# Of (hyper)graphs and functions of binary variables: Old and recent results

Yves Crama
HEC Management School, University of Liège, Belgium

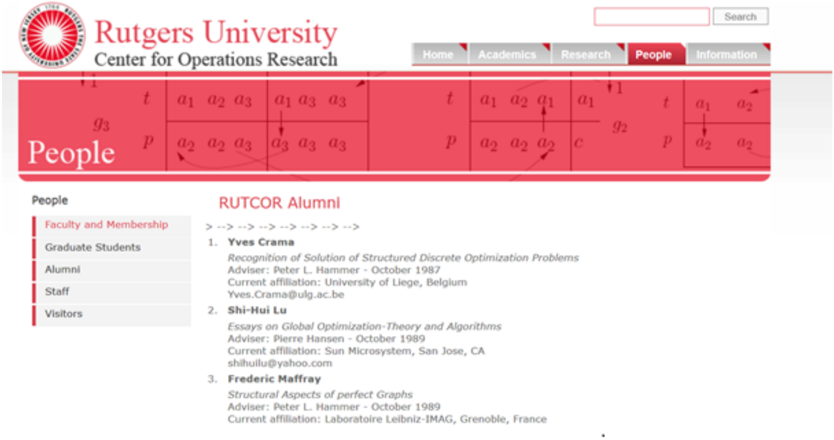
Grenoble, September 2019

# Outline

# At Rutgers University

- Both at Rutgers University, RUTCOR
- Frédéric's PhD degree: 1989

# At Rutgers University



# At Rutgers University

- Both at Rutgers University, RUTCOR
- Frédéric's PhD degree: 1989
- Advisor: Peter L. Hammer
- Interest in combinatorial structures and functions of 0-1 variables:

P.L. Hammer, F. Maffray. Completely separable graphs. *Discrete Applied Mathematics* 27 (1990), 85-99.
P.L. Hammer, F. Maffray, M. Queyranne. Cut-threshold graphs. *Discrete Applied Mathematics* 30 (1991), 163-179.
C. Benzaken, Y. Crama, P. Duchet, P.L. Hammer, F. Maffray. More characterizations of triangulated and cotriangulated graphs. *Journal of Graph Theory* 14 (1990), 413-422.

# Graph-parameter functions

- $G = (V, E)$ is *perfect* if, for all $S \subseteq V$: $\alpha(G[S]) = \theta(G[S])$ (stability number of $G[S]$ = clique cover number of $G[S]$).
- Idea: look at the function $\alpha_G : 2^V \mapsto \mathbb{R} : S \to \alpha(G[S])$.
- Similarly for $\theta_G$. ($G$ is perfect if $\alpha_G = \theta_G$, viewed as functions.)
- Identify $2^V$ with $\{0, 1\}^n$ ($n = |V|$): then $\alpha_G$ is a real-valued function of 0-1 variables (*pseudo-Boolean function*):

$$\alpha_G(x_1, \dots, x_n) = \alpha(G[S]), \quad S \text{ is indexed by } (x_1, \dots, x_n).$$

- $\alpha_G$ has a unique representation as a multilinear polynomial in 0-1 variables.
- What does this polynomial look like??

# Graph-parameter functions

Some examples:

- If $G = K_n$, then $\alpha_G = \theta_G = 1 - (1 - x_1)(1 - x_2) \ldots (1 - x_n)$
- If $G = 2K_2$, then $\alpha_G = \theta_G = x_1 + x_2 + x_3 + x_4 - x_1 x_3 - x_2 x_4$
- If $G = P_4$, then
  $\alpha_G = \theta_G = x_1 + x_2 + x_3 + x_4 - x_1 x_2 - x_2 x_3 - x_3 x_4 + x_1 x_2 x_3 + x_2 x_3 x_4 - x_1 x_2 x_3 x_4$
- If $G = C_4$, then $\alpha_G = \theta_G = x_1 + x_2 + x_3 + x_4 - x_1 x_2 - x_2 x_3 - x_3 x_4 - x_1 x_4 + x_1 x_2 x_3 + x_1 x_2 x_4 + x_1 x_3 x_4 + x_2 x_3 x_4 - 2x_1 x_2 x_3 x_4$

**Theorem (BCDHM 1990)**

The polynomial expression of the stability function of $G$ has all its coefficients equal to 0, - 1, or + 1 if and only if $G$ is triangulated. Moreover, when this is the case, the coefficients alternate in sign between odd and even degree terms.

# Extensions?

- Nice, but anecdotic result?
- Generalization: the hypergraph $H = \{123, 124, 34\}$ has $\alpha_H = x_1 + x_2 + x_3 + x_4 - x_3x_4 - x_1x_2x_3 - x_1x_2x_4 + x_1x_2x_3x_4$.
- Anything special about it? What hypergraphs have all coefficients equal to 0, - 1, or + 1 ?

# Definitions

## Pseudo-Boolean functions

A *pseudo-Boolean function* is a mapping $f : \{0,1\}^n \to \mathbb{R}$, that is, a real-valued function of $0-1$ variables.

## Multilinear polynomials

Every pseudo-Boolean function can be represented – in a unique way – as a *multilinear polynomial* in its variables, of the form

$$f(x_1,\ldots,x_n) = \sum_{S\in\mathcal{S}} a_S \prod_{k\in S} x_k + \sum_{i=1}^{n} a_i x_i$$

where $\mathcal{S} = \{S \in 2^{[n]} \mid a_S \neq 0, |S| \geq 2\}$.

## Example:

$f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3x_4 - 13x_1x_2x_3x_4$

# Co-occurrence hypergraph

## Co-occurrence hypergraph

When

$$f(x_1, \ldots, x_n) = \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k + \sum_{i=1}^{n} a_i x_i,$$

$H_f = ([n], \mathcal{S})$ is the *co-occurrence hypergraph* associated with $f$.

## Example:

If $f = 4 - 9x_1 - 5x_2 - 2x_3 + 13x_1x_2 + 13x_1x_3 + 6x_2x_3x_4 - 13x_1x_2x_3x_4$, then $\mathcal{S} = \{12, 13, 234, 1234\}$.

# Multilinear optimization in binary variables

We are frequently interested in:

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k + \sum_{i=1}^{n} a_i x_i$$

- Multilinear optimization is NP-hard, even if $f$ is quadratic
- Approaches:
  - Direct resolution methods
  - *Linearization*: extensive literature in integer programming.
  - *Quadratization*: more recent approach.

# Multilinear optimization in binary variables

We are frequently interested in:

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k + \sum_{i=1}^{n} a_i x_i$$

- Multilinear optimization is NP-hard, even if $f$ is quadratic
- Approaches:
  - Direct resolution methods
  - ***Linearization***: extensive literature in integer programming.
  - *Quadratization*: more recent approach.

# Standard linearization (SL)

$$\text{(MOB)} \quad \min_{x\in\{0,1\}^n} \sum_{S\in\mathcal{S}} a_S \prod_{k\in S} x_k + \sum_{i=1}^{n} a_i x_i,$$

## 1. Substitute monomials

$$\begin{aligned}
\min \sum_{S\in\mathcal{S}} a_S y_S + \sum_{i=1}^{n} a_i x_i & \\
\text{s.t. } y_S = \prod_{k\in S} x_k, & \qquad \forall S \in \mathcal{S} \\
y_S \in \{0,1\}, & \qquad \forall S \in \mathcal{S} \\
x_k \in \{0,1\} & \qquad \forall k = 1,\dots,n
\end{aligned}$$

# Standard linearization (SL)

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k + \sum_{i=1}^{n} a_i x_i,$$

## 1. Substitute monomials

$$\begin{aligned} \min \; & \sum_{S \in \mathcal{S}} a_S y_S + \sum_{i=1}^{n} a_i x_i \\ \text{s.t. } & y_S = \prod_{k \in S} x_k, && \forall S \in \mathcal{S} \\ & y_S \in \{0,1\}, && \forall S \in \mathcal{S} \\ & x_k \in \{0,1\} && \forall k = 1, \ldots, n \end{aligned}$$

## 2. Linearize constraints

$$\begin{aligned} \min \; & \sum_{S \in \mathcal{S}} a_S y_S + \sum_{i=1}^{n} a_i x_i \\ \text{s.t. } & y_S \leq x_k, && \forall k \in S, \forall S \in \mathcal{S} \\ & y_S \geq \sum_{k \in S} x_k - (|S| - 1), && \forall S \in \mathcal{S} \\ & y_S \in \{0,1\}, && \forall S \in \mathcal{S} \\ & x_k \in \{0,1\} && \forall k = 1, \ldots, n \end{aligned}$$

# Standard linearization (SL)

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k + \sum_{i=1}^{n} a_i x_i,$$

## 1. Substitute monomials

$$\begin{aligned}
\min \; & \sum_{S \in \mathcal{S}} a_S y_S + \sum_{i=1}^{n} a_i x_i \\
\text{s.t. } & y_S = \prod_{k \in S} x_k, & \forall S \in \mathcal{S} \\
& y_S \in \{0,1\}, & \forall S \in \mathcal{S} \\
& x_k \in \{0,1\} & \forall k = 1, \dots, n
\end{aligned}$$

## 3. Linear relaxation

$$\begin{aligned}
\min \; & \sum_{S \in \mathcal{S}} a_S y_S + \sum_{i=1}^{n} a_i x_i \\
\text{s.t. } & y_S \leq x_k, & \forall k \in S, \forall S \in \mathcal{S} \\
& y_S \geq \sum_{k \in S} x_k - (|S| - 1), & \forall S \in \mathcal{S} \\
& 0 \leq y_S \leq 1, & \forall S \in \mathcal{S} \\
& 0 \leq x_k \leq 1 & \forall k = 1, \dots, n
\end{aligned}$$

# Linear relaxation

A natural question: does the **standard linearization polytope**

$$P_{SL} = \{(x, y) \in [0, 1]^{n+|\mathcal{S}|} \mid y_S \leq x_k \ \ \forall k \in S, y_S \geq \sum_{k \in S} x_k - (|S| - 1) \ \ \forall S \in \mathcal{S}\}$$

have fractional vertices?

- For a function containing a single nonlinear monomial: No.
- For two or more nonlinear terms, Yes! $P_{SL}$ is in general very weak!!!
- So, when is $P_{SL}$ integral?

# Co-occurrence hypergraph

## Recall: co-occurrence hypergraph

When

$$f(x_1, \ldots, x_n) = \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k + \sum_{i=1}^{n} a_i x_i,$$

$H_f = ([n], \mathcal{S})$ is the *co-occurrence hypergraph* associated with $f$.

## Definition: Berge cycles

For a hypergraph $H = (V, \mathcal{S})$, a **Berge cycle** of length $p$ is a sequence

$$(i_1, S_1, i_2, S_2, \ldots, i_p, S_p, i_1),$$

where

1. $i_1, i_2, \ldots, i_p$ are pairwise distinct vertices of $V$,
2. $S_1, S_2, \ldots, S_p$ are pairwise distinct edges of $\mathcal{S}$,
3. $i_j, i_{j+1} \in S_j$ for $j = 1, \ldots p-1$, and $i_1, i_p \in S_p$.

# Perfect standard linearization

**(E. Rodríguez-Heck, Ch. Buchheim, Y. Crama, 2016)**

$P_{SL}$ is integral if and only if $H_f$ has no Berge cycles.

**Proof:**

- $\Leftarrow$ If $H_f$ is Berge-acyclic then the constraint matrix of $P_{SL}$ is balanced, a property that guarantees integrality.
- $\Rightarrow$ If $H_f$ has a cycle, then construct an objective function that reaches its optimum at a fractional vertex of $P_{SL}$.

# Perfect standard linearization

**(E. Rodríguez-Heck, Ch. Buchheim, Y. Crama, 2016)**

$P_{SL}$ is integral if and only if $H_f$ has no Berge cycles.

- Generalizes a result of Padberg (1989) for quadratic functions.
- Closely related to a result of Crama (1988,1993) for an "irredundant" relaxation of $P_{SL}$.
- Independently obtained by Del Pia and Khajavirad (2016).

# Multilinear optimization in binary variables

$$\text{(MOB)} \quad \min_{x \in \{0,1\}^n} \sum_{S \in \mathcal{S}} a_S \prod_{k \in S} x_k + \sum_{i=1}^{n} a_i x_i$$

- Multilinear optimization is NP-hard, even if $f$ is quadratic.
- Approaches:
  - Direct resolution methods
  - *Linearization*: extensive literature in integer programming.
  - ***Quadratization***: more recent approach.
  - Idea: can we reduce MOB to the (unconstrained) quadratic case rather than to the (constrained) linear case?
  - Yes, in many ways!

# Quadratization

Observations

- Say $g(x, y)$, $(x, y) \in \{0, 1\}^{n+m}$, is a quadratic function.
- Then, for all $x \in \{0, 1\}^n$,

  $$f(x) := \min\{g(x, y) \mid y \in \{0, 1\}^m\}$$

  is a pseudo-Boolean function.
- $f(x)$ may be quadratic, or not.
- $\min\{f(x) \mid x \in \{0, 1\}^n\} = \min\{g(x, y) \mid (x, y) \in \{0, 1\}^{n+m}\}$.
- Conversely...

# Quadratization

## Quadratization

The quadratic function $g(x, y)$, $(x, y) \in \{0, 1\}^{n+m}$ is an *m-quadratization* of the pseudo-Boolean function $f(x)$, $x \in \{0, 1\}^n$, if

$$f(x) = \min\{g(x, y) \,|\, y \in \{0, 1\}^m\} \quad \text{for all } x \in \{0, 1\}^n.$$

The $y$-variables are called *auxiliary* variables.

- $\min\{f(x) \,|\, x \in \{0, 1\}^n\} = \min\{g(x, y) \,|\, (x, y) \in \{0, 1\}^{n+m}\}$.
- Does every function $f$ have a quadratization?

# Existence

**Existence of quadratizations (Rosenberg 1975)**

Given the multilinear expression of a pseudo-Boolean function $f(x), x \in \{0,1\}^n$, one can find in polynomial time a quadratization $g(x,y)$ of $f(x)$.

- Idea: in each term $\prod_{i \in A} x_i$ of $f$, with $\{1,2\} \subseteq A$, replace the product $x_1 x_2$ by a new variable $y$;
- Introduce a penalty term to force $y = x_1 x_2$ in every minimizer of the transformed expression;
- $t(x,y) = \left( \prod_{i \in A \setminus \{1,2\}} x_i \right) y \; + \; M(x_1 x_2 - 2x_1 y - 2x_2 y + 3y).$
- Potential drawbacks: introduces many auxiliary variables, big $M$.

# Questions arising...

- Many quadratization procedures proposed in recent years. Which ones are “best”? Small number of variables, of positive terms, good properties with respect to persistencies, submodularity?
- Easier question: What if *f* is a single monomial?
- How many variables are needed in a quadratization?
- etc.

Refs: Boros and Gruber (2011); Buchheim and Rinaldi (2007); Fix, Gruber, Boros and Zabih (2011): Freedman and Drineas (2005); Ishikawa (2011); Kolmogorov and Zabih (2004); Ramalingam et al. (2011); Rosenberg (1975); Rother et al. (2009); Živný, Cohen and Jeavons (2009); etc.

# Outline

**Focus of our recent work:**

- lower and upper bounds on size of quadratizations
- the case of symmetric functions

M. Anthony, E. Boros, Y. Crama and M. Gruber, Quadratization of symmetric pseudo-Boolean functions, *Discrete Applied Mathematics* 203 (2016) 1–12.

M. Anthony, E. Boros, Y. Crama and M. Gruber, Quadratic reformulations of nonlinear binary optimization problems *Mathematical Programming* 162 (2017) 115-144.

E. Boros, Y. Crama and E. Rodrìguez-Heck, Compact quadratizations for pseudo-Boolean functions, Working paper, 2018.

# General question

- How many auxiliary variables are needed in general?
- Upper bound based on termwise quadratizations:

## Observation

Every term of the form $a \prod_{i=1}^{n} x_i$ can be quadratized using $n-2$ auxiliary variables (Rosenberg 1975), and even $\lfloor \frac{n-1}{2} \rfloor$ auxiliary variables (Ishikawa 2011).

So:

## Ishikawa (2011)

For every $n$-variable pBf, one can find in polynomial time a quadratization involving $\lfloor \frac{n-1}{2} \rfloor 2^n$ auxiliary variables.

- Best known bound, until recently.

# Upper bound

- Upper bound based on termwise quadratizations:

**Ishikawa (2011)**

For every $n$-variable pBf, one can find in polynomial time a quadratization involving at most $\left\lfloor \frac{n-1}{2} \right\rfloor 2^n$ auxiliary variables.

- We prove:

**Theorem: upper bound (*Math. Prog.* (2017))**

For every $n$-variable pBf, one can find in polynomial time a quadratization involving at most $O(2^{n/2})$ auxiliary variables.

# Pairwise cover

Based on a construction using small *pairwise covers:*

**Pairwise cover**

A hypergraph $\mathcal{H}$ is a *pairwise cover* of $\{1, \ldots, n\}$ if, for every $S \subseteq \{1, \ldots, n\}$ with $|S| \geq 3$, there are sets $A, B \in \mathcal{H}$ such that $|A| < |S|$, $|B| < |S|$ and $A \cup B = S$.

We can prove:

**Theorem: From pairwise cover to quadratization**

If there exists a pairwise cover of $\{1, \ldots, n\}$ of size $m$, then every pseudo-Boolean function has an $m$-quadratization.

- Idea of the proof: write $\prod_{i \in S} x_i = (\prod_{j \in A} x_j)(\prod_{k \in B} x_k)$; substitute $y_A$ for $\prod_{j \in A} x_j$ and $y_B$ for $\prod_{k \in B} x_k$;
- Introduce a penalty term to force the correct values of $y_A$ and $y_B$ in every minimizer of the transformed expression.

# Pairwise covers

Thus:

**Theorem: From pairwise cover to quadratization**

If there exists a pairwise cover of $\{1, \ldots, n\}$ of size $m$, then every pseudo-Boolean function has an $m$-quadratization.

- But... there are pairwise covers with size $O(2^{n/2})$.
- Pairwise covers are (almost) identical to so-called 2-*bases* investigated by Erdös, Füredi and Katona (2006), Frein, Lévêque and Sebö (2008), Ellis and Sudakov (2011).
- $\mathcal{P}(even) =$ all subsets of even integers in $\{1, \ldots, n\}$.
- $\mathcal{P}(odd) =$ all subsets of odd integers in $\{1, \ldots, n\}$.
- $\mathcal{H} = \mathcal{P}(even) \cup \mathcal{P}(odd)$ is a “small” pairwise cover with size $O(2^{n/2})$.

# Lower bound

- Any good lower bound on the number of auxiliary variables?

**Theorem: lower bound (*Math. Prog.* (2017))**

There are pseudo-Boolean functions of $n$ variables for which every quadratization must involve at least $\Omega(2^{n/2})$ auxiliary variables.

- This lower bound matches the $O(2^{n/2})$ upper bound.
- Non constructive proof based on dimensionality argument: if too few auxiliary variables, then we cannot generate the whole vector space of pseudo-Boolean functions.

# Conclusions

- Many fruitful connections between functions of Boolean variables, graphs and hypergraphs.
- Many intriguing questions and conjectures.
- See also

**BOOLEAN FUNCTIONS**
**Theory, Algorithms, and Applications**

Yves CRAMA and Peter L. HAMMER
Cambridge University Press, 2011
710 pages

with contributions by C. Benzaken, E. Boros, N. Brauner, M.C. Golumbic, V. Gurvich, L. Hellerstein, T. Ibaraki, A. Kogan, K. Makino, B. Simeone