Optimization in Graphical Models - NP-complete optimization and its applications

Thomas Schiex

Optimization in Graphical Models

NP-complete optimization and its applications

T. Schiex

Many co-workers and contributors, see bibliography

AFIA General assembly
EurAI Fellowship
October 2016

The second era of AI

Spring of AI

- AI algorithms win chess championship, Deep Blue II, 1997.
- Computing power, RAM, search and a bit of ML (evaluation function)

The second era of AI

Spring of AI

- AI algorithms win chess championship, Deep Blue II, 1997.
- Computing power, RAM, search and a bit of ML (evaluation function)

Now Go !

- 2016, AlphaGo wins over a world champion Go player .
- Computing power, RAM, Monte Carlo Tree Search, deep learning.

The second era of AI

Spring of AI

- AI algorithms win chess championship, Deep Blue II, 1997.
- Computing power, RAM, search and a bit of ML (evaluation function)

Now Go !

- 2016, AlphaGo wins over a world champion Go player .
- Computing power, RAM, Monte Carlo Tree Search, deep learning.

New Spring of AI

- Mostly computing power, ML and data.
- Before that, NP-hard problem solving (CP, SAT, search) has been providing major contributions to AI, in its theory and applications.

The power of NP-hard problem solving

Inside AI and computer science

1. Model Based Diagnosis, Planning,...
2. Scheduling, Configuration, Resource Allocation,...
3. Verification, Cryptography, Testing

Inside AI and computer science

1. Model Based Diagnosis, Planning,...
2. Scheduling, Configuration, Resource Allocation,...
3. Verification, Cryptography, Testing

SAT, CSP, CP, ASP: satisfiability/feasibility

Finding perfect solutions to perfectly defined problems

The power of NP-hard problem solving

Inside AI and computer science

1. Model Based Diagnosis, Planning,...
2. Scheduling, Configuration, Resource Allocation,...
3. Verification, Cryptography, Testing

SAT, CSP, CP, ASP: satisfiability/feasibility

Finding perfect solutions to perfectly defined problems

The world is often more complex
ML gives us the capacity to build models of complex systems from data.

Perfect solution to perfect problems

Early planning for GPS navigation

- A digital model of roads
- A shortest path algorithm for planning
- GPS navigation monitors plan execution

Perfect solution to perfect problems

Early planning for GPS navigation

- A digital model of roads
- A shortest path algorithm for planning
- GPS navigation monitors plan execution

- Static model of reality
- Exact and efficient algorithm to decide from this “exact” model

Modern planning for GPS navigation

- Data on travel time on each road, at all time
- Feeds a stochastic model of duration
- Planning by minimizing expected time
- GPS navigation monitors plan execution

From data to decisions: **data analytics**

Modern planning for GPS navigation

- Data on travel time on each road, at all time
- Feeds a stochastic model of duration
- Planning by minimizing expected time
- GPS navigation monitors plan execution

Data analytics: what do we need ?

- Massive data
- A family of (stochastic) models that can be learnt
- On which optimization algorithms may apply
- Minimum requirement: efficient/tight bounds.

Definition

1. Set $X = \{x_1, \cdots, x_n\}$ of variables, each with a domain D^i
2. Set Φ of functions φ_S involving variables of $S \subset X$ (scope).
3. A joint function on X, $\mathbf{x} \in D^X$:

$$F(\mathbf{x}) = \bigoplus_{\varphi_S \in \Phi} \varphi_S(\mathbf{x}_S)$$

Graphical Model

Definition

1. Set $X = \{x_1, \cdots, x_n\}$ of variables, each with a domain D^i
2. Set Φ of functions φ_S involving variables of $S \subset X$ (scope).
3. A joint function on X, $\mathbf{x} \in D^X$:

$$F(\mathbf{x}) = \bigoplus_{\varphi_S \in \Phi} \varphi_S(\mathbf{x}_S)$$

Why Graphical ?

1. a vertex per variable
2. an edge when 2 variables interact in a $\varphi_S \in \Phi$
3. Allows to describe knowledge on a lot of variables concisely
4. Usually hard to manipulate (NP-hard queries).

A Constraint Network is a graphical model

Constraint Network

1. Variables and domains
2. φ_S are Boolean functions (constraints, $\mathrm{t} \equiv 0 < 1 \equiv \mathrm{f}$)
3. Minimize joint function ($\oplus = \wedge =$ max)

$$\mathsf{F}(\mathbf{x}) = \max_{\varphi_S \in \Phi} \varphi_S(\mathbf{x}_S)$$

A Constraint Network is a graphical model

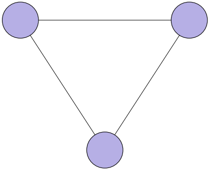
Constraint Network

1. Variables and domains
2. φ_S are Boolean functions (constraints, $t \equiv 0 < 1 \equiv f$)
3. Minimize joint function ($\oplus = \wedge = \max$)

$$F(\mathbf{x}) = \max_{\varphi_S \in \Phi} \varphi_S(\mathbf{x}_S)$$

Graph coloring/RLFAP-feas

1. A graph $G = (V, E)$ and m colors.
2. Can we color all vertices in such a way that no edge connects two vertices of the same color ?



A Constraint Network is a graphical model

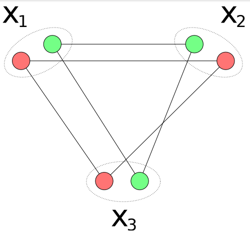
Constraint Network

1. Variables and domains
2. φ_S are Boolean functions (constraints, $t \equiv 0 < 1 \equiv f$)
3. Minimize joint function ($\oplus = \wedge = \max$)

$$F(\mathbf{x}) = \max_{\varphi_S \in \Phi} \varphi_S(\mathbf{x}_S)$$

Graph coloring/RLFAP-feas

1. A graph $G = (V, E)$ and m colors.
2. Can we color all vertices in such a way that no edge connects two vertices of the same color ?



Beyond feasibility... optimization

Biology has very few general laws

Need to shift from perfect knowledge processing to approximate/imperfect knowledge processing

Beyond feasibility... optimization

Biology has very few general laws

Need to shift from perfect knowledge processing to approximate/imperfect knowledge processing

Valued CSP

Shift from a conjunction of boolean functions to a more general combination of functions with a totally ordered co-domain (preferences, costs, priorities...)

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

S. Bistarelli, H. Fargier, U. Montanari, F. Rossi, et al. "Semiring-based CSPs and Valued CSPs: Frameworks, Properties and Comparison". In: Constraints 4 (1999), pp. 199–240

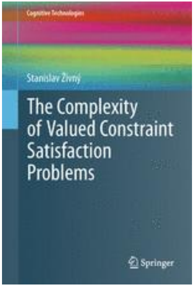
Beyond feasibility... optimization

Biology has very few general laws

Need to shift from perfect knowledge processing to approximate/imperfect knowledge processing

Valued CSP

Shift from a conjunction of boolean functions to a more general combination of functions with a totally ordered co-domain (preferences, costs, priorities...)



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

S. Bistarelli, H. Fargier, U. Montanari, F. Rossi, et al. "Semiring-based CSPs and Valued CSPs: Frameworks, Properties and Comparison". In: Constraints 4 (1999), pp. 199–240

Additive functions

Cost Function Networks - Weighted Constraint Networks - GAI models

- Variables and domains as usual
- Cost functions $\varphi_S : D^S \to \{0, \ldots, k\}$ (k finite or not)
- Combined by bounded addition[8]

$$F(\mathbf{x}) = \sum_{\varphi_S \in \Phi} \varphi_S(\mathbf{x}[S]) \qquad \varphi_\varnothing : \text{lower bound}$$

A solution has cost $< k$. Optimal if minimum cost.

Additive functions

Cost Function Networks - Weighted Constraint Networks - GAI models

- Variables and domains as usual
- Cost functions $\varphi_S : D^S \rightarrow \{0, \ldots, k\}$ (k finite or not)
- Combined by bounded addition[8]

$$F(\mathbf{x}) = \sum_{\varphi_S \in \Phi} \varphi_S(\mathbf{x}[S]) \qquad \varphi_\varnothing : \text{lower bound}$$

A solution has cost $< k$. Optimal if minimum cost.

Benefits

- Defines feasibility and cost homogeneously
- A constraint is a cost function with range $\{0, k\}$ only

Markov Random Fields, Bayesian Networks

- Random variables X with discrete domains
- joint probability distribution p(X) defined through the product of positive real-valued functions:

$$\mathsf{p}(\mathsf{X} = \mathbf{x}) \propto \prod_{\varphi_{\mathsf{S}} \in \Phi} \varphi_{\mathsf{S}}(\mathbf{x}_{\mathsf{S}})$$

Markov Random Fields, Bayesian Networks

- Random variables X with discrete domains
- joint probability distribution p(X) defined through the product of positive real-valued functions:

$$\mathsf{p}(\mathsf{X} = \mathbf{x}) \propto \prod_{\varphi_{\mathsf{S}} \in \Phi} \varphi_{\mathsf{S}}(\mathbf{x}_{\mathsf{S}})$$

Massively used in 2/3D Image Analysis, Statistical Physics, NLP, planning/reasoning under uncertainty...

Stochastic Graphical Models

Markov Random Fields, Bayesian Networks

- Random variables X with discrete domains
- joint probability distribution p(X) defined through the product of positive real-valued functions:

$$\mathrm{p}(\mathrm{X} = \mathbf{x}) \propto \prod_{\varphi_{\mathrm{S}} \in \Phi} \varphi_{\mathrm{S}}(\mathbf{x}_{\mathrm{S}})$$

Massively used in 2/3D Image Analysis, Statistical Physics, NLP, planning/reasoning under uncertainty...

Maximum a Posteriori MAP-MRF

MRF $\equiv$ CFN up to a $(-\log)$ transform.

Binary CFN/MRF as 01LP

1. No constraint or determinism: infinite k, finite costs
2. v_{ia}: value a used for variable x_i.
3. p_{iajb}: pair (a, b) used for x_i and x_j

$$\text{Minimize} \sum_{i,a} \varphi_i(a) \cdot v_{ia} + \sum_{\substack{\varphi_{ij} \in \Phi \\ a \in D^i, b \in D^j}} \varphi_{ij}(a,b) \cdot p_{iajb} \quad \text{subject to}$$

$$\sum_{a \in D^i} v_{ia} = 1 \qquad \forall i \in \{1, \ldots, n\}$$

$$\sum_{b \in D^j} p_{iajb} = v_{ia} \qquad \forall \varphi_{ij} \in \Phi, \forall a \in D^i$$

$$v_{ia}, p_{iajb} \in \{0, 1\}$$

Continuous relaxation : the local polytope[45,25,54]

Exact approaches (beyond ILP)

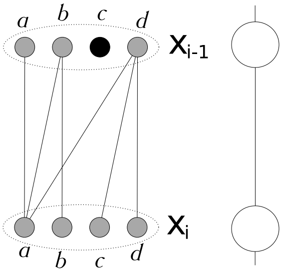
1. Dynamic programming (Variable and join-tree elimination,[2,12] resolution[11,40,13])
2. Tree search + fast, incremental approximate local reasoning (CP, SAT)

Fast, approximate reasoning with some guarantees

- Arc Consistency in CSP/CP
- Message Passing (MRF/BN)
- Soft Arc Consistency in CFN

AC as Dynamic programming

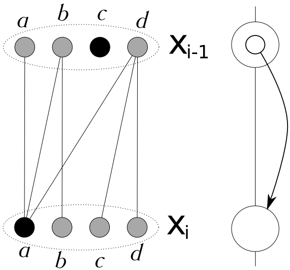
- Imagine a CSP with a linear graph
- Assume we know which values of D^{i-1} belongs to a solution on $x_1 \dots x_{i-1}$
- We can extend this to D^i for $x_1 \dots x_i$
- Let $f_i = \min_{x_{i-1}}(\max(\varphi_{i-1}, \varphi_{i,i-1}))$
- Include f_i in the problem



Arc Consistency = local dynamic programming

AC as Dynamic programming

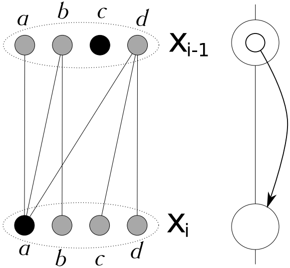
- Imagine a CSP with a linear graph
- Assume we know which values of D^{i-1} belongs to a solution on $x_1 \dots x_{i-1}$
- We can extend this to D^i for $x_1 \dots x_i$
- Let $f_i = \min_{x_{i-1}}(\max(\varphi_{i-1}, \varphi_{i,i-1}))$
- Include f_i in the problem



Arc Consistency = local dynamic programming

AC as Dynamic programming

- Imagine a CSP with a linear graph
- Assume we know which values of D^{i-1} belongs to a solution on $x_1 \ldots x_{i-1}$
- We can extend this to D^i for $x_1 \ldots x_i$
- Let $f_i = \min_{x_{i-1}}(\max(\varphi_{i-1}, \varphi_{i,i-1}))$
- Include f_i in the problem



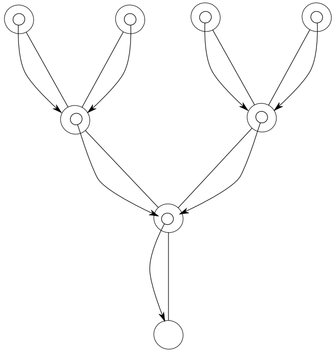
Revise: Equivalence Preserving Transformation

The resulting problem is equivalent (same set of solutions)

(Directional) AC solves Berge-acyclic CN

Rooted tree CN

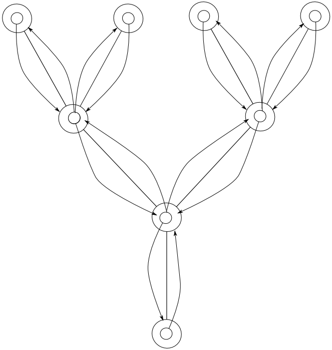
- Revise from leaves to root
- Root domain: values that belong to a solution



(Directional) AC solves Berge-acyclic CN

Tree CN

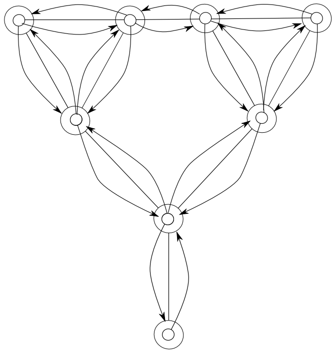
- Revise from leaves and back
- All domains: values that belong to a solution
- Resulting problem solved backtrack-free[19,18]



Can be done on any CN, with arbitrary graph

Arc consistency (Waltz 1972)

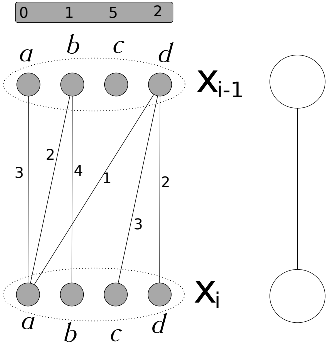
1. Linear time (tables)
2. Unique fixpoint (confluent)
3. Preserves equivalence
4. May detect infeasibility
5. Problem transformation (incremental)



Thomas Schiex, Jean-Charles. Régin, Chistine Gaspin, and Gérard Verfaillie. "Lazy Arc Consistency". In: Proc. of AAAI'96. Portland, OR: AAAI Press, Aug. 1996

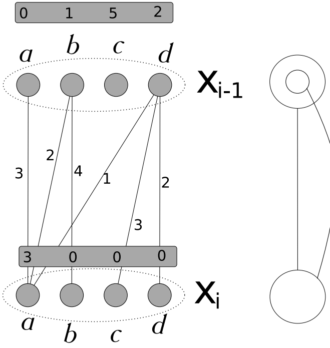
MP - Dynamic Programming[35,28]

1. Assume that the optimum cost on $x_1 \dots x_{i-1}$ is known for every value of D^{i-1}
2. We can extend this to D^i for $x_1 \dots x_i$
3. Let $f_i = \min_{x_{i-1}}(\varphi_{i-1} + \varphi_{i,i-1})$
4. Store f_i as a message



MP - Dynamic Programming[35,28]

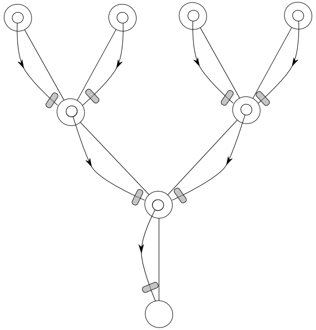
1. Assume that the optimum cost on $x_1 \dots x_{i-1}$ is known for every value of D^{i-1}
2. We can extend this to D^i for $x_1 \dots x_i$
3. Let $f_i = \min_{x_{i-1}}(\varphi_{i-1} + \varphi_{i,i-1})$
4. Store f_i as a message



MRF/BN/Factor graphs (− log domain)

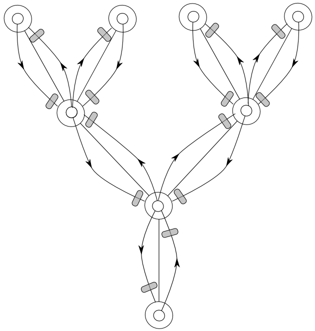
MP - Dynamic Programming[35,28]

1. Assume that the optimum cost on $x_1 \ldots x_{i-1}$ is known for every value of D^{i-1}
2. We can extend this to D^i for $x_1 \ldots x_i$
3. Let $f_i = \min_{x_{i-1}}(\varphi_{i-1} + \varphi_{i,i-1})$
4. Store f_i as a message



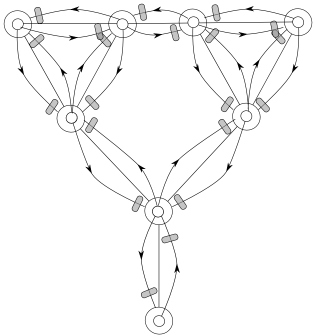
MP - Dynamic Programming[35,28]

1. Assume that the optimum cost on $x_1 \dots x_{i-1}$ is known for every value of D^{i-1}
2. We can extend this to D^i for $x_1 \dots x_i$
3. Let $f_i = \min_{x_{i-1}} (\varphi_{i-1} + \varphi_{i,i-1})$
4. Store f_i as a message



MP - Dynamic Programming[35,28]

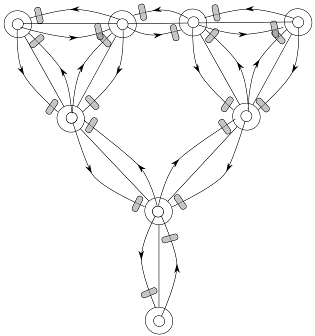
1. Assume that the optimum cost on $x_1 \ldots x_{i-1}$ is known for every value of D^{i-1}
2. We can extend this to D^i for $x_1 \ldots x_i$
3. Let $f_i = \min_{x_{i-1}}(\varphi_{i-1} + \varphi_{i,i-1})$
4. Store f_i as a message



MRF/BN/Factor graphs (− log domain)

MP - Dynamic Programming[35,28]

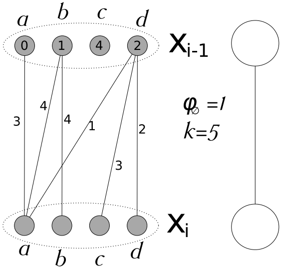
1. Assume that the optimum cost on $x_1 \dots x_{i-1}$ is known for every value of D^{i-1}
2. We can extend this to D^i for $x_1 \dots x_i$
3. Let $f_i = \min_{x_{i-1}}(\varphi_{i-1} + \varphi_{i,i-1})$
4. Store f_i as a message



1. Solves Berge acyclic MRF/BN (acyclic Factor Graphs)
2. Does not converge on graphs (Loopy Belief Propagation)
3. Massively used to produce “good” solutions (turbo-decoding[39])
4. Not an equivalence preserving transformation!

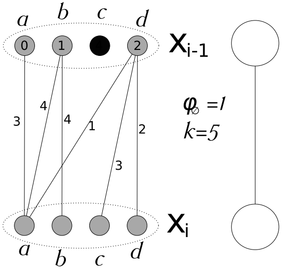
Soft AC as Dynamic programming

1. use $\varphi_\varnothing$ and $\varphi_i(\cdot)$ to store optimum cost over $(x_1, \ldots, x_i)$
2. Preserves equivalence by "cost shifting"[45,53,42]



Soft AC as Dynamic programming

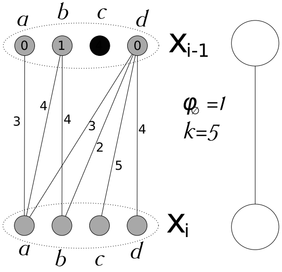
1. use $\varphi_\varnothing$ and $\varphi_i(\cdot)$ to store optimum cost over $(x_1, \ldots, x_i)$
2. Preserves equivalence by "cost shifting"[45,53,42]



Soft Arc Consistency ≈ MP through transformation

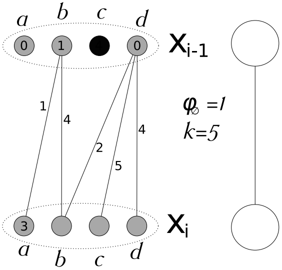
Soft AC as Dynamic programming

1. use $\varphi_\varnothing$ and $\varphi_i(\cdot)$ to store optimum cost over $(x_1, \ldots, x_i)$
2. Preserves equivalence by "cost shifting"[45,53,42]



Soft AC as Dynamic programming

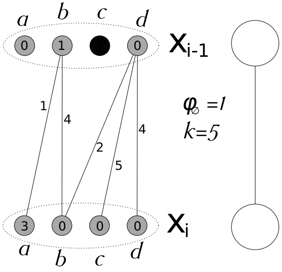
1. use $\varphi_\varnothing$ and $\varphi_i(\cdot)$ to store optimum cost over $(x_1, \ldots, x_i)$
2. Preserves equivalence by "cost shifting"[45,53,42]



Soft Arc Consistency ≈ MP through transformation

Soft AC as Dynamic programming

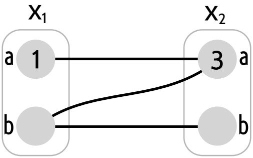
1. use $\varphi_\varnothing$ and $\varphi_i(\cdot)$ to store optimum cost over $(x_1, \ldots, x_i)$
2. Preserves equivalence by "cost shifting"[45,53,42]



T. Schiex. "Arc consistency for soft constraints". In: Principles and Practice of Constraint Programming - CP 2000. Vol. 1894. LNCS. Singapore, Sept. 2000, pp. 411–424

Tightening by cost shifting

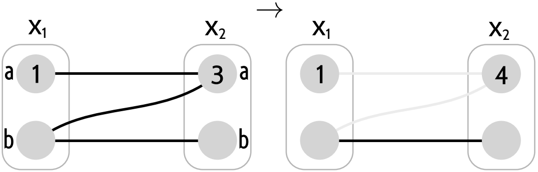
Assume that initially $\varphi_{\varnothing} = 0$, $k = 4$



Tightening by cost shifting

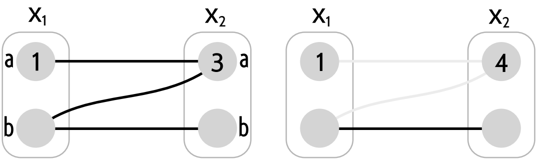
Assume that initially $\varphi_\varnothing = 0$, k $= 4$

Shift 1 to right a



Tightening by cost shifting

Assume that initially $\varphi_\varnothing = 0$, k $= 4$



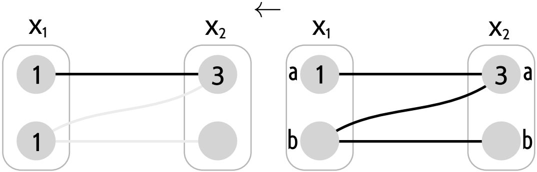
$\leftarrow$

Shift 1 from right a

Tightening by cost shifting

Assume that initially $\varphi_\varnothing = 0$, k $= 4$

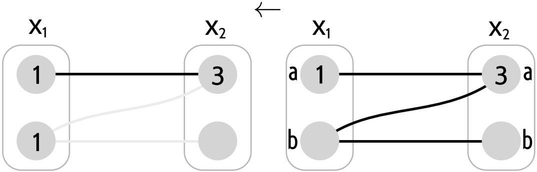
Shift 1 to left b



Tightening by cost shifting

Assume that initially $\varphi_\varnothing = 0, k = 4$

Shift 1 to left b

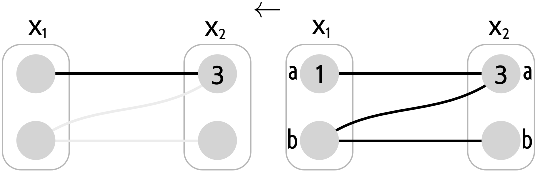


$\Downarrow$ Shift 1 from φ_1 to $\varphi_\varnothing$

Tightening by cost shifting

Assume that initially $\varphi_\varnothing = 0$, k $= 4$

Shift 1 to left b



$\Downarrow$ Shift 1 from φ_1 to $\varphi_\varnothing$

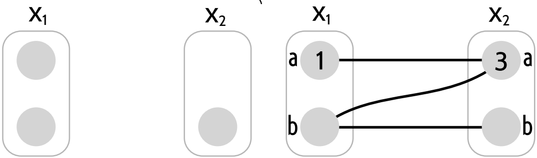
$\varphi_\varnothing = 1$

Tightening by cost shifting

Assume that initially $\varphi_\varnothing = 0$, k $= 4$

Shift 1 to left b

$\leftarrow$



$\Downarrow$ Shift 1 from φ_1 to $\varphi_\varnothing$

$\varphi_\varnothing = 1$

Preserves the joint function F$(\cdot)$ below k

Properties

- Solves tree structured problems, optimum in $\varphi_\varnothing$
- Reformulation: incremental
- May loop indefinitely (graphs)
- No unique fixpoint (when it exists)
- Already used by Ukrainian school[45,27,26,54] and for a subclass of ILP.[53]
- Independently introduced in 2003 in ML as "reparametrizations"[51]

T. Schiex. "Arc consistency for soft constraints". In: Principles and Practice of Constraint Programming - CP 2000. Vol. 1894. LNCS. Singapore, Sept. 2000, pp. 411–424

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

Convergent Soft Arc Consistencies

Breaking the loops, strengthening $\varphi_\varnothing$

1. 2000: Arc consistency[42,29]
2. 2003: Full Directional AC[6,31,32]
3. 2005: Existential DAC[30]
4. 2008: Virtual AC[10,9]
5. 2007: Optimal Soft AC[7] (LP)

Convergent Soft Arc Consistencies

Breaking the loops, strengthening $\varphi_\varnothing$

1. 2000: Arc consistency[42,29]
2. 2003: Full Directional AC[6,31,32]
3. 2005: Existential DAC[30]
4. 2008: Virtual AC[10,9]
5. 2007: Optimal Soft AC[7] (LP)

- $O(ed)$ space, equivalent to the CSP variants on CSP (except OSAC).
- Implemented in **toulbar2**, with many other bells & whistles

Javier Larrosa and Thomas Schiex. "Solving weighted CSP by maintaining arc consistency". In: Artif. Intell. 159.1-2 (2004), pp. 1–26

M. Cooper, S. de Givry, M. Sanchez, T. Schiex, et al. "Soft arc consistency revisited". In: Artificial Intelligence 174 (2010), pp. 449–478

LP variables, binary CFN with no constraint/determinism

1. u_i: amount of cost shifted from φ_i to $\varphi_\varnothing$
2. q_{ija}: amount of cost shifted from φ_{ij} to $a \in D^i$

Optimal Soft Arc Consistency (finite costs, $k = \infty$)

LP variables, binary CFN with no constraint/determinism

1. u_i: amount of cost shifted from φ_i to $\varphi_\varnothing$
2. q_{ija}: amount of cost shifted from φ_{ij} to $a \in D^i$

OSAC[45,25,7,54,9]

Maximize $\sum_{i=1}^{n} u_i$ subject to

$$\varphi_i(a) - u_i + \sum_{(\varphi_{ij} \in \Phi)} q_{ija} \geq 0 \qquad \forall i \in \{1, \dots, n\}, \forall a \in D^i$$

$$\varphi_{ij}(a, b) - q_{ija} - q_{jib} \geq 0 \qquad \forall \varphi_{ij} \in \Phi, \forall (a, b) \in D^{ij}$$

M C. Cooper, S. de Givry, and T. Schiex. "Optimal soft arc consistency". In: Proc. of IJCAI'2007. Hyderabad, India, Jan. 2007, pp. 68–73

Remember the local polytope ?

01 LP Variables, for a binary CFN

1. v_{ia}: value a used for variable x_i.
2. p_{iajb}: pair (a, b) used for x_i and x_j

The MRF local polytope[54]

$$\text{Minimize} \sum_{i,a} c_i(a) \cdot x_{ia} + \sum_{\substack{c_{ij} \in C \\ a \in D^i, b \in D^j}} c_{ij}(a,b) \cdot y_{iajb} \quad \text{s.t}$$

$$\sum_{a \in D^i} x_{ia} = 1 \qquad \forall i \in \{1, \ldots, n\} \qquad (1)$$

$$\sum_{b \in D^j} y_{iajb} - x_{ia} = 0 \qquad \forall c_{ij}/c_{ji} \in C, \forall a \in D^i \qquad (2)$$

Remember the local polytope ?

01 LP Variables, for a binary CFN

1. v_{ia}: value a used for variable x_i.
2. p_{iajb}: pair (a, b) used for x_i and x_j

The MRF local polytope[54]

$$\text{Minimize} \sum_{i,a} c_i(a) \cdot x_{ia} + \sum_{\substack{c_{ij} \in C \\ a \in D^i, b \in D^j}} c_{ij}(a,b) \cdot y_{iajb} \quad \text{s.t}$$

$$\sum_{a \in D^i} x_{ia} = 1 \qquad \forall i \in \{1, \ldots, n\} \qquad (1)$$

$$\sum_{b \in D^j} y_{iajb} - x_{ia} = 0 \qquad \forall c_{ij}/c_{ji} \in C, \forall a \in D^i \qquad (2)$$

OSAC is its dual: u_i multiplier for (1) and q_{ija} for (2).

Better understanding

1. Soft convergent ACs find feasible (but non necessarily optimal) solutions of the dual.
2. Optimal does not mean more efficient for tree search.

1. Soft convergent ACs find feasible (but non necessarily optimal) solutions of the dual.
2. Optimal does not mean more efficient for tree search.

2015: universality of the local polytope

Prusa and Werner[38] showed that any "normal" LP can be reduced to such a polytope in linear time (constructive proof).

1. Soft convergent ACs find feasible (but non necessarily optimal) solutions of the dual.
2. Optimal does not mean more efficient for tree search.

2015: universality of the local polytope

Prusa and Werner[38] showed that any "normal" LP can be reduced to such a polytope in linear time (constructive proof).

Could soft arc consistency (eg. VAC) speed-up LP?

VAC faster than LP, close/same bound

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.
```

VAC faster than LP, close/same bound

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

```
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.
```

Progress

Radio Link Frequency Assignment Problems

Assign frequencies to radio link in order to minimize interferences.

CELAR 06, n $= 100$, d $= 44$ - one core

1. 1997: 26 days of a Sun UltraSparc 167 MHz.
2. 2015: optimum found in 7", proved in 73" (2.1GHz CPU)

- 12.5 fold increase in frequency
- More than 30,000 times faster
- All min-interference instances closed. Most resisted ILP (fap.zib.de)

B. Cabon, S. de Givry, L. Lobjois, T. Schiex, et al. "Radio Link Frequency Assignment". In: Constraints Journal 4 (1999), pp. 79–89

Computational Protein Design[48,1]

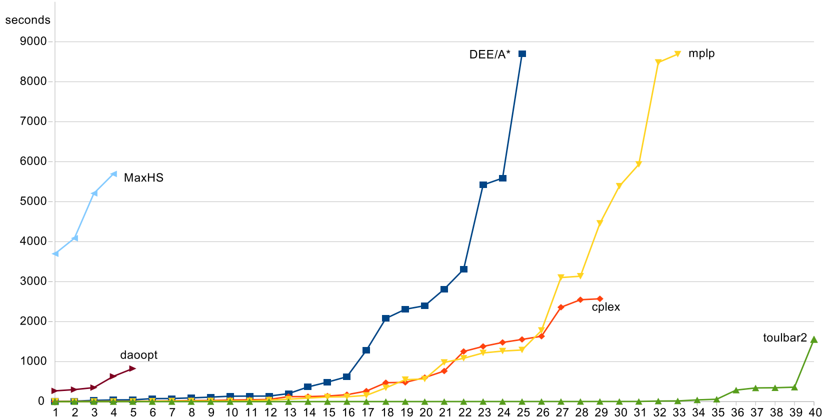
Design new enzymes for biofuels, drugs...cosmetics too

- graphical model capturing molecule stability based on atom-scale forces (electrostatics, solvation, torsion, Van der Waals...)

- Few variables (from 10 to few hundreds)
- Huge domains (typ. d $= 450$)
- Exact solvers: DEE/A* (DEE[14]), ILP[24]
- By far the most used: simulated annealing (Rosetta[21]).

Seydou Traoré, David Allouche, Isabelle André, Simon de Givry, et al. "A new framework for computational protein design through cost function network optimization". In: Bioinformatics 29.17 (2013), pp. 2129–2136

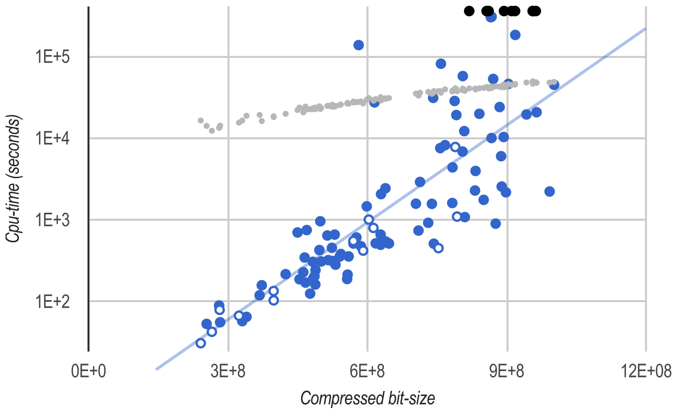
Multi-paradigm comparison - QP,SDP,ILP,Maxsat,MRF,CFN



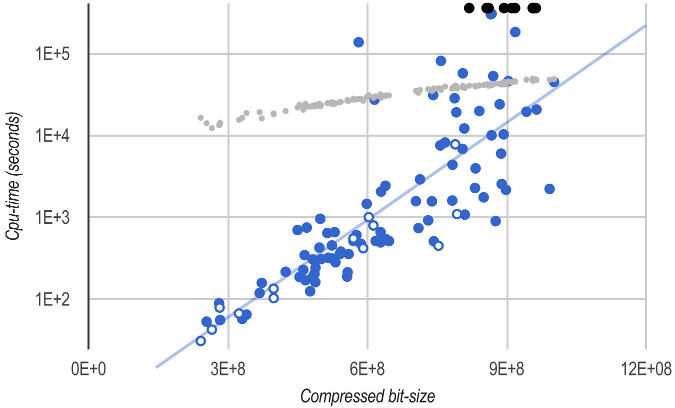
David Allouche, Isabelle André, Sophie Barbe, Jessica Davies, et al. "Computational protein design as an optimization problem". In: Artificial Intelligence 212 (2014), pp. 59–79

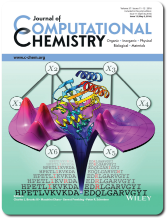
Barry Hurley, Barry O'Sullivan, David Allouche, George Katsirelos, et al. "Multi-language evaluation of exact solvers in graphical model discrete optimization". In: Constraints 21.3 (2016), pp. 413–434

Faster than dedicated simulated annealing[46]



Faster than dedicated simulated annealing[46]





David Simoncini, David Allouche, Simon de Givry, Celine Delmas, et al. "Guaranteed discrete energy optimization on large protein design problems". In: Journal of chemical theory and computation 11.12 (2015), pp. 5980–5989

Seydou Traoré, Kyle E Roberts, David Allouche, Bruce R Donald, et al. "Fast search algorithms for computational protein design". In: Journal of computational chemistry (2016)

Graphical models for biology

Genetic mapping with CarthaGene

GM learning + TSP optimization.

S. de Givry, M. Bouchez, P. Chabrier, D. Milan, et al. "CarthaGene: multipopulation integrated genetic and radiation hybrid mapping." In: Bioinformatics 21.8 (2005), pp. 1703–4

V. Laurent, E. Wajnberg, B. Mangin, T. Schiex, et al. "A composite genetic map of the parasitoid wasp Trichogramma brassicae based on RAPD markers." In: Genetics 150.1 (1998), pp. 275–82

RNA Gene finding: MilPat, Darn!

CSP/CFN + string algorithms

P. Thebault, S. de Givry, T. Schiex, and C. Gaspin. "Searching RNA motifs and their intermolecular contacts with constraint networks." In: Bioinformatics 22.17 (2006), pp. 2074–80

M Zytnicki, C Gaspin, and T Schiex. "DARN! A soft constraint solver for RNA motif localization". In: Constraints 13.1 (2008), pp. 91–109

Gene finding with EuGene

Semi-CRF + optimization.

Tomato Genome Consortium et al. "The tomato genome sequence provides insights into fleshy fruit evolution". In: Nature 485.7400 (2012), pp. 635–641

C. Mathé, M. F. Sagot, T. Schiex, and P. Rouzé. "Current methods of gene prediction, their strengths and weaknesses." In: Nucleic Acids Res 30.19 (2002), pp. 4103–17

MendelSoft: pedigree debugging

CFN for MPE on Bayesian nets

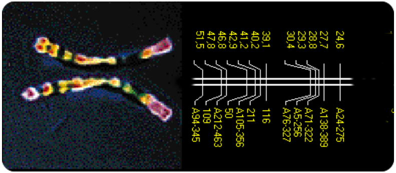
M Sanchez, S de Givry, and T Schiex. "Mendelian error detection in complex pedigrees using weighted constraint satisfaction techniques". In: Constraints 13.1 (2008), pp. 130–154

T. Faraut, S. de Givry, P. Chabrier, T. Derrien, et al. "A comparative genome approach to marker ordering." In: Bioinformatics 23.2 (2007), e50–6

Graphical models for biology

Genetic mapping with CarthaGene

GM learning + TSP optimization.



RNA Gene finding: MilPat, Darn!

CSP/CFN + string algorithms

P. Thebault, S. de Givry, T. Schiex, and C. Gaspin. "Searching RNA motifs and their intermolecular contacts with constraint networks." In: Bioinformatics 22.17 (2006), pp. 2074–80

M Zytnicki, C Gaspin, and T Schiex. "DARN! A soft constraint solver for RNA motif localization". In: Constraints 13.1 (2008), pp. 91–109

Gene finding with EuGene

Semi-CRF + optimization.

Tomato Genome Consortium et al. "The tomato genome sequence provides insights into fleshy fruit evolution". In: Nature 485.7400 (2012), pp. 635–641

C. Mathé, M. F. Sagot, T. Schiex, and P. Rouzé. "Current methods of gene prediction, their strengths and weaknesses." In: Nucleic Acids Res 30.19 (2002), pp. 4103–17

MendelSoft: pedigree debugging

CFN for MPE on Bayesian nets

M Sanchez, S de Givry, and T Schiex. "Mendelian error detection in complex pedigrees using weighted constraint satisfaction techniques". In: Constraints 13.1 (2008), pp. 130–154

T. Faraut, S. de Givry, P. Chabrier, T. Derrien, et al. "A comparative genome approach to marker ordering." In: Bioinformatics 23.2 (2007), e50–6

Graphical models for biology

Genetic mapping with CarthaGene

GM learning + TSP optimization.

S. de Givry, M. Bouchez, P. Chabrier, D. Milan, et al. "CarthaGene: multipopulation integrated genetic and radiation hybrid mapping." In: Bioinformatics 21.8 (2005), pp. 1703–4

V. Laurent, E. Wajnberg, B. Mangin, T. Schiex, et al. "A composite genetic map of the parasitoid wasp Trichogramma brassicae based on RAPD markers." In: Genetics 150.1 (1998), pp. 275–82

RNA Gene finding: MilPat, Darn!

CSP/CFN + string algorithms

P. Thebault, S. de Givry, T. Schiex, and C. Gaspin. "Searching RNA motifs and their intermolecular contacts with constraint networks." In: Bioinformatics 22.17 (2006), pp. 2074–80

M Zytnicki, C Gaspin, and T Schiex. "DARN! A soft constraint solver for RNA motif localization". In: Constraints 13.1 (2008), pp. 91–109

Gene finding with EuGene

Semi-CRF + optimization.

MendelSoft: pedigree debugging

CFN for MPE on Bayesian nets

M Sanchez, S de Givry, and T Schiex. "Mendelian error detection in complex pedigrees using weighted constraint satisfaction techniques". In: Constraints 13.1 (2008), pp. 130–154

T. Faraut, S. de Givry, P. Chabrier, T. Derrien, et al. "A comparative genome approach to marker ordering." In: Bioinformatics 23.2 (2007), e50–6

Graphical models for biology

Genetic mapping with CarthaGene

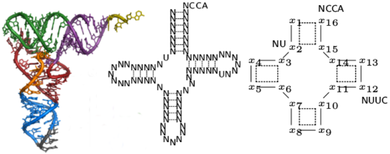
GM learning + TSP optimization.

S. de Givry, M. Bouchez, P. Chabrier, D. Milan, et al. “CarthaGene: multipopulation integrated genetic and radiation hybrid mapping.” In: Bioinformatics 21.8 (2005), pp. 1703–4

V. Laurent, E. Wajnberg, B. Mangin, T. Schiex, et al. “A composite genetic map of the parasitoid wasp Trichogramma brassicae based on RAPD markers.” In: Genetics 150.1 (1998), pp. 275–82

RNA Gene finding: MilPat, Darn!

CSP/CFN + string algorithms



Gene finding with EuGene

Semi-CRF + optimization.

Tomato Genome Consortium et al. “The tomato genome sequence provides insights into fleshy fruit evolution”. In: Nature 485.7400 (2012), pp. 635–641

C. Mathé, M. F. Sagot, T. Schiex, and P. Rouzé. “Current methods of gene prediction, their strengths and weaknesses.” In: Nucleic Acids Res 30.19 (2002), pp. 4103–17

MendelSoft: pedigree debugging

CFN for MPE on Bayesian nets

M Sanchez, S de Givry, and T Schiex. “Mendelian error detection in complex pedigrees using weighted constraint satisfaction techniques”. In: Constraints 13.1 (2008), pp. 130–154

T. Faraut, S. de Givry, P. Chabrier, T. Derrien, et al. “A comparative genome approach to marker ordering.” In: Bioinformatics 23.2 (2007), e50–6

Graphical models for biology

Genetic mapping with CarthaGene

GM learning + TSP optimization.

S. de Givry, M. Bouchez, P. Chabrier, D. Milan, et al. "CarthaGene: multipopulation integrated genetic and radiation hybrid mapping." In: Bioinformatics 21.8 (2005), pp. 1703–4

V. Laurent, E. Wajnberg, B. Mangin, T. Schiex, et al. "A composite genetic map of the parasitoid wasp Trichogramma brassicae based on RAPD markers." In: Genetics 150.1 (1998), pp. 275–82

RNA Gene finding: MilPat, Darn!

CSP/CFN + string algorithms

P. Thebault, S. de Givry, T. Schiex, and C. Gaspin. "Searching RNA motifs and their intermolecular contacts with constraint networks." In: Bioinformatics 22.17 (2006), pp. 2074–80

M Zytnicki, C Gaspin, and T Schiex. "DARN! A soft constraint solver for RNA motif localization". In: Constraints 13.1 (2008), pp. 91–109

Gene finding with EuGene

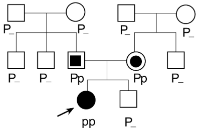
Semi-CRF + optimization.

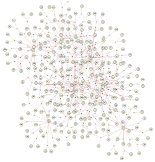
Tomato Genome Consortium et al. "The tomato genome sequence provides insights into fleshy fruit evolution". In: Nature 485.7400 (2012), pp. 635–641

C. Mathé, M. F. Sagot, T. Schiex, and P. Rouzé. "Current methods of gene prediction, their strengths and weaknesses." In: Nucleic Acids Res 30.19 (2002), pp. 4103–17

MendelSoft: pedigree debugging

CFN for MPE on Bayesian nets





Beyond optimization

Graphical models with decision and stochastic variables

- MRF/BN: stochastic variables + probability distribution. Can be learnt (RBM for deep learning).
- CFN: decision variables + cost functions/utilities/feasibility distribution.

Beyond optimization

Graphical models with decision and stochastic variables

- MRF/BN: stochastic variables + probability distribution. Can be learnt (RBM for deep learning).
- CFN: decision variables + cost functions/utilities/feasibility distribution.

Mixed graphical models with decision and stochastic variables

- Influence diagrams[22]
- Stochastic constraint programming[52]
- Plausability-feasibility-utility networks[37]

H. Fargier, J. Lang, R. Martin-Clouaire, and Thomas Schiex. "A constraint satisfaction framework for decision under uncertainty". In: Proc. of the 11th Int. Conf. on Uncertainty in Artificial Intelligence. Montréal, Canada, Aug. 1995

H. Fargier, J. Lang, and T. Schiex. "Mixed Constraint Satisfaction: a framework for decision problems under incomplete knowledge". In: Proc. of AAAI'96. Portland, OR: AAAI Press, Aug. 1996

C Pralet, G Verfaillie, and T Schiex. "An algebraic graphical model for decision with uncertainties, feasibilities, and utilities". In: Journal of Artificial Intelligence Research 29 (2007), pp. 421–489

Beyond NP

- Marginal probability: weighted counting[50]
- Marginal MAP: maximize probability with unobservable variables.
- Maximize expected utility...

Clément Viricel, David Simoncini, Sophie Barbe, and Thomas Schiex. "Guaranteed Weighted Counting for Affinity Computation: Beyond Determinism and Structure". In: International Conference on Principles and Practice of Constraint Programming. Springer. 2016, pp. 733–750

C Pralet, T Schiex, and G Verfaillie. "From influence diagrams to multioperator cluster DAGs". In: Proc. of UAI'2006. Boston, MA, 2006

Beyond optimization

Graphical models

Provide a strong and grounded basis for data analytics, from data to decision.

Beyond optimization

Graphical models

Provide a strong and grounded basis for data analytics, from data to decision.

We need efficient anytime anyspace algorithms, with associated modelling langages and solvers.

Questions ?

References I

[1] David Allouche, Isabelle André, Sophie Barbe, Jessica Davies, et al. "Computational protein design as an optimization problem". In: Artificial Intelligence 212 (2014), pp. 59–79.

[2] Umberto Bertelé and Francesco Brioshi. Nonserial Dynamic Programming. Academic Press, 1972.

[3] S. Bistarelli, H. Fargier, U. Montanari, F. Rossi, et al. "Semiring-based CSPs and Valued CSPs: Frameworks, Properties and Comparison". In: Constraints 4 (1999), pp. 199–240.

[4] B. Cabon, S. de Givry, L. Lobjois, T. Schiex, et al. "Radio Link Frequency Assignment". In: Constraints Journal 4 (1999), pp. 79–89.

[5] Tomato Genome Consortium et al. "The tomato genome sequence provides insights into fleshy fruit evolution". In: Nature 485.7400 (2012), pp. 635–641.

[6] M C. Cooper. "Reduction operations in fuzzy or valued constraint satisfaction". In: Fuzzy Sets and Systems 134.3 (2003), pp. 311–342.

[7] M C. Cooper, S. de Givry, and T. Schiex. "Optimal soft arc consistency". In: Proc. of IJCAI'2007. Hyderabad, India, Jan. 2007, pp. 68–73.

[8] M C. Cooper and T. Schiex. "Arc consistency for soft constraints". In: Artificial Intelligence 154.1-2 (2004), pp. 199–227.

[9] M. Cooper, S. de Givry, M. Sanchez, T. Schiex, et al. "Soft arc consistency revisited". In: Artificial Intelligence 174 (2010), pp. 449–478.

[10] Martin C Cooper, Simon de Givry, Martı Sánchez, Thomas Schiex, et al. "Virtual Arc Consistency for Weighted CSP." In: AAAI. Vol. 8. 2008, pp. 253–258.

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

[12] Rina Dechter. "Bucket Elimination: A Unifying Framework for Reasoning". In: Artificial Intelligence 113.1–2 (1999), pp. 41–85.

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

[14] J Desmet, M De Maeyer, B Hazes, and I Lasters. "The dead-end elimination theorem and its use in protein side-chain positioning." In: Nature 356.6369 (Apr. 1992), pp. 539–42. ISSN: 0028-0836.

[15] T. Faraut, S. de Givry, P. Chabrier, T. Derrien, et al. "A comparative genome approach to marker ordering." In: Bioinformatics 23.2 (2007), e50–6.

[16] H. Fargier, J. Lang, R. Martin-Clouaire, and Thomas Schiex. "A constraint satisfaction framework for decision under uncertainty". In: Proc. of the 11th Int. Conf. on Uncertainty in Artificial Intelligence. Montréal, Canada, Aug. 1995.

[17] H. Fargier, J. Lang, and T. Schiex. "Mixed Constraint Satisfaction: a framework for decision problems under incomplete knowledge". In: Proc. of AAAI'96. Portland, OR: AAAI Press, Aug. 1996.

References III

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

[19] Eugene C. Freuder. "A sufficient Condition for Backtrack-free Search". In: Journal of the ACM 29.1 (1982), pp. 24–32.

[20] S. de Givry, M. Bouchez, P. Chabrier, D. Milan, et al. "CarthaGene: multipopulation integrated genetic and radiation hybrid mapping." In: Bioinformatics 21.8 (2005), pp. 1703–4.

[21] Dominik Gront, Daniel W Kulp, Robert M Vernon, Charlie EM Strauss, et al. "Generalized fragment picking in Rosetta: design, protocols and applications". In: PloS one 6.8 (2011), e23294.

[22] Ronald A Howard and James E Matheson. "Influence diagrams". In: Decision Analysis 2.3 (2005), pp. 127–143.

[23] Barry Hurley, Barry O'Sullivan, David Allouche, George Katsirelos, et al. "Multi-language evaluation of exact solvers in graphical model discrete optimization". In: Constraints 21.3 (2016), pp. 413–434.

[24] Carleton L Kingsford, Bernard Chazelle, and Mona Singh. "Solving and analyzing side-chain positioning problems using linear and integer programming." In: Bioinformatics (Oxford, England) 21.7 (Apr. 2005), pp. 1028–36. ISSN: 1367-4803.

[25] 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.

[26] VK Koval' and Mykhailo Ivanovich Schlesinger. "Two-dimensional programming in image analysis problems". In: Avtomatika i Telemekhanika 8 (1976), pp. 149–168.

[27] VA Kovalevsky and VK Koval. "A diffusion algorithm for decreasing energy of max-sum labeling problem". In: Glushkov Institute of Cybernetics, Kiev, USSR (1975).

[28] Frank R Kschischang, Brendan J Frey, and Hans-Andrea Loeliger. "Factor graphs and the sum-product algorithm". In: Information Theory, IEEE Transactions on 47.2 (2001), pp. 498–519.

[29] J. Larrosa. "On Arc and Node Consistency in weighted CSP". In: Proc. AAAI'02. Edmondton, (CA), 2002, pp. 48–53.

[30] J. Larrosa, S. de Givry, F. Heras, and M. Zytnicki. "Existential arc consistency: getting closer to full arc consistency in weighted CSPs". In: Proc. of the 19th IJCAI. Edinburgh, Scotland, Aug. 2005, pp. 84–89.

[31] J. Larrosa and T. Schiex. "In the quest of the best form of local consistency for Weighted CSP". In: Proc. of the 18th IJCAI. Acapulco, Mexico, Aug. 2003, pp. 239–244.

[32] Javier Larrosa and Thomas Schiex. "Solving weighted CSP by maintaining arc consistency". In: Artif. Intell. 159.1-2 (2004), pp. 1–26.

[33] V. Laurent, E. Wajnberg, B. Mangin, T. Schiex, et al. "A composite genetic map of the parasitoid wasp Trichogramma brassicae based on RAPD markers." In: Genetics 150.1 (1998), pp. 275–82.

[34] C. Mathé, M. F. Sagot, T. Schiex, and P. Rouzé. "Current methods of gene prediction, their strengths and weaknesses." In: Nucleic Acids Res 30.19 (2002), pp. 4103–17.

References V

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

[36] C Pralet, T Schiex, and G Verfaillie. "From influence diagrams to multioperator cluster DAGs". In: Proc. of UAI'2006. Boston, MA, 2006.

[37] C Pralet, G Verfaillie, and T Schiex. "An algebraic graphical model for decision with uncertainties, feasibilities, and utilities". In: Journal of Artificial Intelligence Research 29 (2007), pp. 421–489.

[38] 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.

[39] Thomas J Richardson and Rüdiger L Urbanke. "The capacity of low-density parity-check codes under message-passing decoding". In: Information Theory, IEEE Transactions on 47.2 (2001), pp. 599–618.

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

[41] M Sanchez, S de Givry, and T Schiex. "Mendelian error detection in complex pedigrees using weighted constraint satisfaction techniques". In: Constraints 13.1 (2008), pp. 130–154.

[42] T. Schiex. "Arc consistency for soft constraints". In: Principles and Practice of Constraint Programming - CP 2000. Vol. 1894. LNCS. Singapore, Sept. 2000, pp. 411–424.

[43] 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.

[44] Thomas Schiex, Jean-Charles. Régin, Chistine Gaspin, and Gérard Verfaillie. “Lazy Arc Consistency”. In: Proc. of AAAI’96. Portland, OR: AAAI Press, Aug. 1996.

[45] 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.

[46] David Simoncini, David Allouche, Simon de Givry, Celine Delmas, et al. “Guaranteed discrete energy optimization on large protein design problems”. In: Journal of chemical theory and computation 11.12 (2015), pp. 5980–5989.

[47] P. Thebault, S. de Givry, T. Schiex, and C. Gaspin. “Searching RNA motifs and their intermolecular contacts with constraint networks.” In: Bioinformatics 22.17 (2006), pp. 2074–80.

[48] Seydou Traoré, David Allouche, Isabelle André, Simon de Givry, et al. “A new framework for computational protein design through cost function network optimization”. In: Bioinformatics 29.17 (2013), pp. 2129–2136.

[49] Seydou Traoré, Kyle E Roberts, David Allouche, Bruce R Donald, et al. “Fast search algorithms for computational protein design”. In: Journal of computational chemistry (2016).

[50] Clément Viricel, David Simoncini, Sophie Barbe, and Thomas Schiex. “Guaranteed Weighted Counting for Affinity Computation: Beyond Determinism and Structure”. In: International Conference on Principles and Practice of Constraint Programming. Springer. 2016, pp. 733–750.

[51] Martin J Wainwright, Tommi S Jaakkola, and Alan S Willsky. "Tree-based reparameterization framework for analysis of sum-product and related algorithms". In: IEEE Transactions on information theory 49.5 (2003), pp. 1120–1146.

[52] Toby Walsh. "Stochastic constraint programming". In: ECAI. Vol. 2. 2002, pp. 111–115.

[53] Dag Wedelin. "An algorithm for large scale 0–1 integer programming with application to airline crew scheduling". In: Annals of operations research 57.1 (1995), pp. 283–301.

[54] 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.

[55] M Zytnicki, C Gaspin, and T Schiex. "DARN! A soft constraint solver for RNA motif localization". In: Constraints 13.1 (2008), pp. 91–109.