Graphical Models – Queries, complexity, algorithms and applications

STACS'2020 tutorial

M.C. Cooper[1], S. de Givry[2] & T. Schiex[2]

[1] Université Fédérale de Toulouse, ANITI, IRIT, Toulouse, France
[2] Université Fédérale de Toulouse, ANITI, INRAE MIAT, UR 875, Toulouse, France

Presented by Thomas Schiex

https://doi.org/10.4230/LIPIcs.STACS.2020.4

Informally

A description of a multivariate function as the combination of a set of simple functions.

Informally

A description of a multivariate function as the combination of a set of simple functions.

Propositional logic (CNF aka Conjunctive Normal Form)

A Boolean function of Boolean variables described as the conjunction of disjunction of literals.

What is a graphical model?

ANITI INRAE

Informally

A description of a multivariate function as the combination of a set of simple functions.

Propositional logic (CNF aka Conjunctive Normal Form)

A Boolean function of Boolean variables described as the conjunction of disjunction of literals.

Constraint Networks

A Boolean function of discrete variables described as the conjunction of Boolean tensors.

ANITI INRAE

Informally

A description of a multivariate function as the combination of a set of simple functions.

Propositional logic (CNF aka Conjunctive Normal Form)

A Boolean function of Boolean variables described as the conjunction of disjunction of literals.

Constraint Networks

A Boolean function of discrete variables described as the conjunction of Boolean tensors.

Cost Function Networks

A non negative function of discrete variables described as the sum of non negative tensors.

Informally

A description of a multivariate function as the combination of a set of simple functions.

Propositional logic (CNF aka Conjunctive Normal Form)

A Boolean function of Boolean variables described as the conjunction of disjunction of literals.

Constraint Networks

A Boolean function of discrete variables described as the conjunction of Boolean tensors.

Cost Function Networks

A non negative function of discrete variables described as the sum of non negative tensors.

Discrete Markov Random Fields

A non negative function of discrete variables described as the product of non negative tensors.

Concisely describing complex systems

- Concise: we use a set of *small* functions.
- Complex: the joint function results from the interaction of several small functions.

Example

A digital circuit	value of the output
A Sudoku grid	solution or not
A schedule or a time-table	feasibility, acceptability
A pedigree with genotypes	Mendel consistency, probability
A frequency assignment	interference amount
A 3D molecule	energy, stability

Ideally, we would like to

- Learn them: from a sample [Par+17; PPW18]
- Compute their value: given a variable assignment
- Compute simple statistics:
  - Minimum/Maximum: optimization
  - Average: counting
  - ...

Concise and Complex

Plenty of NP-hard problems.

ANITI INRAE

- Variables: $X, Y, Z, \ldots$, possibly indexed as X_i or just i.
- Domains: D_X for variable X, or D_i for variable X_i.
- Values: $\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{g}, \mathtt{r}, \mathtt{t}, \mathtt{1} \ldots$
- Unknown values: $u, v, w, x, y, z \ldots$.
- Sequence of variables: $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \ldots$
- Sequence of values: acgtgcatggagccacgtcaggta
- Unknown sequence of values: $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \ldots$.
- Domain of a sequence of variables $\boldsymbol{X} : D_{\boldsymbol{X}}$ (Cartesian product of the domains).
- Assignment $\boldsymbol{u}_{\boldsymbol{X}}$: an element of $D_{\boldsymbol{X}}$. Defines an assignment for all the variables in $\boldsymbol{X}$.
- $\boldsymbol{u}_{\boldsymbol{X}}[\boldsymbol{Y}]$ (or $\boldsymbol{u}_{\boldsymbol{Y}}$): *projection* of $\boldsymbol{u}_{\boldsymbol{X}}$ on $\boldsymbol{Y} \subseteq \boldsymbol{X}$ (the sequence of values of $\boldsymbol{Y}$ in $\boldsymbol{u}_{\boldsymbol{X}}$).

A definition parameterized by B and $\bigoplus$

Definition (Graphical Model (GM))

A GM $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$ with co-domain B and combination operator $\oplus$ is defined by:

- a sequence of n variables $\boldsymbol{V}$, each with an associated finite domain of size less than d.
- a set Φ of e functions (or factors).
- Each function $\varphi_{\boldsymbol{S}} \in \Phi$ is a function from $D_{\boldsymbol{S}} \rightarrow B$. $\boldsymbol{S}$ is called the scope of the function and $|\boldsymbol{S}|$ its arity.

Definition (Joint function)

$\mathcal{M}$ defines a joint function:

$$\Phi_{\mathcal{M}}(\boldsymbol{v}) = \bigoplus_{\varphi_{\boldsymbol{S}} \in \Phi} \varphi_{\boldsymbol{S}}(\boldsymbol{v}[\boldsymbol{S}])$$

ANITI INRAE

B

- B is assumed to be totally ordered by $\prec$.
- With a minimum element $\mathbf{0}$ and a maximum element denoted as $\top$.

$\bigoplus$

- Associative, commutative, monotonic. $(\alpha \succeq \beta \Rightarrow (\alpha \oplus \gamma) \succeq (\beta \oplus \gamma))$
- $\mathbf{0}$ as an identity. $(\alpha \oplus \mathbf{0} = \alpha)$
- $\top$ as an absorbing element. $(\alpha \oplus \top = \top)$

Optional

- Idempotency. $(\alpha \oplus \alpha = \alpha)$
- Fairness. $(\forall \beta \preccurlyeq \alpha, \exists \gamma \text{ s.t. } (\beta \oplus \gamma) = \alpha)$
- Denoted as $\gamma = (\alpha \ominus \beta)$ $(\beta \oplus (\alpha \ominus \beta) = \alpha)$

Structure (GM)	B	$a \oplus b$	$\prec$	$\mathbf{0}$	$\top$	Idemp.	$a \ominus b$
Boolean	$\{t, f\}$	$a \wedge b$	$t < f$	t	f	yes	a
Possibilistic	$[0, 1]$	$\max(a, b)$	$<$	0	1	yes	$\max(a, b)$
Additive	$\bar{\mathbb{N}}$	$a + b$	$<$	0	$+\infty$	no	$a - b$
Weighted	$\{0,1,\ldots,k\}$	$\min(k, a+b)$	$<$	0	k	no	$(a=k\ ?\ k\ :\ a-b)$
Probabilistic	$[0, 1]$	$a \times b$	$>$	1	0	no	a/b

Fair countable structures exhaustively analyzed [CS04; Coo05]

- Stack of additive/weighted structures
- Interacting as idempotent structures

How are functions $\varphi_{\boldsymbol{S}} \in \Phi$ represented?

- Default: as tensors over B. (multidimensional tables)
- Boolean vars: (weighted) clauses. (disjunction of literals: variables or their negation)
- Using a specific language, subset of all tensors or clauses or dedicated (All-Different).

This influences complexities

- We assume a constant time $\oplus$ and constant space representation of elements of B.
- We mostly use tensors (universal): $\varphi_{\boldsymbol{S}}$ represented in space $O(d^{|\boldsymbol{S}|})$.

A variety of well-studied frameworks

- Propositional Logic (PL): Boolean domains and co-domain, conjunction of clauses
- Constraint Networks (CN): Finite domains, Boolean co-domain, conjunction of tensors
- Cost Function Networks (CFN): Finite domains, numerical co-domain, sum of tensors.
- Markov Random Fields (MRF): Finite domains, $\mathbb{R}^+$ as co-domain, product of tensors.
- Bayesian Networks (BN): MRF + normalized functions and scopes following a DAG.
- Generalized Additive Independence [BG95], Weighted PL, QPBO [BH02], ILP…

Excluded

- Gaussian Graphical Models or Linear Programming.
- Totally ordered B excludes e.g. Ceteris Paribus networks (CP-nets [Dom+03])

Definition (Equivalence)

Two graphical models $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$ and $\mathcal{M}' = \langle \boldsymbol{V}, \Phi' \rangle$, with the same variables and valuation structure are equivalent iff they define the same joint function:

$$\forall \boldsymbol{v} \in D_{\boldsymbol{V}}, \Phi_{\mathcal{M}}(\boldsymbol{v}) = \Phi_{\mathcal{M}'}(\boldsymbol{v})$$

Definition (Relaxation)

Given two graphical models $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$ and $\mathcal{M}' = \langle \boldsymbol{V}, \Phi' \rangle$, with the same variables and valuation structure, $\mathcal{M}$ is a relaxation of $\mathcal{M}'$ iff

$$\forall \boldsymbol{v} \in D_{\boldsymbol{V}}, \Phi_{\mathcal{M}}(\boldsymbol{v}) \preccurlyeq \Phi_{\mathcal{M}'}(\boldsymbol{v})$$

ANITI INRAE

Definition ((Hyper)graph of $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$)

One vertex per variable, one (hyper)edge per scope $\boldsymbol{S}$ of function $\varphi_{\boldsymbol{S}} \in \Phi$.

Definition (Factor graph of $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$)

The bi-partite incidence graph of the hypergraph above. One vertex per variable or function, an edge connects the vertex $\varphi_{\boldsymbol{s}}$ to all variables in $\boldsymbol{S}$.

Definition (Primal/Moral graph of $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$)

The 2-section of its hypergraph.

Definition (Micro-structure graph of $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$)

Weighted n-partite graph with one vertex per value and a weighted hyper-edge on $\boldsymbol{s} \in D_{\boldsymbol{S}}$ for every $\varphi_{\boldsymbol{S}} \in \Phi$ and $\boldsymbol{s}$ such that $\varphi_{\boldsymbol{S}}(\boldsymbol{s}) \neq \boldsymbol{0}$.

CFN $\mathcal{M} = \langle V, \Phi \rangle$, parameterized by $k = \top$

$\mathcal{M}$ defines a non negative joint function $\Phi_{\mathcal{M}} = \min(\sum_{\varphi_S \in \Phi} \varphi_S, k)$

Flexible

- $k = 1$ same as Constraint Networks
- $k = \infty$ same as GAI, $-\log()$ transform of MRFs
- k finite k is a known upper bound
- $\varphi_\varnothing$ is a naive lower bound on the minimum cost

ANITI INRAE

Optimization queries

- SAT/PL: is the minimum of $\Phi_{\mathcal{M}} = t$?
- CSP/CN: is the minimum of $\Phi_{\mathcal{M}} = t$?
- WCSP/CFN: is the minimum of $\Phi_{\mathcal{M}} \prec \alpha$?
- MAP/MRF: is the minimum of $\Phi_{\mathcal{M}} \prec \alpha$?
- MPE/BN: is the minimum of $\Phi_{\mathcal{M}} \prec \alpha$?

Counting queries

- #-SAT/PL: how many assignments satisfy $\Phi_{\mathcal{M}} = t$?
- MAR/MRF: compute $Z = \sum(\Phi_{\mathcal{M}})$ or $P_{\mathcal{M}}(X = u)$ where $X \in \boldsymbol{V}$
- MAR/BN: compute $P_{\mathcal{M}}(X = u)$ where $X \in \boldsymbol{V}$

ANITI INRAE

Using $\bigotimes$ as a marginalization or elimination operator

$$\bigotimes_{v \in D_V} \left[\bigoplus_{\varphi_S \in \Phi} (\varphi_S(v[S])) \right]$$

$\otimes$ associative, commutative, distributive $\alpha \oplus (\beta \otimes \gamma) = (\alpha \oplus \beta) \otimes (\alpha \oplus \gamma)$

Axioms for dynamic programming

Proposed in similar forms a number of times [BMR97; AM00; KW08; KMP00; GM08], possibly first by Shafer and Shenoy [Sha91].

ANITI INRAE

WCSP/CFN with one variable X_i per vertex i

- Min-Cut: $D_i = \{\mathtt{l}, \mathtt{r}\}, D_s = \{\mathtt{l}\}, D_t = \{\mathtt{r}\}$ $\quad \forall (i,j) \in \boldsymbol{E}, \varphi_{ij} = \mathbf{1}(X_i \neq X_j)$
- Max-Cut: same $\quad \varphi_{ij} = \mathbf{1}(X_i = X_j)$
- Vertex Cover: $D_i = \{\mathtt{a}, \mathtt{r}\}$ $\quad \forall i, \varphi_i = \mathbf{1}(X_i = \mathtt{a}), \forall (i,j) \in \boldsymbol{E}, \varphi_{ij} = \top(X_i = X_j = \mathtt{r})$
- Max-Clique: $D_i = \{\mathtt{a}, \mathtt{r}\}$ $\quad \forall i, \varphi_i = \mathbf{1}(X_i = \mathtt{r}), \forall (i,j) \notin \boldsymbol{E}, \varphi_{ij} = \top(X_i = X_j = \mathtt{a})$
- 3-coloring: $D_i = \{\mathtt{r}, \mathtt{g}, \mathtt{b}\}$ $\quad \forall (i,j) \in \boldsymbol{E}, \varphi_{ij} = \top(X_i = X_j)$
- Min-Sum 3-coloring: $D_i = \{1, 2, 3\}$ $\quad \forall i, \varphi_i(u) = u, \forall (i,j) \in \boldsymbol{E}, \varphi_{ij} = \top(X_i = X_j)$
- ...

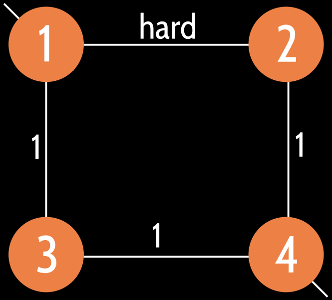
Graph $G = (V, E)$ with edge weight function w

- A boolean variable x_i — per vertex $i \in V$
- A cost function $w_{ij} = w(i, j) \times \mathbb{1}[x_i \neq x_j]$ — per edge $(i, j) \in E$
- Hard edges: $w_{ij} = k$

Graph $G = (V, E)$ with edge weight function w

- A boolean variable x_i — per vertex $i \in V$
- A cost function $w_{ij} = w(i,j) \times \mathbb{1}[x_i \neq x_j]$ — per edge $(i,j) \in E$
- Hard edges: $w_{ij} = k$

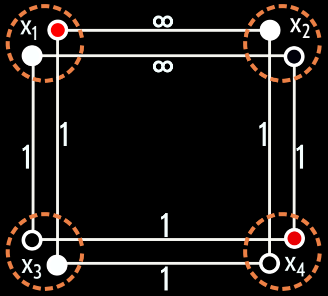
- vertices $\{1, 2, 3, 4\}$
- cut weights 1
- but edge $(1, 2)$ hard



Graph $G = (V, E)$ with edge weight function w

- A boolean variable x_i — per vertex $i \in V$
- A cost function $w_{ij} = w(i,j) \times \mathbb{1}[x_i \neq x_j]$ — per edge $(i,j) \in E$
- Hard edges: $w_{ij} = k$

- vertices $\{1, 2, 3, 4\}$
- cut weights 1
- but edge $(1, 2)$ hard



MinCut on a 3-clique with hard edge

```
{
  problem :{name: MinCut, mustbe: <100.0},
    variables:
      {x1: [l], x2: [l,r], x3: [l,r], x4: [r]}
    functions: {
      cut12:
        {scope: [x1,x2], costs: [0.0, 100.0, 100.0, 0.0]},
      cut13:
        {scope: [x1,x3], costs: [0.0,1.0,1.0,0.0]},
      cut23:
        {scope: [x2,x3], costs: [0.0,1.0,1.0,0.0]}
  ...
}
```

The so called "local polytope" [Sch76; Kos99; Wer07] (w/o last line)

Function $\sum_{i,a} \varphi_i(a) \cdot x_{ia} + \sum_{\substack{\varphi_{ij} \in \Phi \\ a \in D_i, b \in D_j}} \varphi_{ij}(a,b) \cdot y_{iajb}$ such that

$$\sum_{a \in D_i} x_{ia} = 1 \qquad \forall i \in \{1, \ldots, n\}$$

$$\sum_{b \in D_j} y_{iajb} = x_{ia} \qquad \forall \varphi_{ij} \in \Phi, \forall a \in D_i$$

$$\sum_{a \in D_i} y_{iajb} = x_{jb} \qquad \forall \varphi_{ij} \in \Phi, \forall b \in D_j$$

$$x_{ia} \in \{0, 1\} \qquad \forall i \in \{1, \ldots, n\}$$

$nd + e.d^2$ variables. $n + 2ed$ constraints

Conditioning $\varphi_{\boldsymbol{S}}$ by $X = a \quad (X \in \boldsymbol{S})$	Assignment
Let $\boldsymbol{T} = \boldsymbol{S} - \{X\}$, this gives $\varphi_{\boldsymbol{T}}(\boldsymbol{v}) = \varphi_{\boldsymbol{S}}(\boldsymbol{v} \cup \{X = a\})$	Negligible complexity

Combination of $\varphi_{\boldsymbol{S}}$ and $\varphi_{\boldsymbol{S}'}$	Join
$(\varphi_{\boldsymbol{S}} \oplus \varphi_{\boldsymbol{S}'})(\boldsymbol{v}) = \varphi_{\boldsymbol{S}}(\boldsymbol{v}[\boldsymbol{S}]) \oplus \varphi_{\boldsymbol{S}'}(\boldsymbol{v}[\boldsymbol{S}'])$	Space/time $O(d^{\lvert\boldsymbol{S} \cup \boldsymbol{S}'\rvert})$ for tensors

Elimination of $X \in \boldsymbol{S}$ from $\varphi_{\boldsymbol{S}}$	Marginalization/Projection
$\varphi_{\boldsymbol{S}}[-X](\boldsymbol{u}) = \bigotimes_{\boldsymbol{v} \in D_X} \varphi_{\boldsymbol{S}}(\boldsymbol{u} \cup \boldsymbol{v})$	Time $O(d^{\lvert\boldsymbol{S}\rvert})$, space $O(d^{\lvert\boldsymbol{S}\rvert-1})$ for tensors

ANITI INRAE

Tree exploration — Time $O(d^n)$, linear space

- If all $|D_X| = 1, \Phi_{\mathcal{M}}(\boldsymbol{v}), \boldsymbol{v} \in D_{\boldsymbol{V}}$ is the answer
- Else choose $X \in \boldsymbol{V}$ s.t. $|D_X| > 1$ and $u \in D_X$ and reduce to
  1. one query where we condition on $X_i = u$
  2. one where u is removed from D_X
- The result of these queries is combined using $\otimes$

Optimization ($\otimes = \min$) — Branch and Bound

If a lower bound on the current query is $\succeq$ a known upper bound on $\Phi_{\mathcal{M}}$... Prune!

NB: $\varphi_{\varnothing}$ is always a lower bound.

Variable ordering

Drastic empirical effects on efficiency.

Definition (Message sent by variable X)

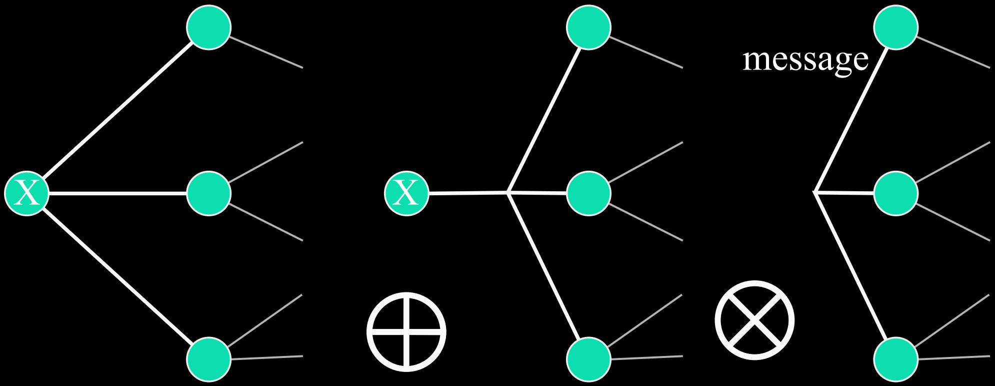
Let $X \in V$, and Φ^X be the set $\{\varphi_{\boldsymbol{S}} \in \Phi \text{ s.t. } X \in \boldsymbol{S}\}$, $\boldsymbol{T}$, the neighbors of X.

The message $m_{\boldsymbol{T}}^{\Phi_X}$ from Φ^X to $\boldsymbol{T}$ is:

$$m_{\boldsymbol{T}}^{\Phi_X} = \Big(\bigoplus_{\varphi_{\boldsymbol{S}} \in \Phi^X} \varphi_{\boldsymbol{S}} \Big)[-X] \qquad (1)$$

Eliminating a variable — Distributivity

$$\bigotimes_{\boldsymbol{v} \in D_{\boldsymbol{V}}} \left[\bigoplus_{\varphi_{\boldsymbol{S}} \in \Phi} (\varphi_{\boldsymbol{S}}(\boldsymbol{v}[\boldsymbol{S}])) \right] = \bigotimes_{\boldsymbol{v} \in D_{\boldsymbol{V}-\{X\}}} \left[\bigoplus_{\varphi_{\boldsymbol{S}} \in \Phi - \Phi^X \cup \{m_{\boldsymbol{T}}^{\Phi_X}\}} (\varphi_{\boldsymbol{S}}(\boldsymbol{v}[\boldsymbol{S}])) \right]$$


X
X
message

Complexity of one elimination for tensors

Computing $m_{\boldsymbol{T}}^{X}$ is $O(d^{|T+1|})$ time, $O(d^{|T|})$ space $|T|$ is the degree of X

The overall complexity is dominated by the largest degree encountered during elimination

Clauses $\boldsymbol{L}, \boldsymbol{L}'$ clauses

If $\Phi^X = \{(X \vee \boldsymbol{L}), (\neg X \vee \boldsymbol{L'})\}$ $m_{\boldsymbol{T}}^{\Phi X}$ is $(\boldsymbol{L} \vee \boldsymbol{L'})$.

The resolution principle [Rob65] is an efficient variable elimination process [DR94; DP60].

Dimension	induced/tree-width
Dimension of an elimination order for G	Largest set $\|T\|$ encountered
Dimension of G	minimum Dimension over all orders

Introduced in 1969 by Bertelé and Brioschi [BB69b; BB69a] (cited 19 and 31 times on GS)
Proved to be equivalent to tree-width by Bodlaender [Bod98].

The secondary optimization problem — Min degree, Minfill, MCS [Ros70]

Finding an optimal order is NP-hard, but useful heuristics exist [BK08].

Tractability

First tractable class for our general query: GMs with bounded tree-width.

Computing marginals — Stochastic Graphical Models

We want $P(X), \forall X \in \boldsymbol{V}$ — Counting

One variable X_i

- Root in X_i and eliminate all variables but X_i, from leaves.
- The elimination of X_i produces a message m_j^i involving just X_j.

All variables — Variables preserved, time & space $O(ed^2)$

Messages are kept as auxiliary functions.

- When a variable X_i has received messages from all its neighbors but one (X_j)
- Send message m_j^i to X_j

$$m_j^i = \bigotimes_{X_i}(\varphi_i \oplus \varphi_{ij} \bigoplus_{X_o \in neigh(X_i), o \neq j} m_i^o) \quad (2)$$

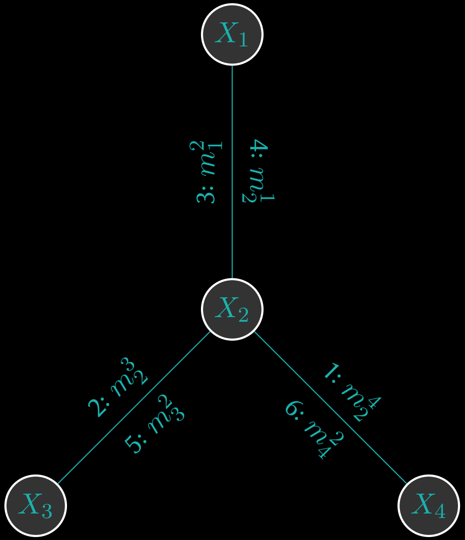


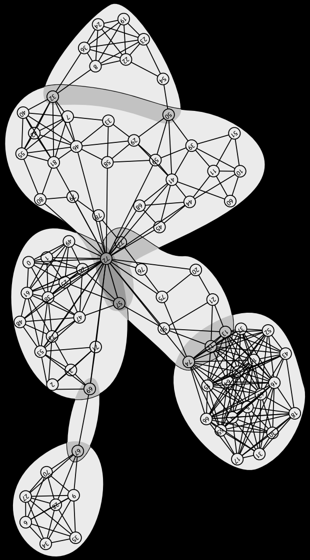
Figure 1: Message passing on a tree, a possible message schedule

The exact approach

Find a (good) tree decomposition and use the previous algorithms on the resulting tree.

Properties

- Space complexity exponential in the separator size only $\theta(d^s)$
- Many variants: block-by-block elimination [BB72], Cluster/Join tree elimination [LS88; DP89],...



ANITI INRAE

The heuristic approach

Starting from e.g., empty messages, apply the message passing equation (2)

$$m_j^i = \bigotimes_{X_i} (\varphi_i \oplus \varphi_{ij} \bigoplus_{X_o \in neigh(X_i), o \neq j} m_i^o)$$

on each function until quiescence or maximum number of iterations (synchronous or asynchronous update schemes exist).

Loopy Belief Propagation [Pea88]

- At the core of Turbo-decoding [BGT93], implemented in all cell phones.
- Widely studied [YFW01], but known to not always converge.
- Often denoted as the "max-sum/min-sum/sum-prod" algorithm.

Assume $\oplus$ is idempotent

If $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$ is a relaxation of $\mathcal{M}' = \langle \boldsymbol{V}, \Phi' \rangle$ then $\mathcal{M}'' = \langle \boldsymbol{V}, \Phi \cup \Phi' \rangle$ is equivalent to $\mathcal{M}'$.

Property

If $\otimes = \min$, any message $m_{\boldsymbol{T}}^{X}$ computed by elimination is a relaxation of Φ^{X} and hence of $\mathcal{M}$.

Equivalence preserving messages

- $\min - \max$ messages can be directly added to the processed graphical model
- This preserves the joint function (equivalence, so for counting too)
- Applies to Boolean, possibilistic and fuzzy structures

Variable elimination/ Resolution based

- Using variable elimination messages: David and Putnam algorithm [DP60] aka Directional Resolution [DR94].
- Using all possible messages: saturation by Resolution [Rob65].

ANITI INRAE

Definition (Arc consistency (closure property))

A graphical model $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$ with idempotent $\oplus$ is arc-consistent iff every variable $X \in \boldsymbol{V}$ is arc consistent w.r.t. every function $\varphi_{\boldsymbol{S}}$ s.t. $X \in \boldsymbol{S}$.

A variable X_i is arc-consistent w.r.t. a function φ_{ij} iff the message m_i^j is a relaxation of φ_i.

Arc consistency (filtering)

A graphical model $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$ with idempotent $\oplus$ can be transformed in polynomial time in a unique equivalent arc consistent graphical model.

Local consistency provides an incremental lower bound on consistency

If the equivalent Arc Consistent graphical model has an empty domain ($\forall a \in D_i, \varphi_i(a) = \top$), then it is infeasible/inconsistent.

Arc consistency filtering is achieved by Loopy BP

- AC-3 [Mac77] is time $O(ed^3)$, space $O(ed)$,
- AC-4 [MH86] is time $O(ed^2)$, space $O(ed^2)$,
- AC-6 [Bes94] is $O(ed^2)$, space $O(ed)$,
- AC2001/3.1 [BR01; ZY01], also optimal, empirically faster and far simpler to implement.

ANITI INRAE

Obvious issue

Without idempotency, messages can not be included in the graphical model without loosing equivalence, hence practical significance.

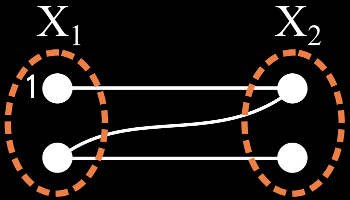
Equivalence Preserving Transformations with $\ominus$

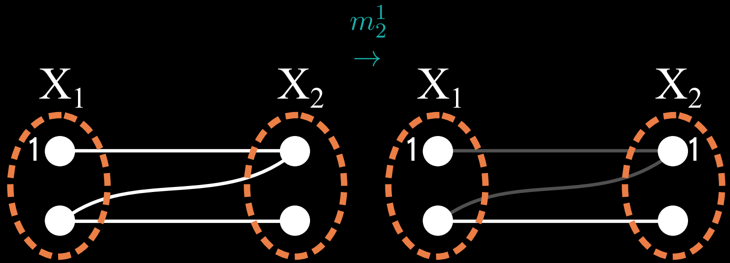
- Consider a set of functions $\Psi \subset \Phi$ and the message $m^{\Psi}_{\boldsymbol{Y}}$
- Replace Ψ by $((\bigoplus_{\varphi_{\boldsymbol{S}} \in \Psi} \varphi_{\boldsymbol{S}}) \ominus m^{\Psi}_{\boldsymbol{Y}})$ and $m^{\Psi}_{\boldsymbol{Y}}$

Any relaxation of $m^{\Psi}_{\boldsymbol{Y}}$ can be used instead.

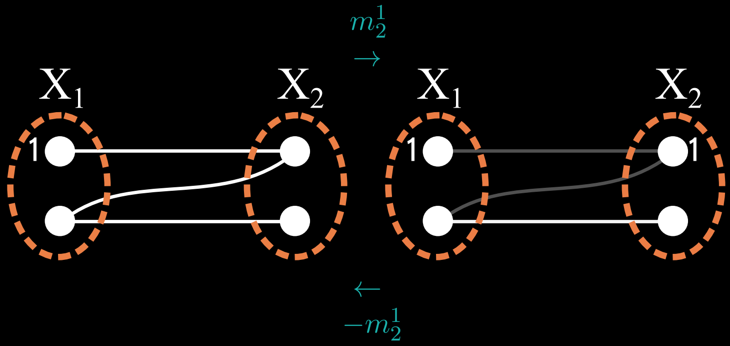
Scope preserving EPTs for tensors — Not for clauses!

If Ψ contains at most one non unary function and $|\boldsymbol{Y}| = 1$ (MRFs: reparametrizations).

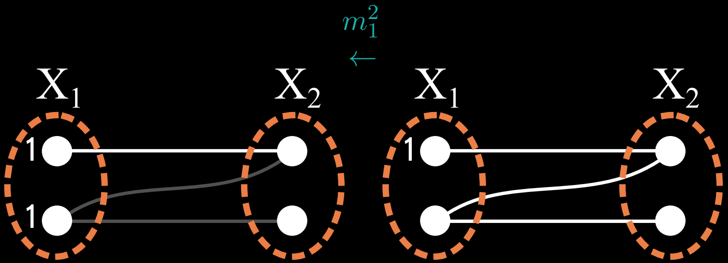

X1
X2
1

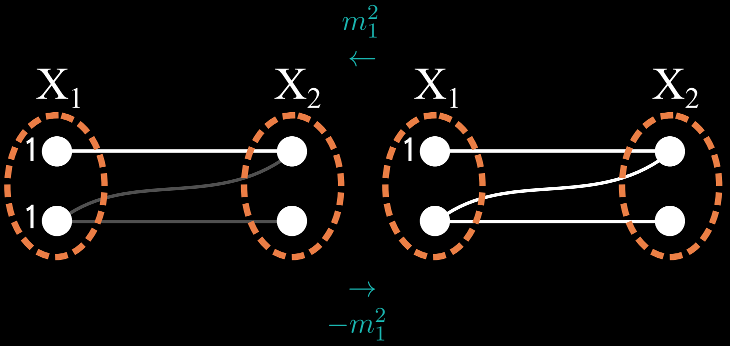

m_2^1
$\rightarrow$
X_1
X_2
1
X_1
X_2
1
1

A small example that may increase $\varphi_\varnothing$



A small example that may increase $\varphi_\varnothing$

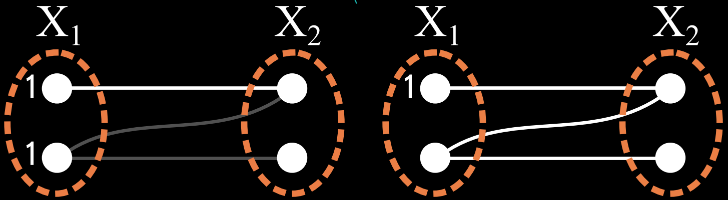



m_1^2
$\leftarrow$
X_1
X_2
1
1
X_1
X_2
1
$\rightarrow$
$-m_1^2$

A small example that may increase $\varphi_\varnothing$

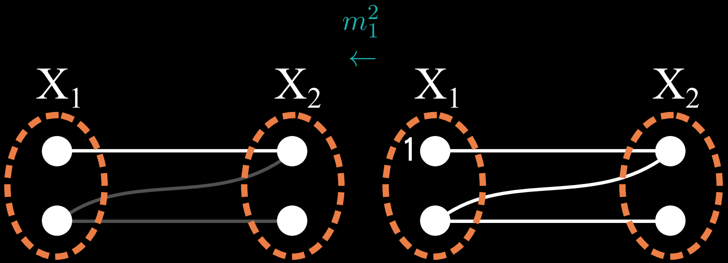
m_1^2

$\leftarrow$



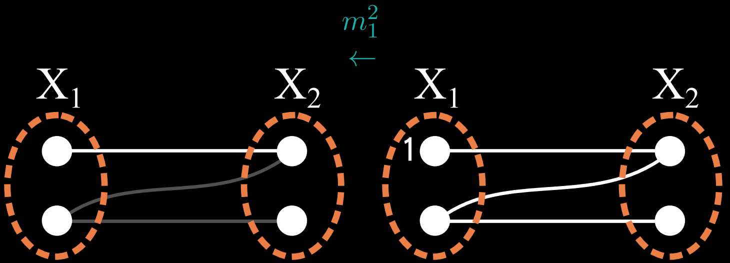
$\Downarrow \quad m_\varnothing^1$

A small example that may increase $\varphi_\varnothing$



$\Downarrow$ $m_\varnothing^1$

$\varphi_\varnothing = 1$



$\Downarrow \quad m_\varnothing^1$

$\varphi_\varnothing = 1$

(Loss of) properties

Preserves equivalence but fixpoints may be non unique (or not guaranteed to exist for some $\Psi / \boldsymbol{Y}$ configurations).

Sequence of integer EPTs

Computing a sequence of integer EPTs that maximizes $\varphi_\varnothing$ is decision NP-complete [CS04].

Set of rational EPTs (OSAC [Sch76; Coo07; Wer07])

Computing a set of rational EPTs maximizing $\varphi_\varnothing$ is in P, solvable by Linear Prog. + AC.

Essentially reduces to solving the dual of the local polytope (+ managing constraints with AC).

Universality of the Local Polytope [PW15]

Any (reasonable) LP can be reduced in linear time to a graphical model whose local polytope has the same optimum as the LP (constructive proof).

OSAC: associated polynomial classes — Empirically slow

- Tree-structured problems
- Submodular problems

Definition (Submodular function over ordered domains)

φ_S submodular if $\forall \boldsymbol{u}, \boldsymbol{v} \in D_S, \varphi_S(\min(\boldsymbol{u}, \boldsymbol{v})) + \varphi_S(\max(\boldsymbol{u}, \boldsymbol{v})) \leq \varphi_S(\boldsymbol{u}) + \varphi(\boldsymbol{v})$

Definition ($\text{Bool}(\varphi_S)$[Coo+08; Coo+10])

$\text{Bool}(\varphi_S)(\boldsymbol{u})$ is $\mathbf{0}$ iff $\varphi_S(\boldsymbol{u}) = \mathbf{0}$.

Definition ($\text{Bool}(\mathcal{M})$[Coo+08; Coo+10])

Given a weighted GM (CFN) $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$, the constraint network

$$\text{Bool}(\mathcal{M}) = \langle \boldsymbol{V}, \{\text{Bool}(\varphi_S) \text{ such that } |\boldsymbol{S}| > 0\} \rangle$$

Definition (Virtual Arc Consistency (VAC)[Coo+08])

A weighted GM $\mathcal{M} = \langle \boldsymbol{V}, \Phi \rangle$ is Virtual Arc Consistent iff enforcing AC on $\text{Bool}(\mathcal{M})$ does not prove inconsistency.

Algorithm loop sketch $O(ed^2k/\varepsilon)$

- Enforce AC on $\mathrm{Bool}(\mathcal{M})$
- If not proved inconsistent, done
- Extract a minimal set of messages proving inconsistency
- Apply these as EPTs on $\mathcal{M}$ (with suitable costs)
- This is guaranteed to increase $\varphi_\varnothing$

Related work

- Convergent MP in MRFs (same family of fixpoints) [Kol06; Kol15]
- Reduces to MaxFlow in the Boolean variable case
- Produces the *roof-dual* lower bound of QPBO [BH02]

Presentation Outline

ANITI INRAE

Combines — Time $O(exp(n))$

- Branch and Bound (aka Backtrack in the Boolean case)
- Incremental Local Consistency enforcing at each node (lower bound)

Variable (and value) ordering heuristics

- Crucial for empirical efficiency
- Are now adaptive (learned while searching) [Mos+01; Bou+04]
- Little theory if any.

ANITI INRAE

Additional ingredients

- Search strategies: Best/Depth First [All+15], restarts [GSC97]
- Stronger preprocessing at the root node
- Dominance analysis [Fre91; DPO13; All+14], ...

Learning from conflicts (Boolean) [Bie+09]

Extracts an informative relaxation at dead-ends using resolution (non serial DP).
Led to CDCL solvers, obsoleted DPLL (Davis, Putnam, Logemann, Loveland [DLL62]).

The power of learning [AFT11; JP12]

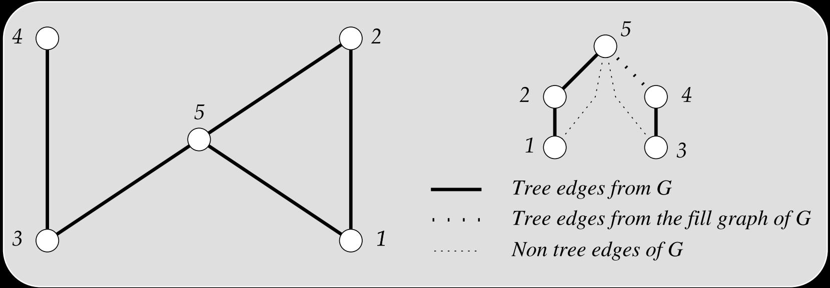
A randomized CDCL solver can decide the consistency of any pairwise CN instance with treewidth w with $O(n^{2w}d^{2w})$ restarts.

Pseudo-tree [Fre85; Sch99]

A pseudo-tree arrangement of a graph G is a rooted tree with the same vertices as G and the property that adjacent vertices in G reside in the same branch of the tree.

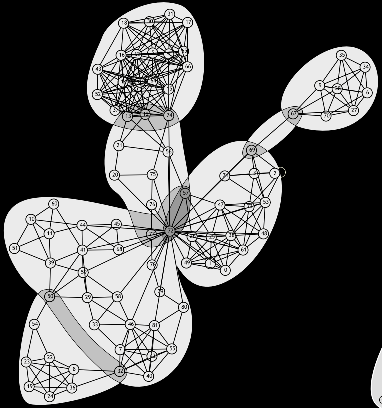
Pseudo-tree [Fre85; Sch99]

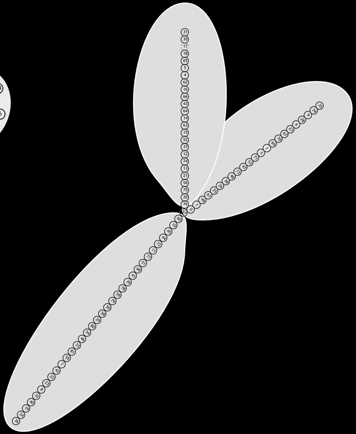
A pseudo-tree arrangement of a graph G is a rooted tree with the same vertices as G and the property that adjacent vertices in G reside in the same branch of the tree.



Pseudo-tree [Fre85; Sch99]

A pseudo-tree arrangement of a graph G is a rooted tree with the same vertices as G and the property that adjacent vertices in G reside in the same branch of the tree.





ANITI INRAE

Pseudo-tree search [Fre85]

- Solve using tree search, assigning variables from the root of the pseudo tree downwards.
- Split resolution when several connected components appear
- space efficient, time $O(exp(h))$

Pseudo-tree height h [Fre85; Sch99] $\equiv$ tree-depth [ND06]

The pseudo-tree height of G is the minimum, over all pseudo-tree arrangements of G of the height of the pseudo-tree arrangement.

Pruning using lower bounds

- AND/OR search uses mini-buckets [MD05]
- BTD uses Arc Consistency [JT03] hyper-treewidth for free [JNT08]

Caching subproblem optima (same separator assignment) time $O(exp(w))$

- AND/OR graph search [MD09]
- Backtrack with tree decompositions (BTD) [JT03; TJ03]

A difficult marriage

- Tree-decompositions constrain the variable ordering
- Variable ordering heuristics crucial for tree search

ANITI INRAE

Languages

- Boolean: A P/NP-complete dichotomy for the CSP [Bul17; Zhu17]
- Additive: the CSP dichotomy implies dichotomy for the additive case [KKR17].
- Submodularity: $\min$ and $\max$ can be replaced by any commutative, conservative functions [CCJ08].
- Finite costs: tight connection with LP [TZ16].

Hybrid tractable class — Joint Winner Property

A binary CFN satisfies the JWP iff for any three variable-value assignment, the multi-set of pairwise costs has not a unique minimum. Related to M-convex functions [TZ16].

Presentation Outline

No universal exact solver

SAT solvers: verification[1], planification, diagnosis, theorem proving,…

[1] Small neural nets too.

[2] Oliver Kullmann. “The Science of Brute Force”. In: *Communications of the ACM* (2017).

ANITI INRAE

No universal exact solver

SAT solvers: verification[1], planification, diagnosis, theorem proving,…

2017: proving an "alien" theorem? ∞

When one splits $\mathbb{N}$ in 2, one part must contain a Pythagorean triple $(a^2 = b^2 + c^2)$

[1] Small neural nets too.

[2] Oliver Kullmann. "The Science of Brute Force". In: *Communications of the ACM* (2017).

No universal exact solver

SAT solvers: verification[1], planification, diagnosis, theorem proving,…

2017: proving an "alien" theorem? ∞

When one splits $\mathbb{N}$ in 2, one part must contain a Pythagorean triple $(a^2 = b^2 + c^2)$

No known proof, puzzled mathematicians for decades (one offered a 100 $ reward)

[1] Small neural nets too.

[2] Oliver Kullmann. "The Science of Brute Force". In: *Communications of the ACM* (2017).

ANITI INRAE

No universal exact solver

SAT solvers: verification[1], planification, diagnosis, theorem proving,...

2017: proving an "alien" theorem? ∞

When one splits $\mathbb{N}$ in 2, one part must contain a Pythagorean triple $(a^2 = b^2 + c^2)$

No known proof, puzzled mathematicians for decades (one offered a 100 $ reward)

SAT solver proof[HKM16; Lam16]

$200TB$ proof, compressed to $86GB$ (stronger proof system)[2]

[1]Small neural nets too.

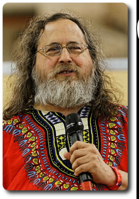
[2]Oliver Kullmann. "The Science of Brute Force". In: *Communications of the ACM* (2017).

Whether it's maths or not… Size matters!

- Not only there exists true unprovable statements (in powerful enough consistent sets of axioms[Göd31])
- There may be true provable statements we will never be able to prove because of their extremely long proofs[Kul17]

A lot of free data and free code...

- International competitions ($> 50,000$ benchmarks with many real problems)
- Open source solvers (autocatalytic)



Different application areas

- CP solvers: resource management in time and or space (eg. scheduling)
- MRFs: image processing (huge problems: heuristics or primal/dual approaches, OpenGM2 [And+10], graph-cuts)
- CFNs: NLP, Computational biology, music composition, resource management (toulbar2 [Hur+16])

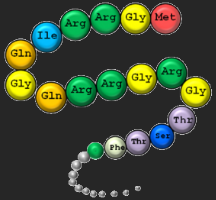
Kind words from OpenGM2 developpers

"ToulBar2 variants were superior to CPLEX variants in all our tests"[HSS18]

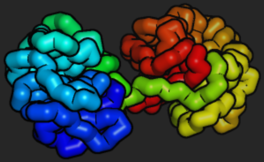
Most active molecules of life

Sequence of amino acids, 20 natural ones each defined by a specific flexible side-chain

Folding



→



→ Function

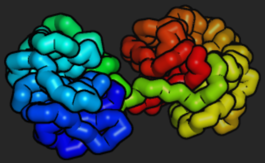
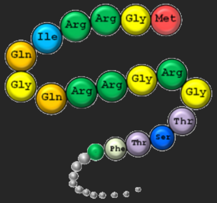
Transporter, binder/regulator, motor, catalyst…
Hemoglobine, TAL effector, ATPase, dehydrogenases…

Protein Design

Most active molecules of life

Sequence of amino acids, 20 natural ones each defined by a specific flexible side-chain

Inverse folding

Function →  → 

Transporter, binder/regulator, motor, catalyst…
Hemoglobine, TAL effector, ATPase, dehydrogenases…

Eco-friendly chemical/structural nano-agents

- Biodegradable (have been mass produced for billions of year)

Eco-friendly chemical/structural nano-agents

- Biodegradable (have been mass produced for billions of year)
- "Easy" to produce (transformed E. coli)

ANITI INRAE

Eco-friendly chemical/structural nano-agents

- Biodegradable (have been mass produced for billions of year)
- “Easy” to produce (transformed E. coli)
- Useful for health, green chemistry[Röt+08] (biofuels, plastic recycling, food and feed, cosmetics…), nanotechnologies[Nog+19]…

ANITI INRAE

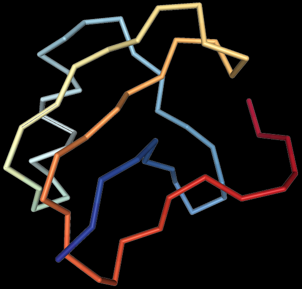
Eco-friendly chemical/structural nano-agents

- Biodegradable (have been mass produced for billions of year)
- "Easy" to produce (transformed E. coli)
- Useful for health, green chemistry[Röt+08] (biofuels, plastic recycling, food and feed, cosmetics…), nanotechnologies[Nog+19]…

20^n sequences! intractable for experimental techniques

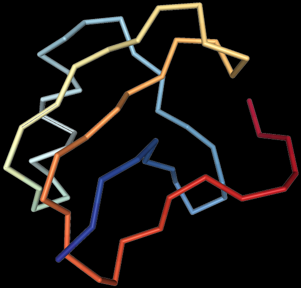
Molecular modeling

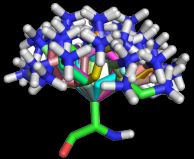
- Full atom model of a protein backbone (assumed to be rigid)



Molecular modeling

- Full atom model of a protein backbone (assumed to be rigid)
- Catalog of all side-chains in different conformations (≈ 400 overall)





Molecular modeling

- Full atom model of a protein backbone (assumed to be rigid)
- Catalog of all side-chains in different conformations ($\approx$ 400 overall)
- Sequence-conformation space: 400^n (or more)

Molecular modeling

- Full atom model of a protein backbone (assumed to be rigid)
- Catalog of all side-chains in different conformations ($\approx$ 400 overall)
- Sequence-conformation space: 400^n (or more)
- Approximate decomposable energy function (intermolecular force field)

Central problem	(plenty of tricky/harder variants)
Maximum stability $\equiv$ Minimum energy	NP-hard[PW02]

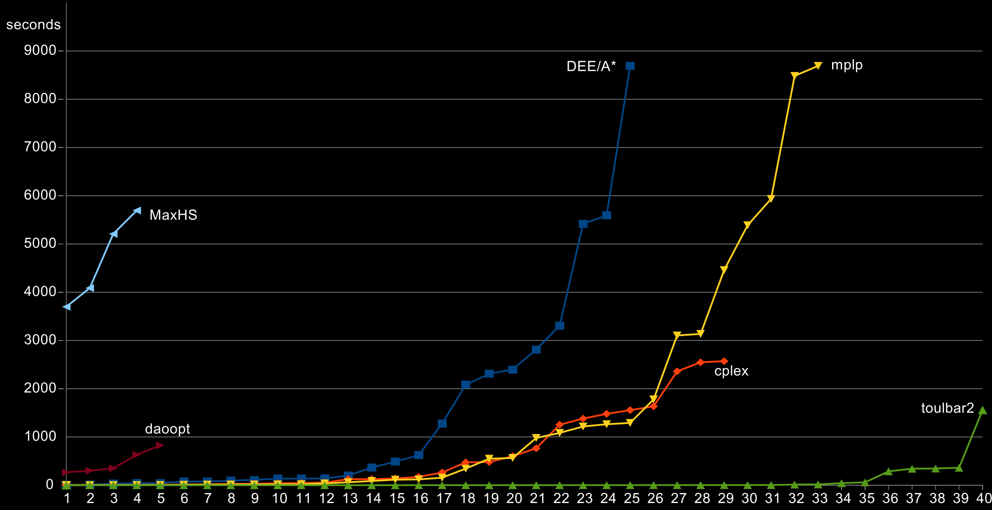
Central problem (plenty of tricky/harder variants)

Maximum stability ≡ Minimum energy NP-hard[PW02]

As a Cost Function Network[Tra+13; All+14]

- One variable per position in the protein sequence
- Domain: catalog of a few hundred amino acids conformations
- Functions: decomposed energy (pairwise terms)
- Treewidth may be less than n (depends on the protein shape)
- Empirically, functions are not permutated submodular

Toulbar2 vs. CPLEX, MaxHS...(real instances)



# of instances solved (X) within a per instance cpu-time limit (Y)

VAC vs. LP on Protein design problems

ANITI INRAE

CPLEX V12.4.0.0

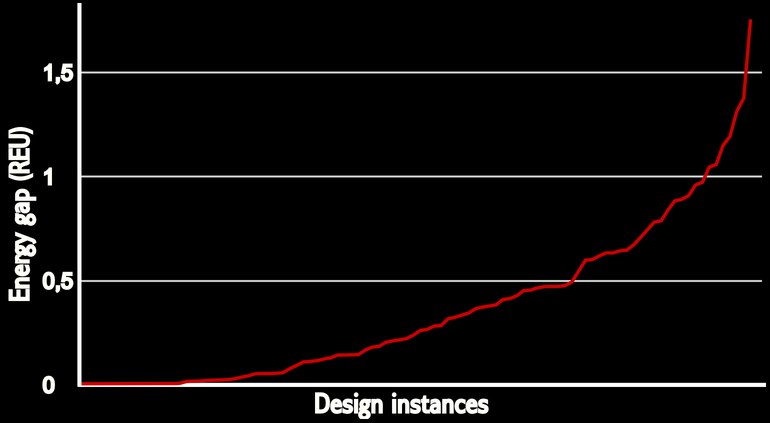
```
Problem '3e4h.LP' read.
Root relaxation solution time =  811.28 sec.
...
MIP - Integer optimal solution:  Objective =  150023297067
Solution time =  864.39 sec.
```

tb2 and VAC (AC3 based)

```
loading CFN file: 3e4h.wcsp
Lb after VAC: 150023297067
Preprocessing time: 9.13 seconds.
Optimum: 150023297067 in 129 backtracks, 129 nodes and 9.38 seconds.
```

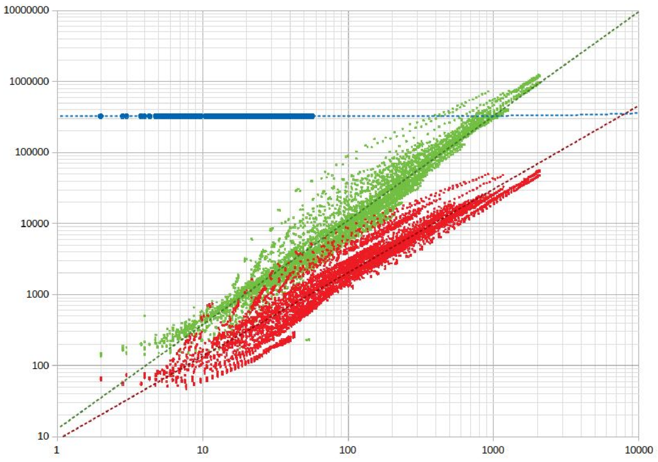
Could this be useful for ILP?

Reversing Prusa-Werner construction somehow?



Optimality gap of the Simulated annealing solution as problems get harder

Asymptotic convergence, close to infinity is arbitrarily far

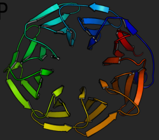


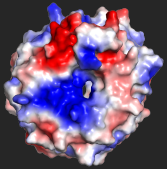
Exact vs. heuristic solvers [Mul+19]

DWave within 1.16 kcal/mol of the optimum 10% of the time, 4.35 kcal/mol 50% of the time, 8.45 kcal/mol 90% of the time.

C8 pseudo-symetric 2OVP symmetrized into a nano-component

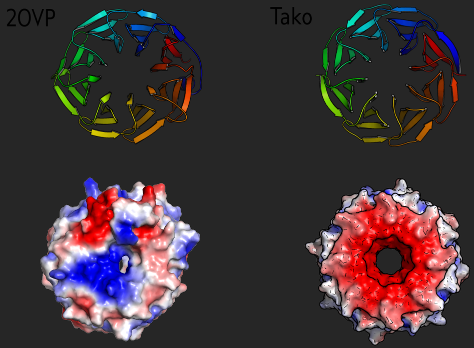
2OVP





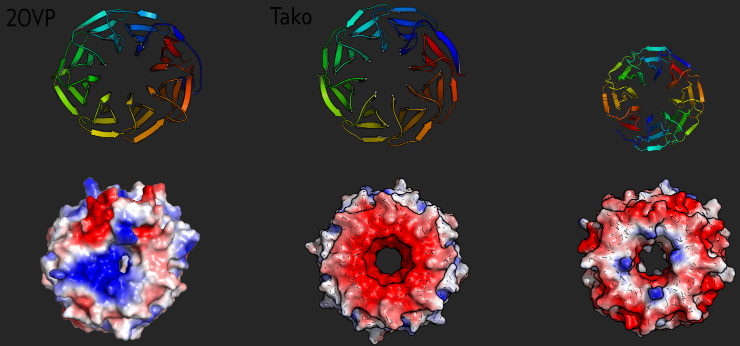
C8 pseudo-symetric 2OVP symmetrized into a nano-component

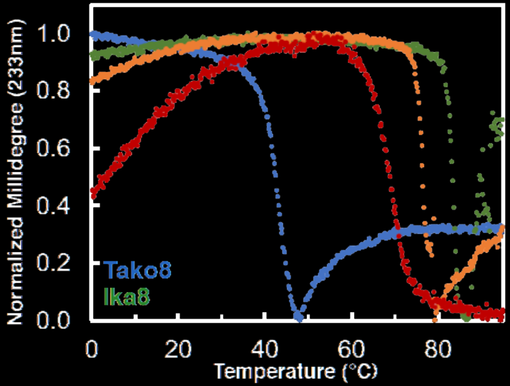
-  Tako: (R)evolution + Rosetta/talaris14 — 8 fold



C8 pseudo-symetric 2OVP symmetrized into a nano-component

-  Tako: (R)evolution + Rosetta/talaris14 — 8 fold
-  Ika: toulbar2 + talaris14 — 4 fold





Compares Tako and Ika structural stability as temperature increases (circular dichroism)

Thank You!

Questions?

Albert Atserias, Andrei Bulatov, and Victor Dalmau. “On the power of k-consistency”. In: *International Colloquium on Automata, Languages, and Programming*. Springer. 2007, pp. 279–290.

Albert Atserias, Johannes Klaus Fichte, and Marc Thurley. “Clause-learning algorithms with many restarts and bounded-width resolution”. In: *Journal of Artificial Intelligence Research* 40 (2011), pp. 353–373.

David Allouche et al. “Computational protein design as an optimization problem”. In: *Artificial Intelligence* 212 (2014), pp. 59–79.

David Allouche et al. “Anytime Hybrid Best-First Search with Tree Decomposition for Weighted CSP”. In: *Principles and Practice of Constraint Programming*. Springer. 2015, pp. 12–29.

Srinivas M Aji and Robert J McEliece. “The generalized distributive law”. In: *IEEE transactions on Information Theory* 46.2 (2000), pp. 325–343.

Björn Andres et al. “An empirical comparison of inference algorithms for graphical models with higher order factors using OpenGM”. In: *Joint Pattern Recognition Symposium*. Springer. 2010, pp. 353–362.

Umberto Bertele and Francesco Brioschi. “A new algorithm for the solution of the secondary optimization problem in non-serial dynamic programming”. In: *Journal of Mathematical Analysis and Applications* 27.3 (1969), pp. 565–574.

Umberto Bertele and Francesco Brioschi. "Contribution to nonserial dynamic programming". In: *Journal of Mathematical Analysis and Applications* 28.2 (1969), pp. 313–325.

Umberto Bertelé and Francesco Brioshi. *Nonserial Dynamic Programming*. Academic Press, 1972.

Christian Bessière. "Arc-Consistency and Arc-Consistency Again". In: *Artificial Intelligence* 65 (1994), pp. 179–190.

Fahiem Bacchus and Adam Grove. "Graphical models for preference and utility". In: *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc. 1995, pp. 3–10.

Claude Berrou, Alain Glavieux, and Punya Thitimajshima. "Near Shannon limit error-correcting coding and decoding: Turbo-codes. 1". In: *Proceedings of ICC'93-IEEE International Conference on Communications*. Vol. 2. IEEE. 1993, pp. 1064–1070.

E. Boros and P. Hammer. "Pseudo-Boolean Optimization". In: *Discrete Appl. Math.* 123 (2002), pp. 155–225.

Armin Biere et al. “Conflict-driven clause learning sat solvers”. In: *Handbook of Satisfiability, Frontiers in Artificial Intelligence and Applications* (2009), pp. 131–153.

H L Bodlaender and A M C A Koster. *Treewidth Computations I. Upper Bounds.* Tech. rep. UU-CS-2008-032. Utrecht, The Netherlands: Utrecht University, Department of Information and Computing Sciences, Sept. 2008. URL: http://www.cs.uu.nl/research/techreps/repo/CS-2008/2008-032.pdf.

Stefano Bistarelli, Ugo Montanari, and Francesca Rossi. “Semiring-based constraint satisfaction and optimization”. In: *Journal of the ACM (JACM)* 44.2 (1997), pp. 201–236.

Hans L Bodlaender. “A partial k-arboretum of graphs with bounded treewidth”. In: *Theoretical computer science* 209.1-2 (1998), pp. 1–45.

Frédéric Boussemart et al. “Boosting systematic search by weighting constraints”. In: *ECAI*. Vol. 16. 2004, p. 146.

C. Bessière and J-C. Régin. “Refining the basic constraint propagation algorithm”. In: *Proc. IJCAI'2001*. 2001, pp. 309–315.

Andrei A. Bulatov. “A Dichotomy Theorem for Nonuniform CSPs”. In: *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*. Ed. by Chris Umans. IEEE Computer Society, 2017, pp. 319–330. ISBN: 978-1-5386-3464-6. DOI: 10.1109/FOCS.2017.37. URL: https://doi.org/10.1109/FOCS.2017.37.

David A. Cohen, Martin C. Cooper, and Peter Jeavons. “Generalising submodularity and Horn clauses: Tractable optimization problems defined by tournament pair multimorphisms”. In: *Theor. Comput. Sci.* 401.1-3 (2008), pp. 36–51. DOI: 10.1016/j.tcs.2008.03.015. URL: https://doi.org/10.1016/j.tcs.2008.03.015.

Martin C Cooper et al. “Virtual Arc Consistency for Weighted CSP”. In: *AAAI*. Vol. 8. 2008, pp. 253–258.

M. Cooper et al. “Soft arc consistency revisited”. In: *Artificial Intelligence* 174 (2010), pp. 449–478.

M C. Cooper. “High-Order Consistency in Valued Constraint Satisfaction”. In: *Constraints* 10 (2005), pp. 283–305.

M C. Cooper. “On the minimization of locally-defined submodular functions”. In: *Constraints* (2007). To appear.

M C. Cooper. “An Optimal k-Consistency Algorithm”. In: *Artificial Intelligence* 41 (1989), pp. 89–95.

M C. Cooper and T. Schiex. “Arc consistency for soft constraints”. In: *Artificial Intelligence* 154.1-2 (2004), pp. 199–227.

Rina Dechter. “Bucket Elimination: A Unifying Framework for Reasoning”. In: *Artificial Intelligence* 113.1–2 (1999), pp. 41–85.

Martin Davis, George Logemann, and Donald Loveland. “A machine program for theorem-proving”. In: *Communications of the ACM* 5.7 (1962), pp. 394–397.

C. Domshlak et al. “Reasoning about soft constraints and conditional preferences: complexity results and approximation techniques”. In: *Proc. of the 18th IJCAI*. Acapulco, Mexico, 2003, pp. 215–220.

Martin Davis and Hilary Putnam. “A computing procedure for quantification theory”. In: *Journal of the ACM (JACM)* 7.3 (1960), pp. 201–215.

Rina Dechter and Judea Pearl. “Tree Clustering for Constraint Networks”. In: *AI* 38 (1989), pp. 353–366.

Simon De Givry, Steven D Prestwich, and Barry O’Sullivan. “Dead-end elimination for weighted CSP”. In: *Principles and Practice of Constraint Programming*. Springer. 2013, pp. 263–272.

Rina Dechter and Irina Rish. "Directional resolution: The Davis-Putnam procedure, revisited". In: *KR* 94 (1994), pp. 134–145.

Eugene C. Freuder. "A sufficient Condition for Backtrack-Bounded Search". In: *Journal of the ACM* 32.14 (1985), pp. 755–761.

Eugene C. Freuder. "Eliminating Interchangeable Values in Constraint Satisfaction Problems". In: *Proc. of AAAI'91*. Anaheim, CA, 1991, pp. 227–233.

Michel Gondran and Michel Minoux. *Graphs, dioids and semirings: new models and algorithms*. Vol. 41. Springer Science & Business Media, 2008.

Kurt Gödel. "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I". In: *Monatshefte für mathematik und physik* 38.1 (1931), pp. 173–198.

Carla P Gomes, Bart Selman, and Nuno Crato. "Heavy-tailed distributions in combinatorial search". In: *International Conference on Principles and Practice of Constraint Programming*. Springer. 1997, pp. 121–135.

Marijn JH Heule, Oliver Kullmann, and Victor W Marek. "Solving and verifying the boolean pythagorean triples problem via cube-and-conquer". In: *International Conference on Theory and Applications of Satisfiability Testing*. Springer. 2016, pp. 228–245.

Stefan Haller, Paul Swoboda, and Bogdan Savchynskyy. “Exact MAP-Inference by Confining Combinatorial Search with LP Relaxation”. In: *Thirty-Second AAAI Conference on Artificial Intelligence*. 2018.

Barry Hurley et al. “Multi-language evaluation of exact solvers in graphical model discrete optimization”. In: *Constraints* (2016), pp. 1–22.

Philippe Jégou, Samba Ndojh Ndiaye, and Cyril Terrioux. “A new Evaluation of Forward Checking and its Consequences on Efficiency of Tools for Decomposition of CSPs”. In: *2008 20th IEEE International Conference on Tools with Artificial Intelligence*. Vol. 1. IEEE. 2008, pp. 486–490.

Peter Jeavons and Justyna Petke. “Local consistency and SAT-solvers”. In: *Journal of Artificial Intelligence Research* 43 (2012), pp. 329–351.

Philippe Jégou and Cyril Terrioux. “Hybrid backtracking bounded by tree-decomposition of constraint networks”. In: *Artificial Intelligence* 146.1 (2003), pp. 43–75.

Vladimir Kolmogorov, Andrei A. Krokhin, and Michal Rolínek. “The Complexity of General-Valued CSPs”. In: *SIAM J. Comput.* 46.3 (2017), pp. 1087–1110. DOI: `10.1137/16M1091836`. URL: `https://doi.org/10.1137/16M1091836`.

E.P. Klement, R. Mesiar, and E. Pap. *Triangular Norms*. Kluwer Academic Publishers, 2000.

Vladimir Kolmogorov. "Convergent tree-reweighted message passing for energy minimization". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 28.10 (2006), pp. 1568–1583.

Vladimir Kolmogorov. "A new look at reweighted message passing". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 37.5 (2015), pp. 919–930.

A M C A. Koster. "Frequency assignment: Models and Algorithms". Available at www.zib.de/koster/thesis.html. PhD thesis. The Netherlands: University of Maastricht, Nov. 1999.

Oliver Kullmann. "The Science of Brute Force". In: *Communications of the ACM* (2017).

Juerg Kohlas and Nic Wilson. "Semiring induced valuation algebras: Exact and approximate local computation algorithms". In: *Artificial Intelligence* 172.11 (2008), pp. 1360–1399.

Evelyn Lamb. "Maths proof smashes size record: supercomputer produces a 200-terabyte proof–but is it really mathematics?" In: *Nature* 534.7605 (2016), pp. 17–19.

S.L. Lauritzen and D.J. Spiegelhalter. “Local computations with probabilities on graphical structures and their application to expert systems”. In: *Journal of the Royal Statistical Society – Series B* 50 (1988), pp. 157–224.

A. K. Mackworth. “Consistency in networks of relations”. In: *Artificial Intelligence* 8 (1977), pp. 99–118.

R. Marinescu and R. Dechter. “AND/OR branch-and-bound for graphical models”. In: *Proc. of the* 19^{th} *IJCAI*. Edinburgh, Scotland, 2005, p. 224.

Radu Marinescu and Rina Dechter. “Memory intensive AND/OR search for combinatorial optimization in graphical models”. In: *Artificial Intelligence* 173.16-17 (2009), pp. 1492–1524.

R. Mohr and T.C. Henderson. “Arc and Path Consistency Revisited”. In: *Artificial Intelligence* 28.2 (1986), pp. 225–233.

Matthew W Moskewicz et al. “Chaff: Engineering an efficient SAT solver”. In: *Proceedings of the 38th annual Design Automation Conference*. ACM. 2001, pp. 530–535.

Vikram Khipple Mulligan et al. “Designing Peptides on a Quantum Computer”. In: *bioRxiv* (2019), p. 752485.

Jaroslav Nešetřil and Patrice Ossona De Mendez. “Tree-depth, subgraph coloring and homomorphism bounds”. In: *European Journal of Combinatorics* 27.6 (2006), pp. 1022–1041.

Hiroki Noguchi et al. “Computational design of symmetrical eight-bladed β-propeller proteins”. In: *IUCrJ* 6.1 (2019).

Youngsuk Park et al. “Learning the network structure of heterogeneous data via pairwise exponential Markov random fields”. In: *Proceedings of machine learning research* 54 (2017), p. 1302.

Judea Pearl. *Probabilistic Reasoning in Intelligent Systems, Networks of Plausible Inference.* Palo Alto: Morgan Kaufmann, 1988.

Rasmus Palm, Ulrich Paquet, and Ole Winther. “Recurrent relational networks”. In: *Advances in Neural Information Processing Systems.* 2018, pp. 3368–3378.

Niles A Pierce and Erik Winfree. “Protein design is NP-hard.”. In: *Protein Eng.* 15.10 (Oct. 2002), pp. 779–82. ISSN: 0269-2139. URL: `http://www.ncbi.nlm.nih.gov/pubmed/12468711`.

Daniel Prusa and Tomas Werner. “Universality of the local marginal polytope”. In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 37.4 (2015), pp. 898–904.

J. Alan Robinson. "A machine-oriented logic based on the resolution principle". In: *Journal of the ACM* 12 (1965), pp. 23–44.

D.J. Rose. "Tringulated Graphs and the elimination process". In: *Journal of Mathematical Analysis and its Applications* 32 (1970).

Daniela Röthlisberger et al. "Kemp elimination catalysts by computational enzyme design.". In: *Nature* 453.7192 (May 2008), pp. 190–5. ISSN: 1476-4687. DOI: 10.1038/nature06879. URL: http://www.ncbi.nlm.nih.gov/pubmed/18354394.

M.I. Schlesinger. "Sintaksicheskiy analiz dvumernykh zritelnikh signalov v usloviyakh pomekh (Syntactic analysis of two-dimensional visual signals in noisy conditions)". In: *Kibernetika* 4 (1976), pp. 113–130.

Thomas Schiex. *A note on CSP graph parameters.* Tech. rep. Citeseer, 1999.

T. Schiex, H. Fargier, and G. Verfaillie. "Valued Constraint Satisfaction Problems: hard and easy problems". In: *Proc. of the* 14^{th} *IJCAI*. Montréal, Canada, Aug. 1995, pp. 631–637.

G. Shafer. *An Axiomatic Study of Computation in Hypertrees.* Working paper 232. Lawrence: University of Kansas, School of Business, 1991.

David Simoncini et al. "Guaranteed Discrete Energy Optimization on Large Protein Design Problems". In: *Journal of Chemical Theory and Computation* 11.12 (2015), pp. 5980–5989. DOI: 10.1021/acs.jctc.5b00594.

C. Terrioux and P. Jegou. "Bounded backtracking for the valued constraint satisfaction problems". In: *Proc. of the Ninth International Conference on Principles and Practice of Constraint Programming (CP-2003)*. 2003.

Seydou Traoré et al. "A New Framework for Computational Protein Design through Cost Function Network Optimization". In: *Bioinformatics* 29.17 (2013), pp. 2129–2136.

Johan Thapper and Stanislav Zivny. "The Complexity of Finite-Valued CSPs". In: *J. ACM* 63.4 (2016), 37:1–37:33. DOI: 10.1145/2974019. URL: https://doi.org/10.1145/2974019.

Chris Umans, ed. *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*. IEEE Computer Society, 2017. ISBN: 978-1-5386-3464-6. URL: https://ieeexplore.ieee.org/xpl/conhome/8100284/proceeding.

T. Werner. "A Linear Programming Approach to Max-sum Problem: A Review.". In: *IEEE Trans. on Pattern Recognition and Machine Intelligence* 29.7 (July 2007), pp. 1165–1179. URL: http://dx.doi.org/10.1109/TPAMI.2007.1036.

Jonathan S Yedidia, William T Freeman, and Yair Weiss. "Bethe free energy, Kikuchi approximations, and belief propagation algorithms". In: *Advances in neural information processing systems* 13 (2001).

Dmitriy Zhuk. "A Proof of CSP Dichotomy Conjecture". In: *58th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2017, Berkeley, CA, USA, October 15-17, 2017*. Ed. by Chris Umans. IEEE Computer Society, 2017, pp. 331–342. ISBN: 978-1-5386-3464-6. DOI: 10.1109/FOCS.2017.38. URL: https://doi.org/10.1109/FOCS.2017.38.

Yuanlin Zhang and Roland HC Yap. "Making AC-3 an optimal algorithm". In: *IJCAI*. Vol. 1. 2001, pp. 316–321.