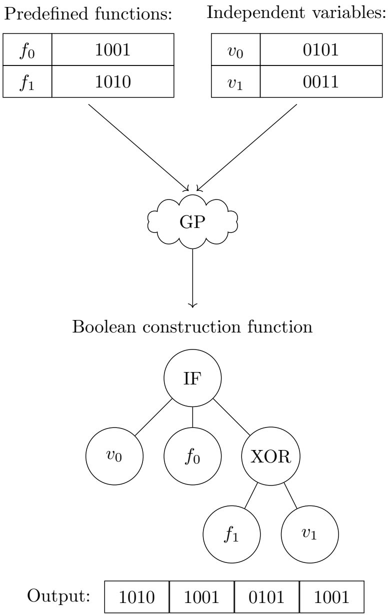
- If g, h, k and $g \oplus h \oplus k$ are bent (maximally nonlinear) on $\mathbb{F}_2^n$, then the following function is bent:

$$f(x_1, x_2, x) = g(x)h(x) \oplus g(x)k(x) \oplus h(x)k(x) \oplus \\ \oplus [g(x) \oplus h(x)]x_1 \oplus [g(x) \oplus k(x)]x_2 \oplus x_1 x_2$$

where $(x_1, x_2, x) \in \mathbb{F}_2^{n+2}$ with $x_1, x_2 \in \mathbb{F}_2$, $x \in \mathbb{F}_2^n$

Goal: Evolve secondary constructions using GP

GP Representation



- **Idea:** represent a secondary construction as a GP tree
- f_0, f_1: *seed* functions
- v_0 v_1: additional *independent variables*
- The GP tree yields a new function of $n+2$ variables
- Seed functions are obtained through direct GP search

Experimental Settings

Problem-related parameters:

Parameter description	Parameter value
Number of variables	5, 6, 7, 8
Independent variables	1, 2
Number of seed functions	2, 4
Number of seed function groups	4
Seed functions type	balanced, bent
Type of fitness function	first group, sum of all groups, minimum of all groups

GP-related parameters:

- Population size: 500
- Mutation probability: 0.5
- Fitness budget: 500 000
- Max tree depth: 5
- Tournament size: 3
- Independent runs: 30

Table: Results for search-based GP

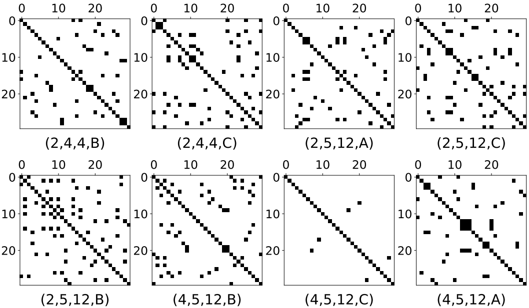
GP search	n = 9	n = 10	n = 11	n = 12	n = 13	n = 14	n = 15	n = 16	n = 17	n = 18
min NL	240	480	992	1 953	3 905	8 001	16 192	32 512	65 280	130 561
avg NL	240	484.24	992	1 996.69	4 028.39	8 087.3	16 253.5	32 542.2	65 280	130 622
max NL	240	492	992	2 008	4 032	8 120	16 256	32 608	65 280	130 753
# max	100%	2%	100%	14%	96%	1%	97%	2%	100%	2%

Table: Results for secondary constructions

constr.	n = 8	n = 9	n = 10	n = 11	n = 12	n = 13	n = 14	n = 15	n = 16	n = 17	n = 18
seed NL	26	56	116	240	488	992	2 000	4 032	8 096	16 256	32 576
res. NL	116	240	488	992	2 000	4 032	8 096	16 256	32 576	65 280	130 688

Simplification of GP Solutions

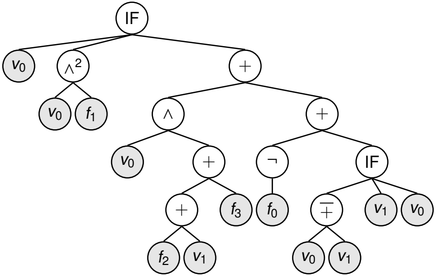
- We used the ESPRESSO tool [R87] to *minimize* the best GP trees
- Performed an **equivalence check** among the best solutions



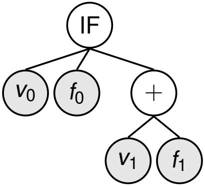
- **Result**: many solutions turn out to be the same construction, especially when 2 seeds are used

Interpretation of Simplest Solutions

Example of bloated GP construction:



Main Remark: many constructions are equivalent to the well-known *indirect sum construction* [C21]



$$F(v_0, v_1, v) = \begin{cases} f_0(v) \ , & \text{if } v_0 = 1 \ , \\ f_1(v) \oplus v_1 \ , & \text{if } v_0 = 0 \ . \end{cases}$$

Conclusions and Future Works

Conclusions:

- GP can be used to evolve secondary constructions for Boolean functions, rather than Boolean functions directly
- Most of the solutions evolved by GP are equivalent, and correspond to the indirect sum construction

Future work:

- Test whether the indirect sum is the *only* construction achievable by GP under this encoding
- Experiment with *non-independent* additional variables

References

[C21] Carlet, C.: Boolean functions for cryptography and coding theory. Cambridge University Press (2021)

[D74] Dillon, J.F.: Elementary Hadamard difference sets. Ph.D. thesis, Univ. of Maryland (1974)

[M73] McFarland, R.L.: A family of difference sets in non-cyclic groups. J. Comb. Theory, Ser. A 15(1): 1–10 (1973)

[M20] Manzoni, L., Mariot, L., Tuba, E.: Balanced crossover operators in Genetic Algorithms. Swarm Evol. Comput. 54: 100646 (2020)

[M19a] Mariot, L., Jakobovic, D., Leporati, A., Picek, S.: Hyper-bent Boolean Functions and Evolutionary Algorithms. In: Proceedings of *EuroGP 2019*: 262–277 (2019)

[M19b] Mariot, L., Picek, S., Leporati, A., Jakobovic, D.: Cellular automata based S-boxes. Cryptography and Communications 11(1):41–62 (2019)

[M15a] Mariot, L., Leporati, A.: Heuristic Search by Particle Swarm Optimization of Boolean Functions for Cryptographic Applications. In: GECCO 2015 (Companion): 1425–1426. ACM (2015)

[M15b] Mariot, L., Leporati, A.: A Genetic Algorithm for Evolving Plateaued Cryptographic Boolean Functions. In: Proceedings of *TPNC 2015*: 33–45 (2015)

[M98] Millan, W., Clark, J., Dawson, E.: Heuristic Design of Cryptographically Strong Balanced Boolean Functions. Proceedings of EUROCRYPT 1998, pp. 489–499 (1998)

[P16] Picek, S., Jakobovic, D., Miller, J.F., Batina, L., Cupic, M.: Cryptographic Boolean functions: One output, many design criteria. Appl. Soft Comput. 40: 635–653 (2016)

[R76] Rothaus, O.S.: On “bent” functions. J. Comb. Theory, Ser. A 20(3): 300–305 (1976)

[R87] Rudell, R.L., Sangiovanni-Vincentelli, A.L.: Multiple-Valued Minimization for PLA Optimization. IEEE Trans. Comput. Aided Des. Integr. Circuits Syst. 6(5): 727–750 (1987)