Function representation for benchmarking

Sébastien Verel
Univ. du Littoral Côte d'Opale
LISIC Lab
Calais, France
verel@univ-littoral.fr

Lorentz Center workshop, Leiden, Nov. 8th, 2020.

Introduction

Usually, there is an falsifiable hypothesis,
then, the researcher designs a benchmark to test the hypothesis

Benchmark : a tool to test futur hypotheses

Set of functions that make sense:

- for the designer(s) of the benchmark
- according to the targeted real wold applications

- SAT community:
  all problems are SAT problems (NP-complete), competitions
- IOHProfiler:
  oneMax, leadingOne, W-model, LABS, Ising model, vertex set, nk-landscapes ; $af(\sigma(x+z))+b$

Representation of functions

$\{f : \mathcal{X} \to \mathbb{R}\}$ is a vector space

Pseudo-boolean functions

$\{f : \{0,1\}^n \to \mathbb{R}\}$ is vector space of dimension 2^n

Any function f can be expanded:

$$f(x) = \sum_{i=0}^{2^n-1} \beta_i \, \varphi_i(x)$$

Consequence

Replace a "clever way" to design a function

by a vector $\beta = (\beta_0, \ldots, \beta_{2^n-1}) \in \mathbb{R}^n$.

Benchmark design

$$f(x) = \sum_{i=0}^{2^n-1} \beta_i \, \varphi_i(x)$$

- What (φ_i) basis?
- What β coefficients?

- **Basis**:

  A number of basis are well known

  Legendre, Walsh, group representation, etc.

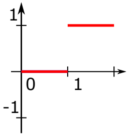
- **Coefficients**:

  In learning context, mean and variance of a property P according to the mean, and variance of β distribution

Basis of pseudo-boolean functions

Multi-linear :

$n = 1$, $\psi_1(x) = x$

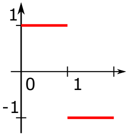


Orthogonal : No

x	ψ_0	ψ_1
0	1	0
1	1	1

Walsh :

$n = 1$, $\varphi_1(x) = (-1)^x$



Orthogonal : Yes

x	φ_0	φ_1
0	1	1
1	1	-1

Extension to dimension n by tensoriel product:

$$\psi_{k_1 \dots k_\ell}(x) = x_{k_1} \dots x_{k_\ell}$$

$$\varphi_{k_1 \dots k_\ell}(x) = (-1)^{x_{k_1}} \dots (-1)^{x_{k_\ell}}$$

Walsh basis

$$f(x) = 1 + (-1)^{x_1} + 3\,(-1)^{x_2} + (-1)^{x_3} + 4\,(-1)^{x_1+x_2} + 2\,(-1)^{x_1+x_2+x_3}$$

Properties that can be controled

- Non-linearity : $(-1)^{x_{i_1}} \dots (-1)^{x_{i_p}}$
- Intensity of non-linearities : β_k
- Complexity of the model: number of non-zero terms

Discussion

- Close to the SAT philosophy
- Easy to convert oneMax, nk, SAT, etc. into Walsh basis.
- Real-world problems correspond to $\beta \in \mathbb{R}^n$ distributions
- Very fast to compute (additive property, gradient)
- Straightforward file format
- Maybe can be extended to other search spaces