PFIA'19

Toulbar2 solver demonstration

Simon de Givry, Thomas Schiex, David Allouche, George Katsirelos, Matthias Zytnicki, MIAT – INRA, Toulouse, France

Abdelkader Ouali, Samir Loudni, GREYC, University of Caen, France

Installation

https://github.com/toulbar2/toulbar2

- Open-source C++ exact solver **toulbar2 v1.0.1**

git clone --branch 1.0.1 **https://github.com/toulbar2/toulbar2.git**

cd toulbar2

cmake –DMPI=ON **.**

make –j 5

- Debian installation (v1.0.0)

echo "deb http://ftp.fr.debian.org/debian sid main" | sudo tee -a /etc/apt/sources.list

sudo apt-get update

sudo apt-get install toulbar2

Weighted n-queen problem

	Q0	Q1	Q2	Q3
Row0	4	4	2	1
Row1	4	3	1	2
Row2	3	4	3	3
Row3	4	4	2	4

Weighted n-queen problem

	Q0	Q1	Q2	Q3
Row0	4	**4**	2	1
Row1	4	3	1	**2**
Row2	**3**	4	3	3
Row3	4	4	**2**	4

Weighted n-queen problem

https://forgemia.inra.fr/thomas.schiex/cost-function-library/tree/master/random/wqueens

Instances	cplex		localsolver		toulbar2 VNS	
8x8	18	0,00 %	18	0,00 %	18	0,00 %
10x10	25	0,00 %	25	0,00 %	25	0,00 %
20x20	63	0,00 %	63	0,00 %	63	0,00 %
30x30	101	0,00 %	144	29,86 %	109	7,92 %
40x40	154	0,00 %	284	45,77 %	205	33,12 %
50x50	176	0,00 %	340	48,24 %	208	18,18 %

60 seconds on 30-core computer

Uncapacitated Warehouse Location Problem

(Kratica et al., RAIRO OR 2001)

Search nodes

	toulbar2 1.0.0	cplex 12.7.1
Capmo1 100x100	**155**	7,581
Capmo2 100x100	**25**	2,024
Capmo3 100x100	**93**	5,439
Capmo4 100x100	**23**	4,055
Capmo5 100x100	**28**	2,664

CPU time (sec. on PC i7 3GHz)

	toulbar2 1.0.0	cplex 12.7.1
Capmo1 100x100	**13.01**	20.13
Capmo2 100x100	3.06	**3.02**
Capmo3 100x100	13.32	**11.40**
Capmo4 100x100	**3.26**	7.45
Capmo5 100x100	**2.68**	4.62

Latin Square N x N with costs

Example of solution for N = 5:

2	1	3	5	4
4	2	1	3	5
1	5	4	2	3
5	3	2	4	1
3	4	5	1	2

All variables take a different value in each row and each column

A unary cost function for each cell $f^{i,j}(x_{i,j}) : D \rightarrow [0,MaxCost[$

Objective: 49

$$\text{Objective} = \sum_i \sum_j f^{i,j}(x_{i,j})$$

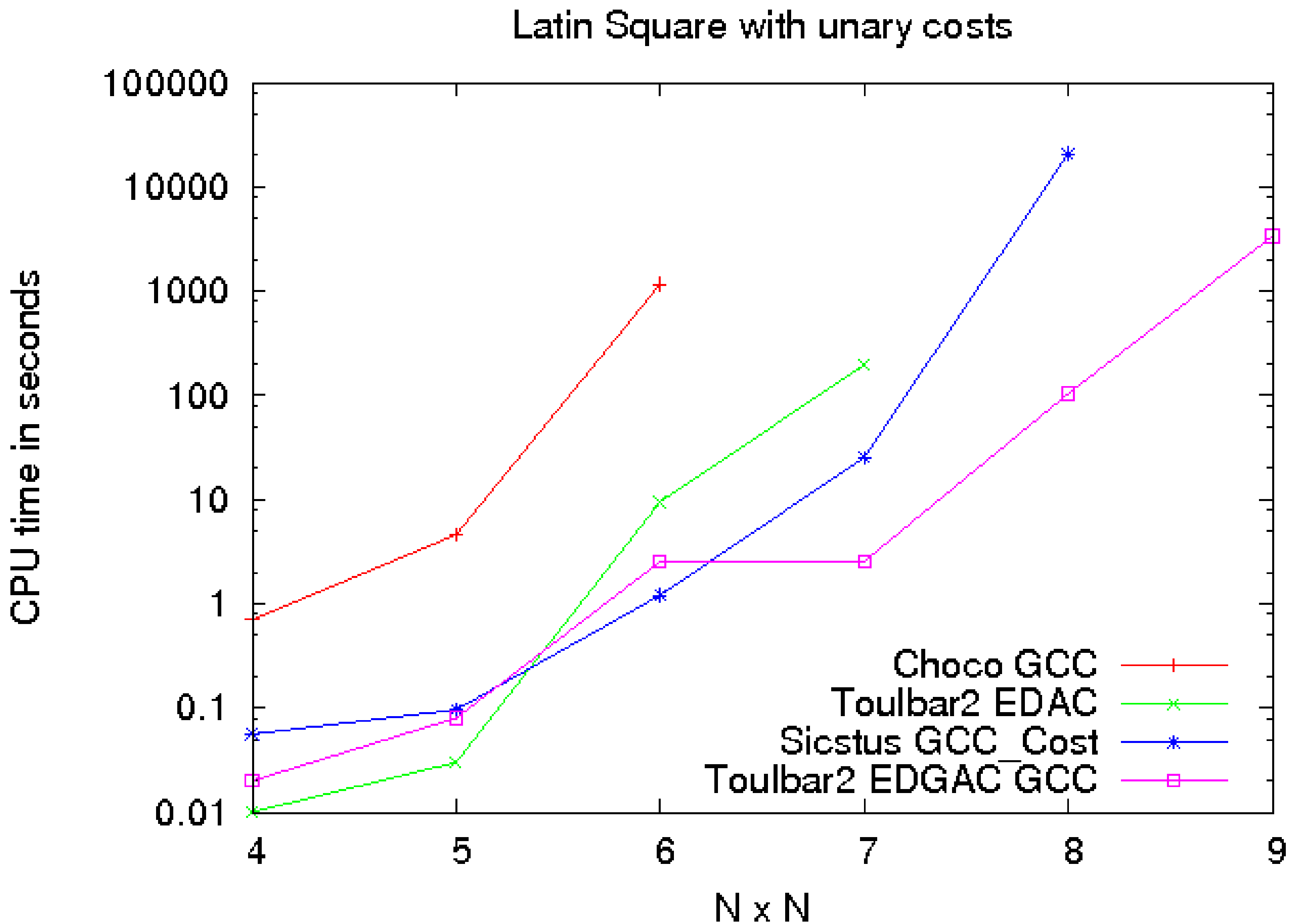
Weighted latin square problem

4, 4, 3, 4	4, 3, 4, 4	2, 1, 3, 2	1, 2, 3, 4
3, 1, 3, 3	4, 1, 1, 1	4, 1, 1, 3	4, 4, 1, 4
1, 3, 3, 2	2, 1, 3, 1	3, 4, 2, 2	2, 3, 1, 3
3, 4, 4, 2	3, 2, 4, 4	4, 1, 3, 4	4, 4, 4, 3

Weighted latin square problem

4, 4, 3, **4** 3	4, 3, **4**, 4 2	**2**, 1, 3, 2 0	1, **2**, 3, 4 1
3, **1**, 3, 3 1	4, 1, 1, **1** 3	4, 1, **1**, 3 2	**4**, 4, 1, 4 0
1, 3, 3, 2 0	2, **1**, 3, 1 1	3, 4, 2, **2** 3	2, 3, **1**, 3 2
3, 4, **4**, 2 2	**3**, 2, 4, 4 0	4, **1**, 3, 4 1	4, 4, 4, **3** 3

Latin Square with costs



GCC_Cost (Régin, *Constraints* 2002)   EDGAC (Lee & Leung, *AAAI* 2010)

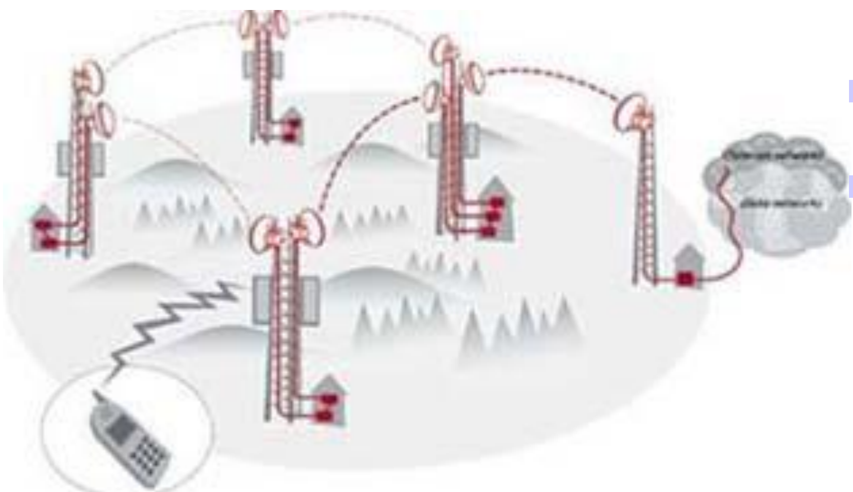
Radio Link Frequency Assignment (CELAR)

(Cabon *et al*, Constraints 1999 ; Givry *et al*, CP97 – AAAI06 – IJCAI07 – IJCAI09 – CP10)

$n \leq 458$, $d=44$, $e(2) \leq 5{,}000$

Radio Link Frequency Assignment Problem

(Cabon et al., *Constraints* 1999) (Koster *et al.*, *4OR* 2003)



- Given a telecommunication network
- ...find the best frequency for each communication link, avoiding interferences

- Best can be:
  - Minimize the maximum frequency, no interference (max operator)
  - **Minimize the global interference (sum operator)**
- Generalizes graph coloring problems: $|f_i - f_j| \geq a$

***CELAR** problem size: n=100—458 ; d=44 ; e=1,000—5,000*

CELAR 6 results since 1993

n. of vars: n=100, domain size: d=44, n. of cost functions: e=1222

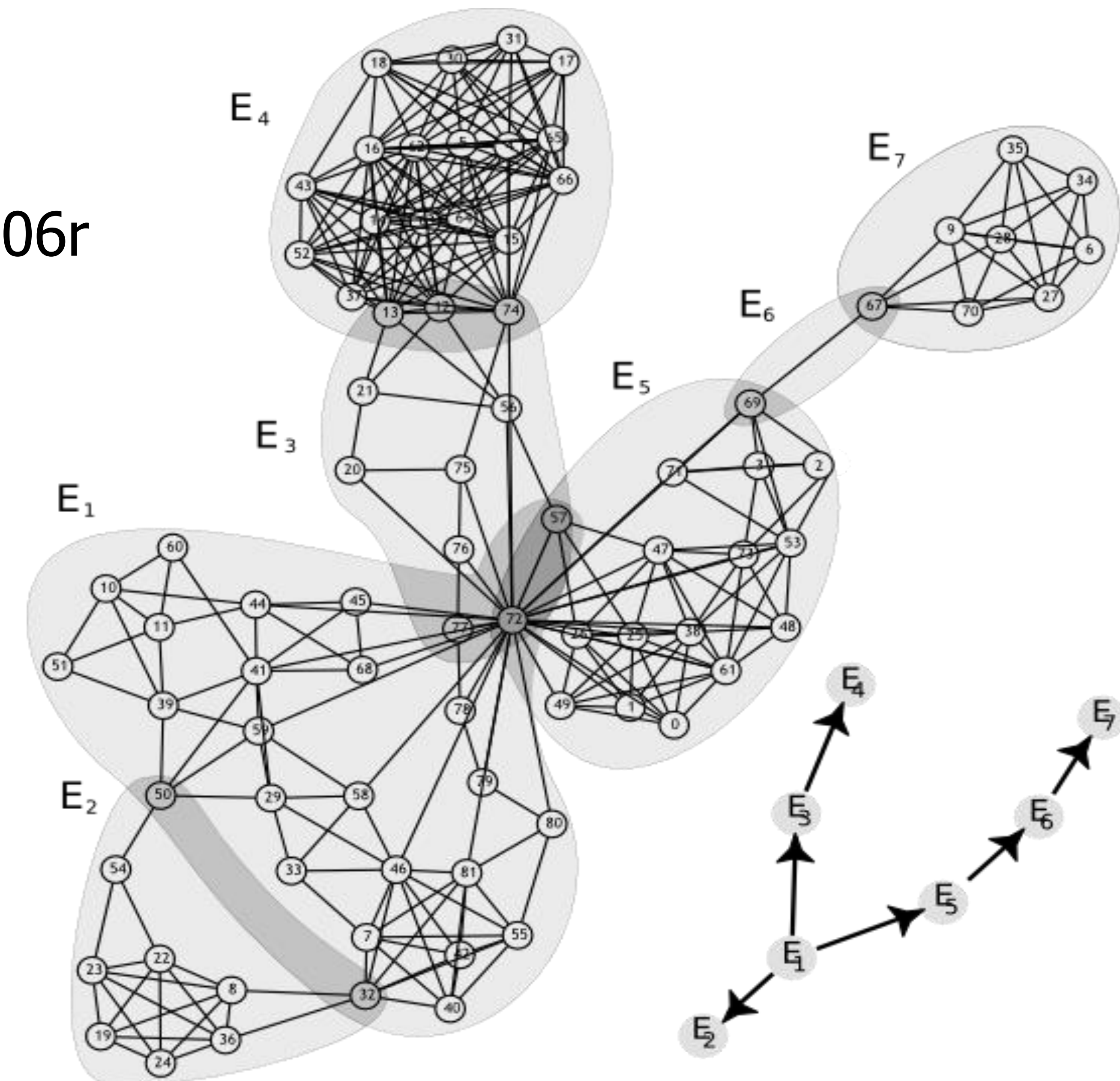
Time of optimality proof	Method(s) used	Publication
26 days (SUN UltraSparc 167 MHz)	Ad-hoc problem decomposition & Russian Doll Search *(22 vars only)*	(de Givry, Verfaillie, Schiex, CP 1997)
3 days (SUN Sparc 2)	Ad-hoc problem decomposition & PFC-MRDAC *(22 vars only)*	(Larrosa, Meseguer, Schiex, AIJ 1998)
8 hours (DEC Alpha 500MP)	Preprocessing rules & Cluster Tree Elimination	(Koster PhD thesis, 1999)
3 hours (PC 2.4 GHz)	B&B with EDAC & tree decomposition (BTD)	(de Givry, Schiex, Verfaillie, AAAI 2006)
1' 26'' **25000x** (PC 2.5 GHz) **16 x**	BTD & variable ordering heuristisc &dicho branching	(Sanchez, Allouche, de Givry,Schiex, IJCAI 2009)

CELAR 7 (n=200) solved in 4.5 days (Sanchez et al, IJCAI 2009)

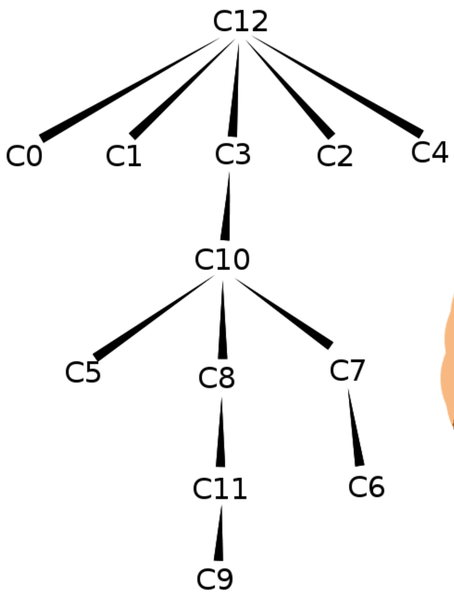
CELAR 8 (n=458) solved in 127 days (Allouche et al, CP 2010)

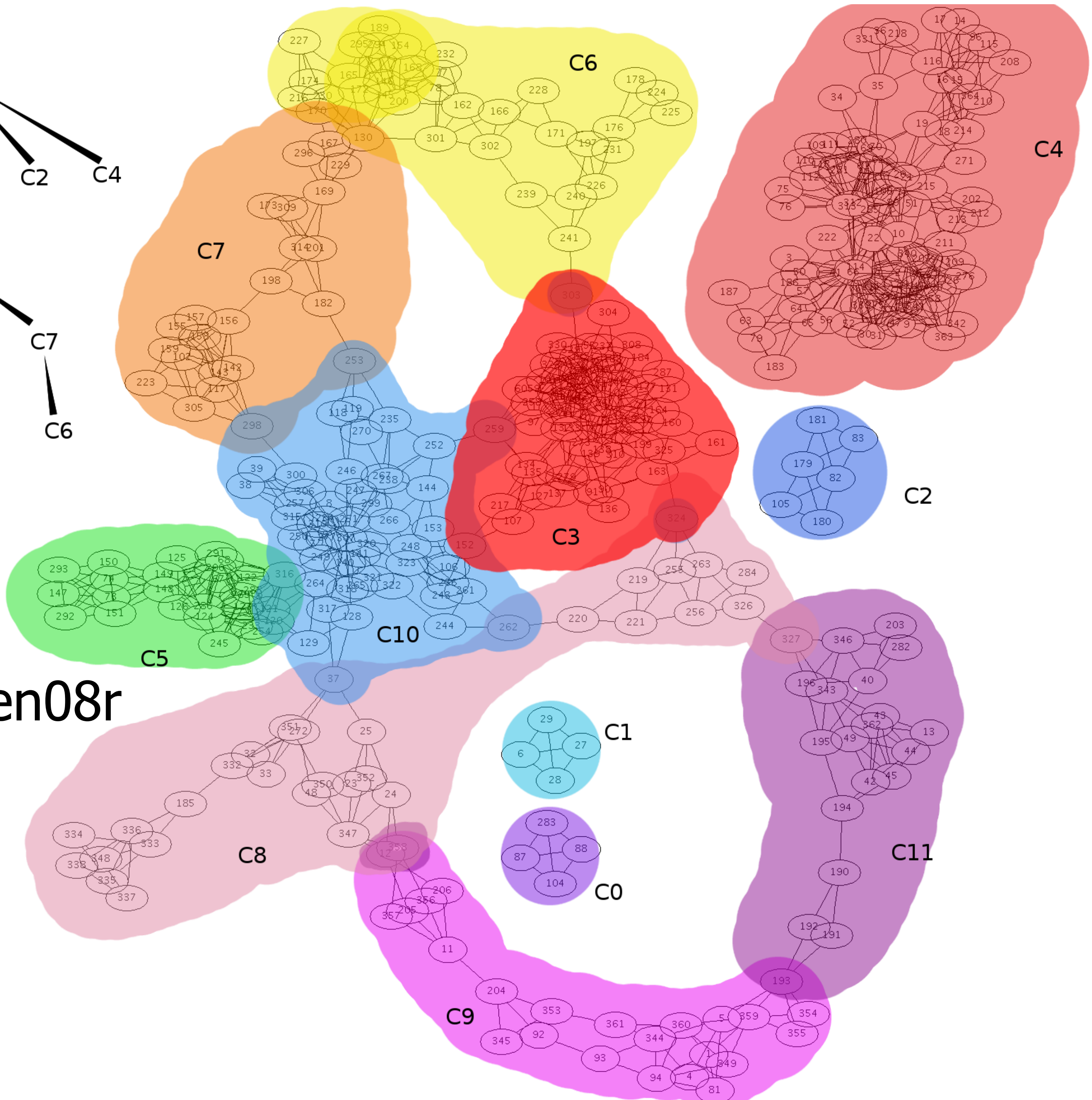
Tree decomposition example

CELAR scen06r
n = 82
d = 44
e = 327
w = 26
s = 3



(Allouche et al, CP 2010)





CELAR scen08r

n = 365

d = 44

e = 1539

w = 85

s = 4

Mendelian error correction in complex pedigree (MendelSoft)

(Sanchez *et al*, Constraints 2008)

n≤20,000, d≤66, e(3)≤30,000

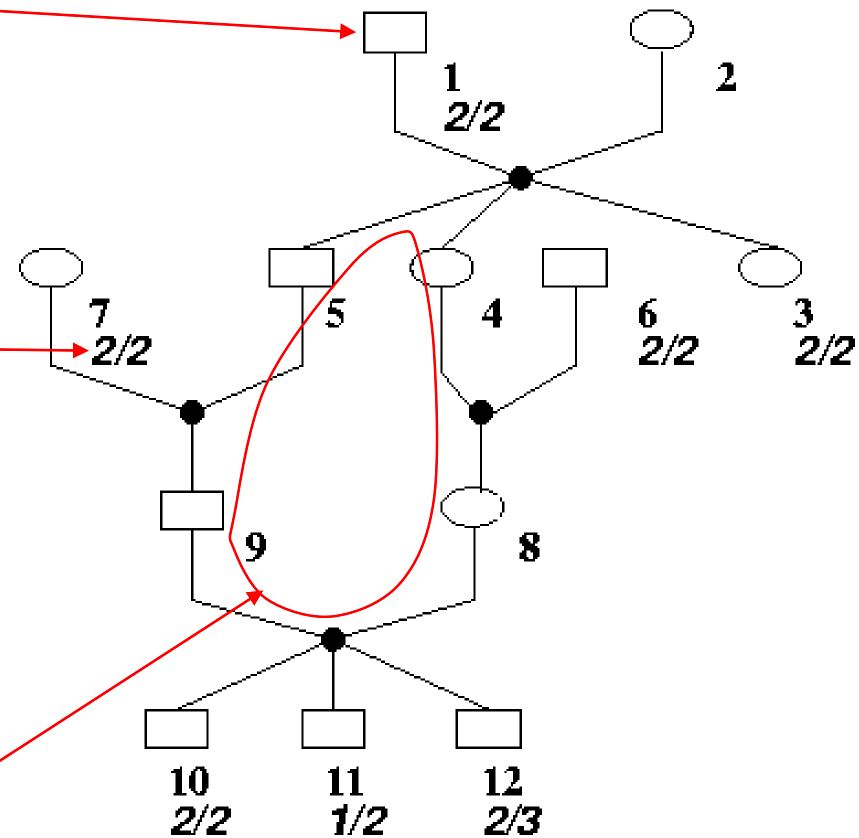
Cleaning the data after genotyping

Today, about 1% errors remain after SNP genotyping

→ For each marker, detects Mendelian inheritance errors

Individual (founder)

Genotype: unordered pair of alleles (possibly unobserved)

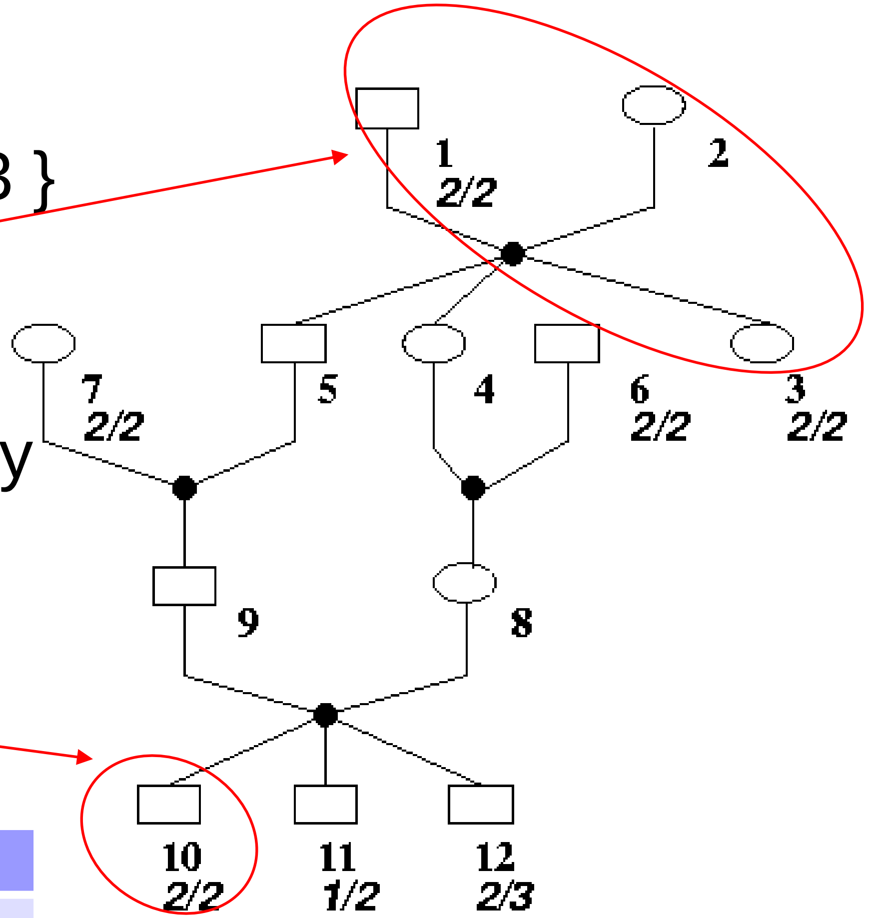


Marriage loop

Cost Function Network

- **X**: one variable per individual
- **D**: domain of every variable is defined as the set of all possible genotypes

  Here: { 1/1, 1/2, 1/3, 2/2, 2/3, 3/3 }
- **F**:
  - Ternary hard constraints to encode Mendelian laws for any non founder
  - Unary soft constraints to encode genotyping data

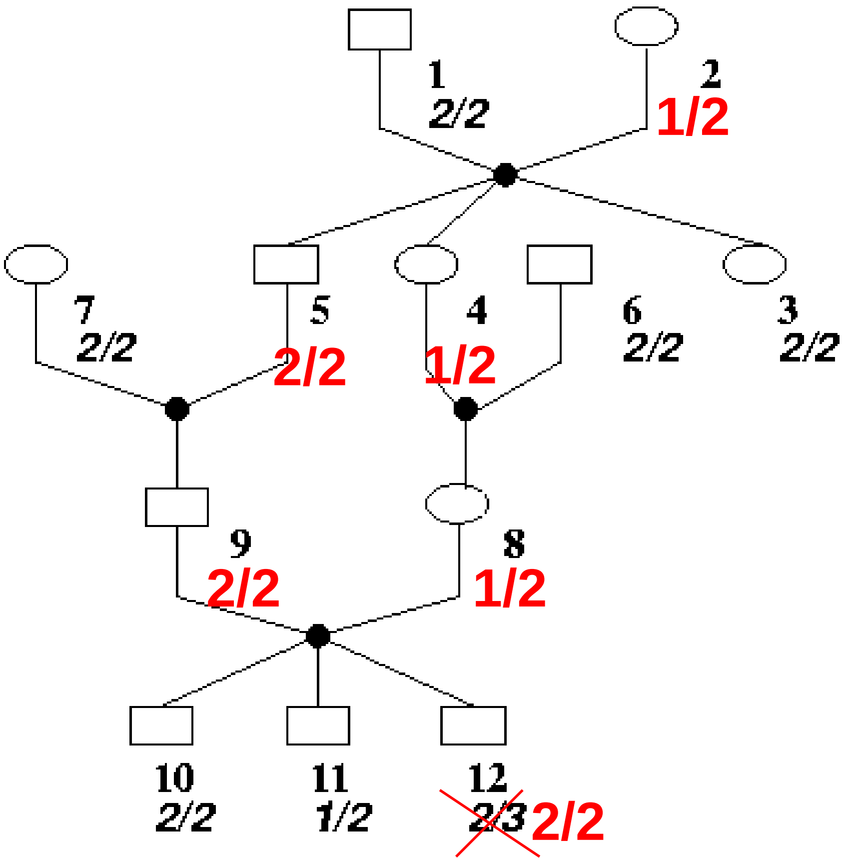


	1/1	1/2	1/3	2/2	2/3	3/3
f_{10}	1	1	1	0	1	1

Task 2: Error Detection

- Finds a complete assignment with the minimum number of errors

→Follows the parsimony principle



Real data

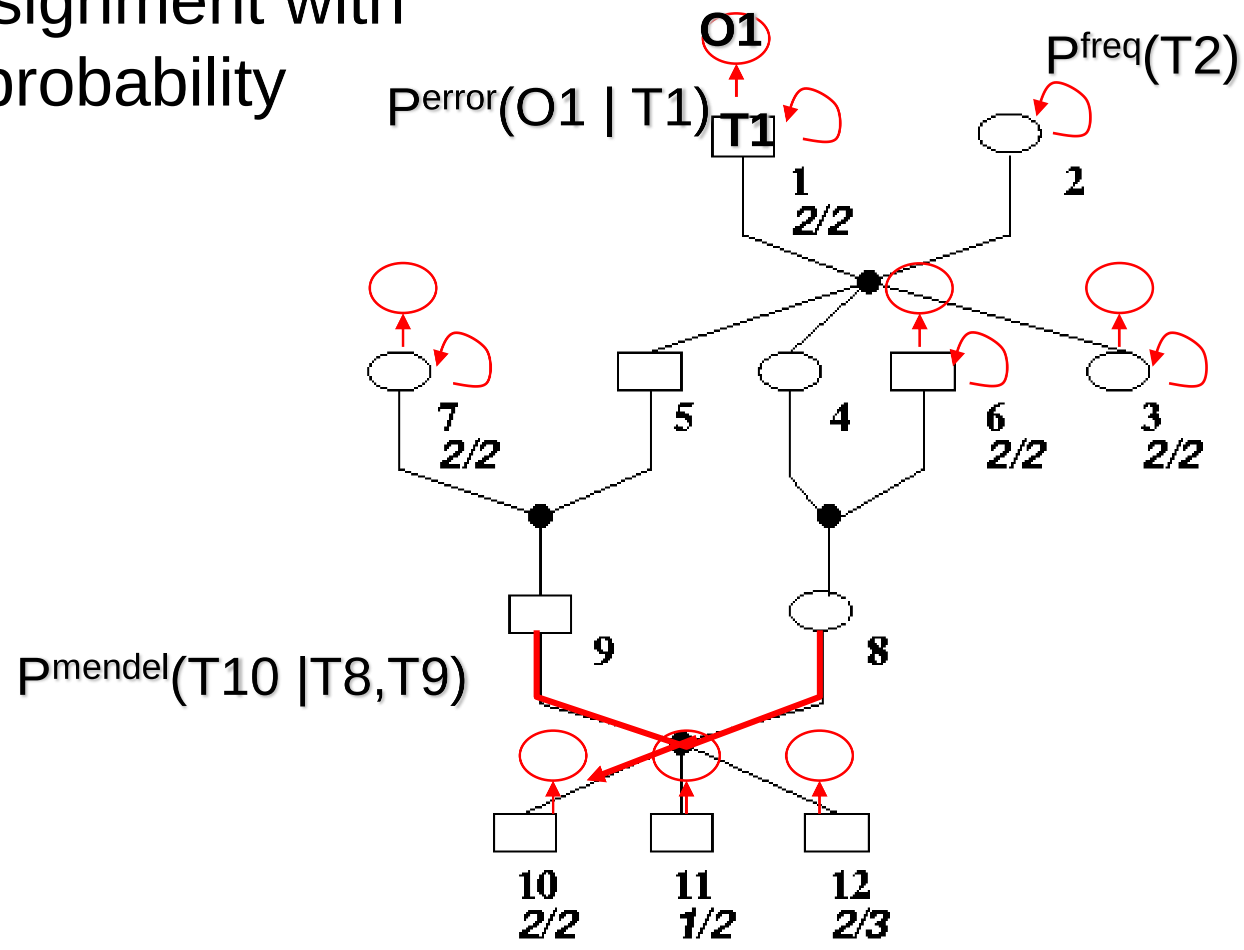
CPU time in seconds to find and prove optimality
on a linux PC 3 GHz with 16 GB using toulbar2 v0.5

								B&B-VE(2)		
	ind	*vars*	*genotyped*	alleles	*nf*	*ngen*	treewidth ub	errors	time	nodes
eye	36	36	28	6	11	4	2	1	0.02	0
cancer	49	48	37	8	18	5	2	1	0.21	0
parkinson	37	34	13	4	7	7	5	0	0	6
$berrichon_{1nc}$	129516	9947	2448	4	8821	17	262	2	4.73	8805
$berrichon_{1}$	129516	10017	2483	4	8786	17	330	23	5.81	8384
$berrichon_{2nc}$	27255	19337	10215	4	4719	19	-	41	5.89	6170
$berrichon_{2}$	27255	19562	10215	4	2381	19	-	106	17.23	15445
$langlade_{1}$	1355	1209	711	9	298	13	84	38	12.28	391
$langlade_{2}$	1355	1223	715	7	298	13	82	89	60.56	17857
$langlade_{3}$	1355	1258	787	5	298	13	85	39	14.19	6731
$langlade_{4}$	1355	1186	672	8	298	13	83	43	59.7	3520
$moissac_{1}$	283	260	183	2	81	5	6	0	0	5
$moissac_{2}$	283	244	167	7	81	5	6	0	0.51	6
$moissac_{3}$	283	225	151	3	81	5	6	0	0	4
$moissac_{4}$	283	256	179	2	81	5	6	0	0	5
$moissac_{5}$	283	237	161	8	81	5	6	0	1.02	5
$moissac_{6}$	283	201	131	11	81	5	5	0	5.64	6

Error Correction using Probabilistic Model

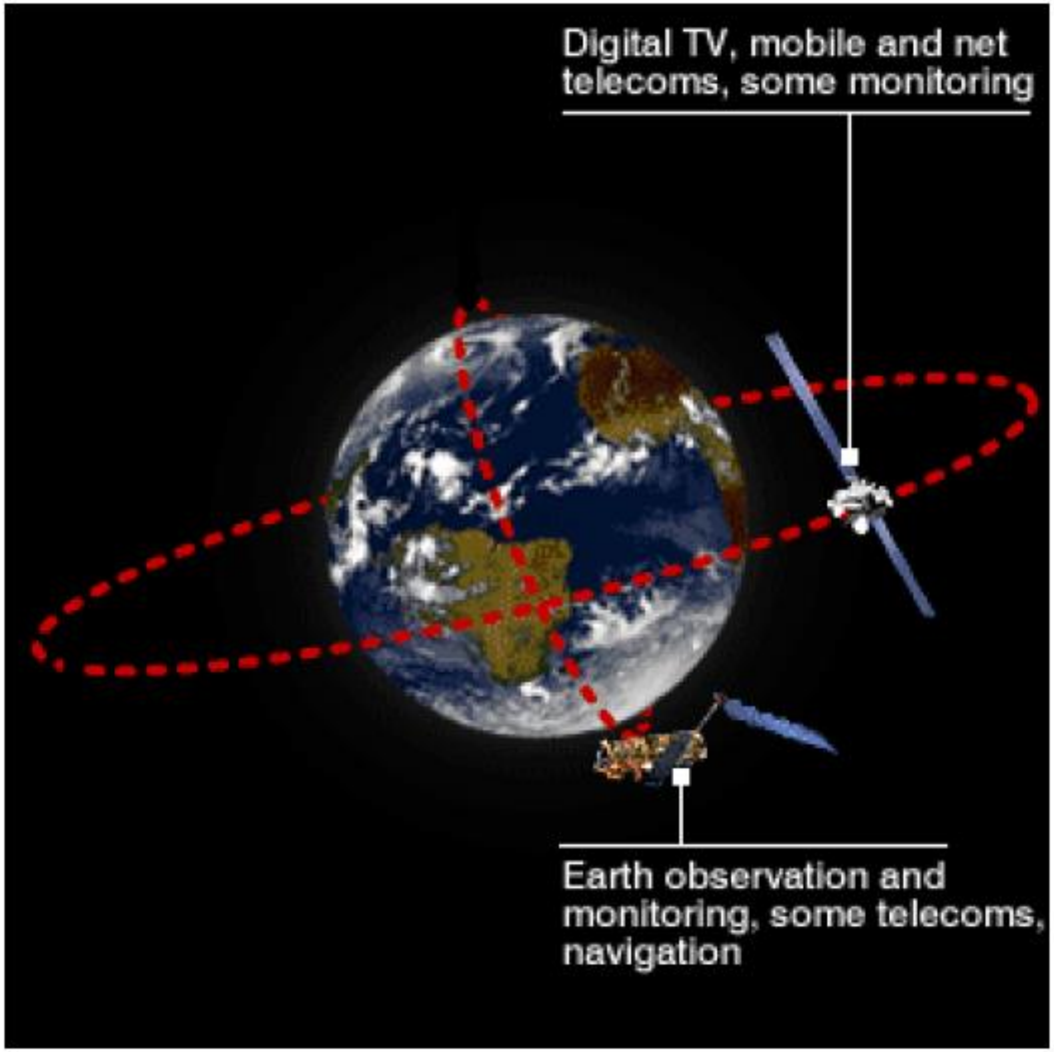
- Finds a complete assignment with maximum posterior probability

→Bayesian network



$$P(O, T) = \prod P^{error}(O_i \mid T_i) \times \prod P^{mendel}(T_i \mid parents(i)) \times \prod P^{freq}(T_i)$$

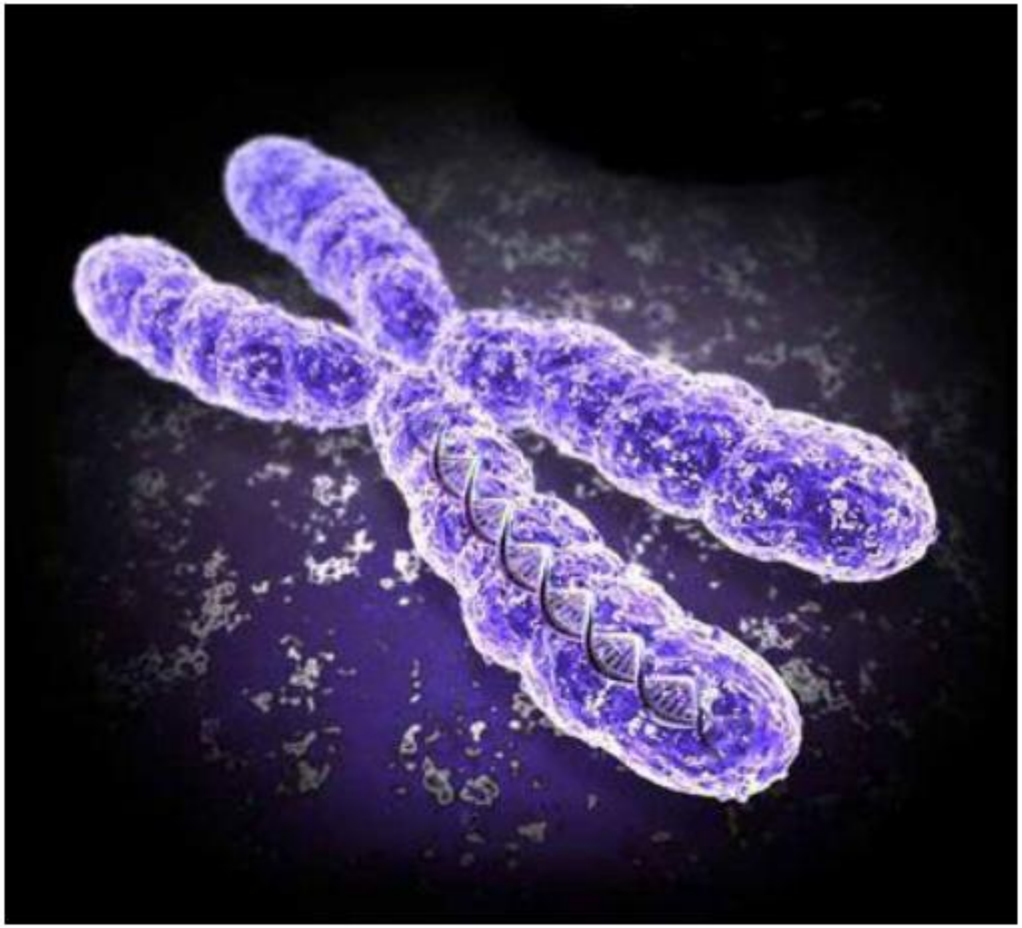
Earth Observation Satellite Management (SPOT5)



(Bensana *et al*, Constraints 1999 ; Sanchez *et a*l, IJCAI 2009)

$n \leq 364$, $d=4$, $e(2\text{-}3) \leq 10{,}108$

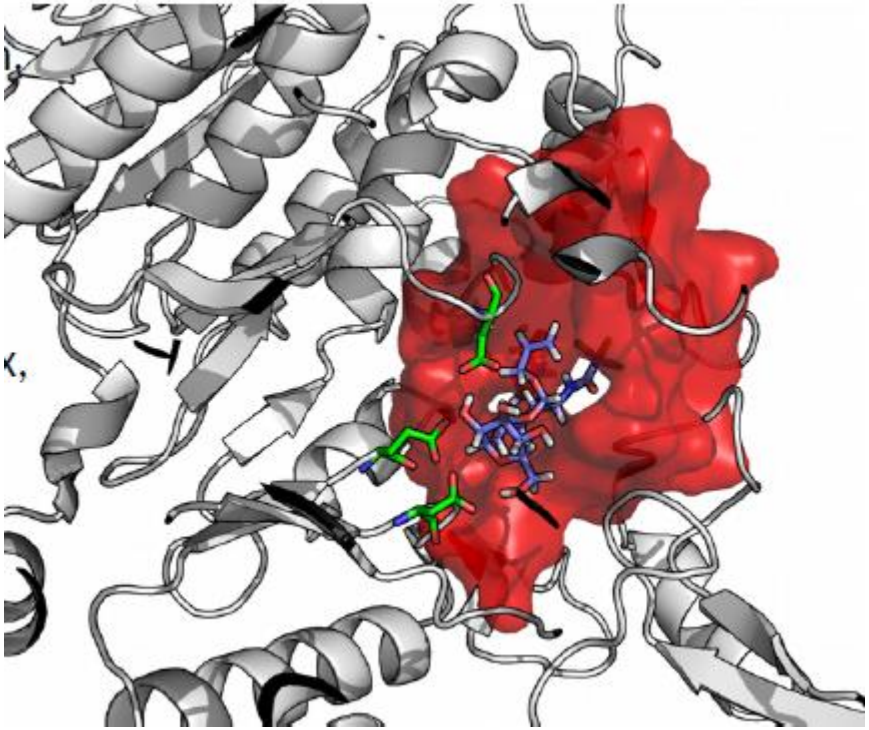
Genetic Linkage Analysis



(Marinescu & Dechter, AAAI 2006 ; Favier *et al*, IJCAI 2011)

n≤1,200, d≤7, e(2-5)≤2,000

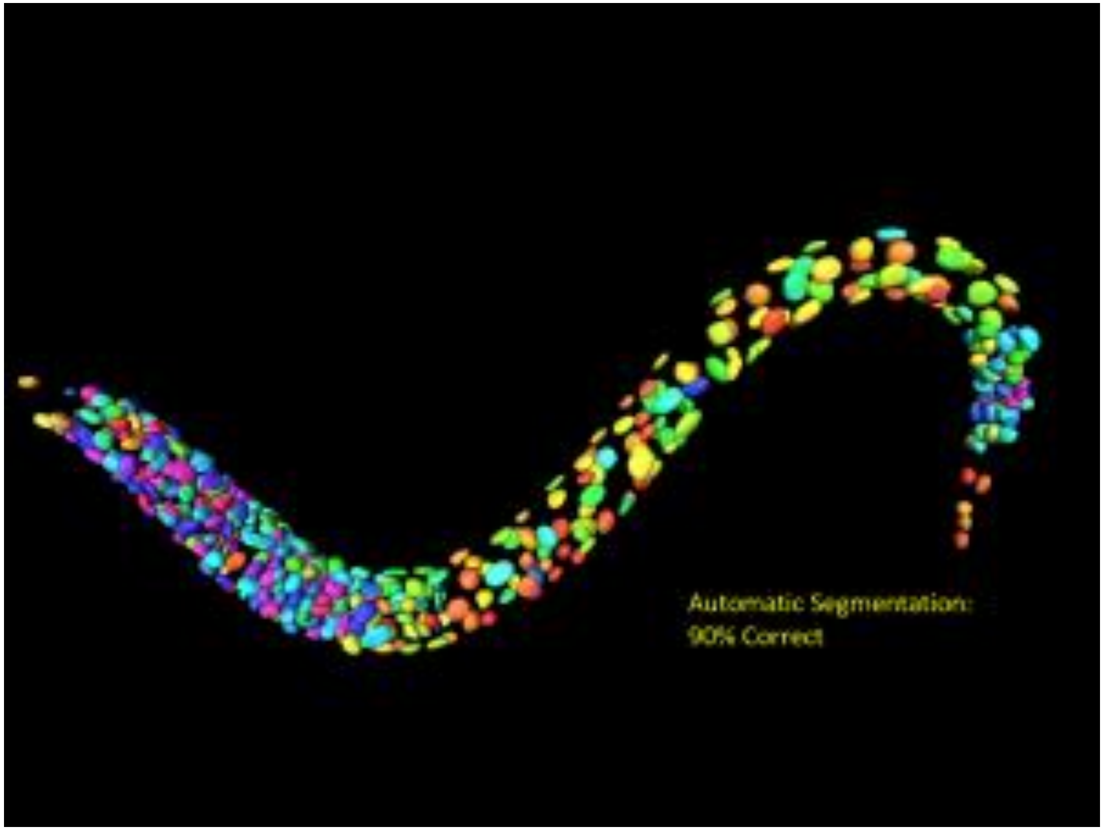
Protein Design



(Schiex *et al*, CP 2012 – Bioinformatics13 - **AIJ14** – JCTC15 – ISMP18)

$n \leq 120$, $d \leq 190$, $e(2) \leq 7{,}260$

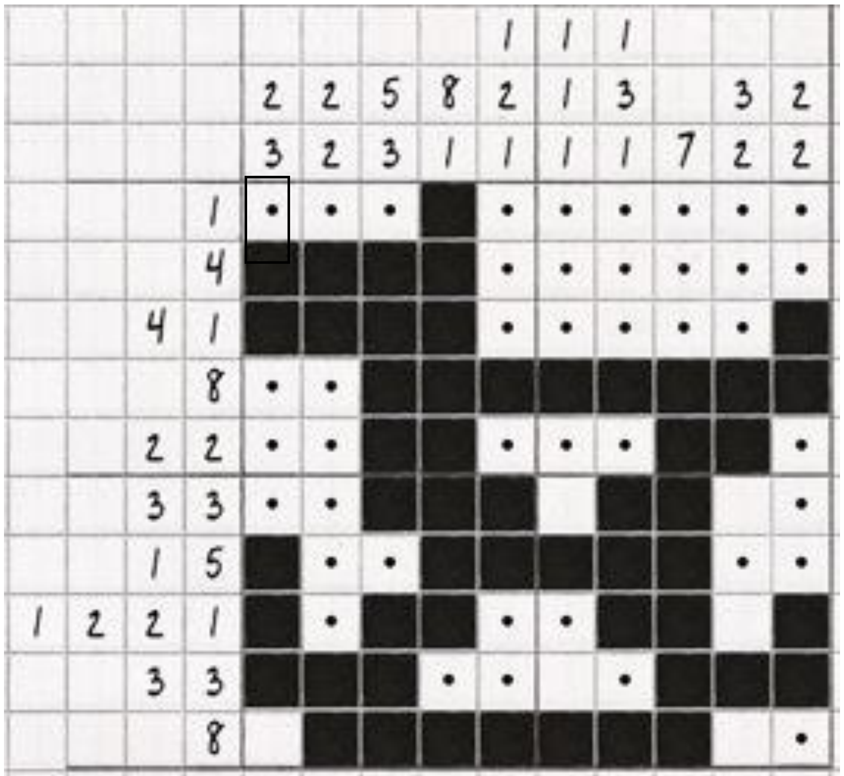
Graph Matching (worms segmentation)

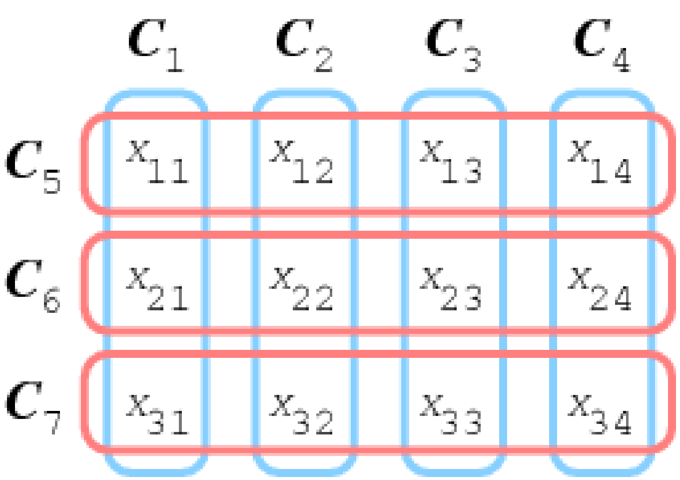


(Kainmueller *et al*, Med Image Comput 2014 ; Haller *et al*, AAAI 2018)

$n \leq 558$, $d \leq 128$, $e(2) \leq 23{,}407$

n × n Nonograms





- Each cell is black or white
- The lengths of successive black segments is fixed on every row and column (*NP-hard)*
- One boolean variable per cell
- Length specifications can be described by a regular language (□*■■■□*■■ □*)
- One Regular per row/column

★ *A DAC order compatible with all Berge-acyclic regulars exists*

White noise nonograms

2n-Regular, $n \times n$ $f(X_i)$ unary costs by cell

Size	choco (scalar)		toulbar2	
	Solved	Time	Solved	Time
20×20	100%	1.88	100%	0.93
25×25	100%	14.78	100%	3.84
30×30	96%	143.6	96%	99.01
35×35	80%	459.9	94%	218.2
40×40	46%	1,148	66%	760.8
45×45	14%	1,627	32%	1.321

(Schiex *et al*, AAAI 2012)

CPU limit: ½ hour

Crop allocation problem

(Akplogan *et al*, RAIRO 2013)

Spatio-temporal assignment of crops to landunits over a fixed horizon H, by enforcing farmer’s hard and soft constraints

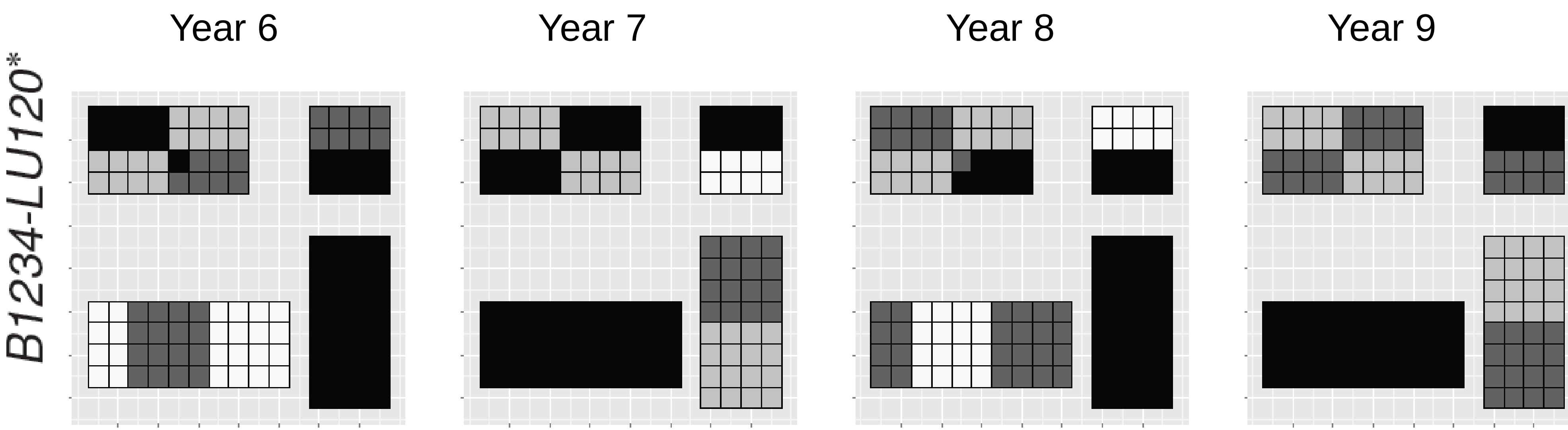


FIGURE: Crop allocation : Winter wheat, Spring barley, Maize, Winter rape

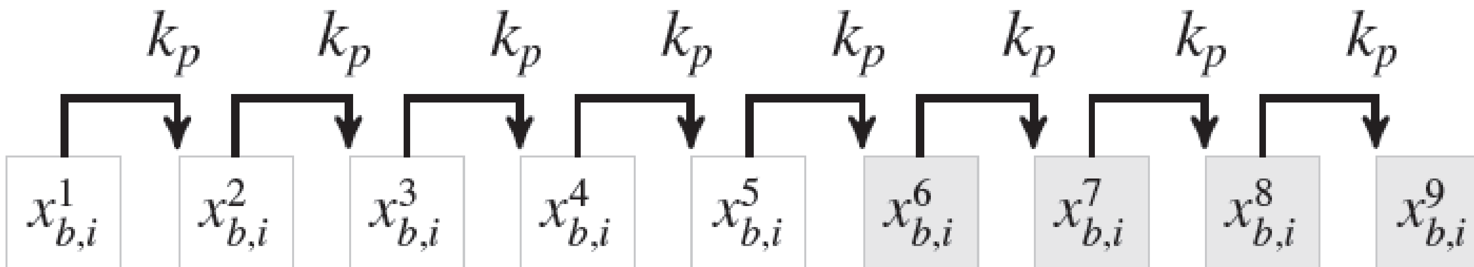
Constraints and preferences

	Category	Name	Scale	Hard/Soft	
Space	Management Unit	Growing area	All	Hard	**1**
		Farm topology	Blocks	Soft	**5**
		Crop acreages	All	Soft	**N**
		Landunit equality	Plot	Hard	**B**
	Resources	Resources capacity	All	Hard	**B**
Time	Agronomic	Minimum return time	All	Hard	**H**
		Crop rotation	Block	Hard	**H**
		Historic	All	Hard	**1**
		Preceding effects costs	Block	Soft	**2**
	Management unit	Same crops assigned	Block	Hard	**2H**
		Crop acreages in sequence	All	Soft	**H**

H : time horizon
B : number of landunits in a block
N : total number of landunits

Crop allocation problem as a CFN

$\mathcal{X}$ Variables $x^t_{b,i}$ define the landunit i in block b at year t ($t \in [1, \mathcal{H}]$).



$\mathcal{D}$ the domains $D_{b,i}$ of variables $x^t_{b,i}$ is the set of possible crops

$\mathcal{W}$ five different types of constraints

- simple tabular cost functions (arity up to 5), (preceding effect costs)
- SAME global constraint, (same set of crops assigned to two landunits)
- REGULAR global constraint, (minimum returned time)
- GCC global cardinality constraint, (resource capacity constraint)
- SOFT-GCC soft global cardinality constraint. (crop balance)

Experimental results

Search effort for finding the optimum and proving optimality

	n	e	Optimum	toulbar2 Decomposed Time(s)	toulbar2 Decomposed Nodes	SCIP ILP Time(s)	SCIP ILP Nodes
B1234-LU15*	135	465	1,824	**0.11**	38	0.12	1
B1234-LU30*	270	922	3,660	**0.40**	92	0.79	2
B1234-LU60*	540	1,804	7,492	**2.19**	210	2.64	8
B1234-LU120*	1,080	3,568	14,800	23.90	3,153	**7.25**	9

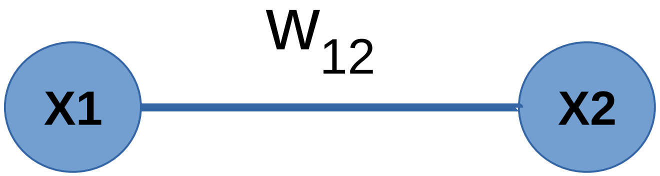
Each global cost function same & gcc is decomposed into 4 among.

1. 10,144 variables after decomposition, $\max_i |D_i| = 121$, needs 52 GB!
2. 85% CPU time used in preprocessing to remove functional variables.

Small example

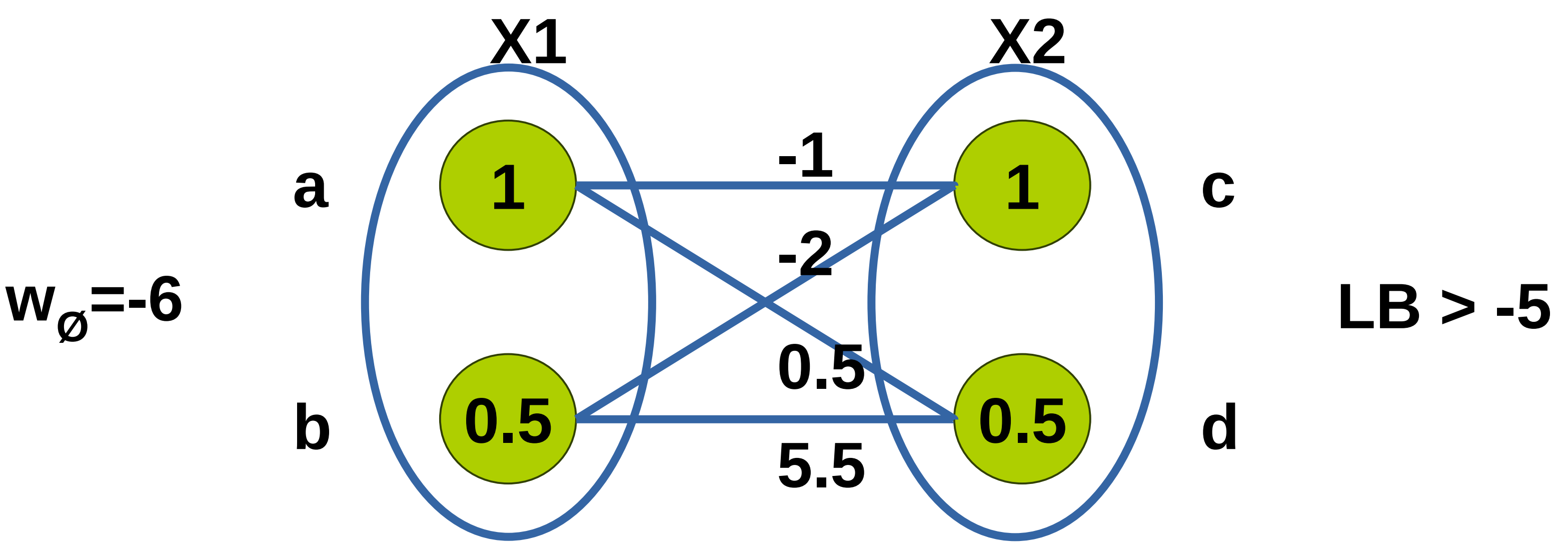
In JSON compatible toulbar2 *cfn* format

```
{
   problem: { name: "maximization", mustbe: ">-5.0"},
   variables: { "X1": ["a", "b"], "X2": ["c", "d"] },
   functions: {
    "w0": {scope: [], costs: [-6.0]},
    "w1": {scope: ["X1"], costs: [1.0, 0.5]},
    "w2": {scope: ["X2"], costs: [1.0, 0.5]},
    "w12": {scope: ["X1", "X2"], costs: [-1.0, 0.5, -2.0, 5.5]}
   }
}
```



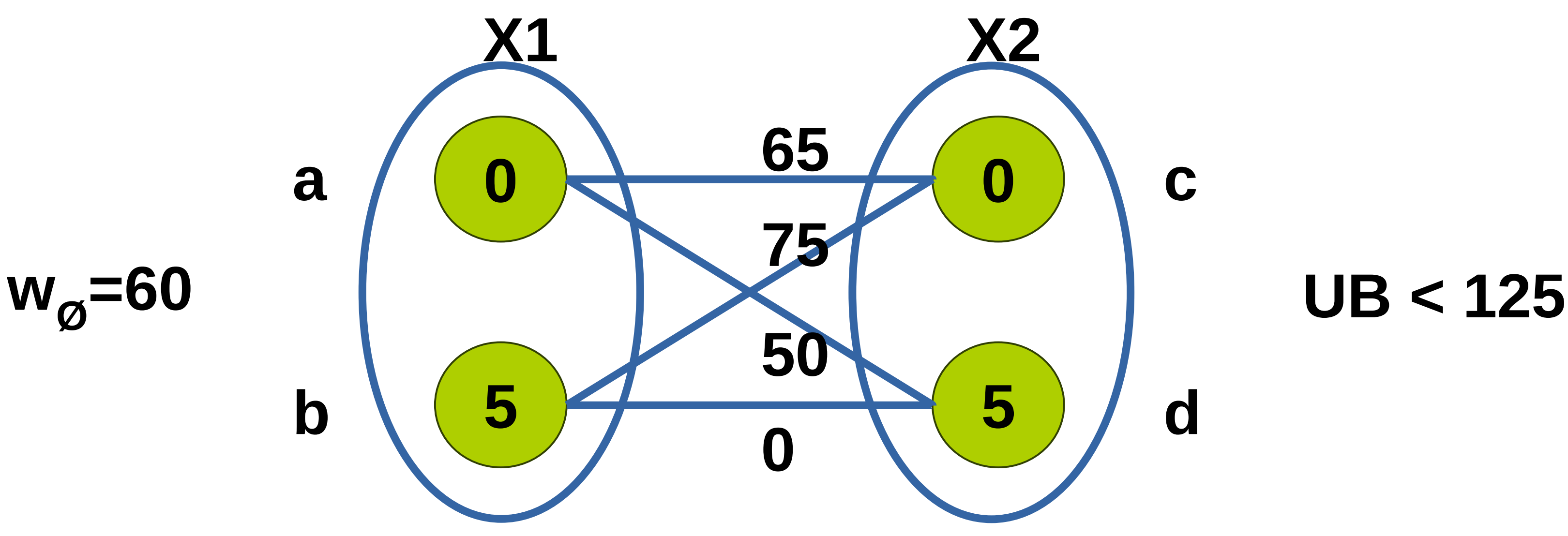
Micro-Structure

```
{
    problem: { name: "maximization", mustbe: ">-5.0"},
    variables: { "X1": ["a", "b"], "X2": ["c", "d"] },
    functions: {
     "w0": {scope: [], costs: [-6.0]},
     "w1": {scope: ["X1"], costs: [1.0, 0.5]},
     "w2": {scope: ["X2"], costs: [1.0, 0.5]},
     "w12": {scope: ["X1", "X2"], costs: [-1.0, 0.5, -2.0, 5.5]}
    }
}
```



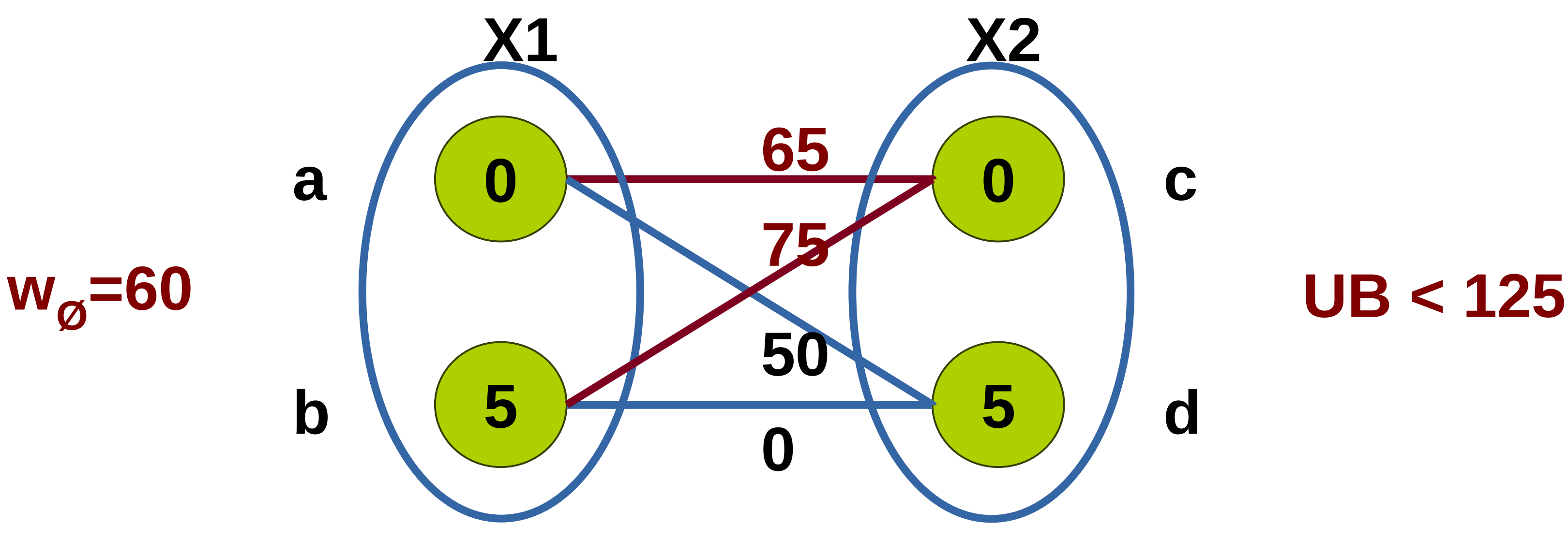
Minimization with non-negative integer costs

```
{
    problem: { name: "maximization", mustbe: ">-5.0"},
    variables: { "X1": ["a", "b"], "X2": ["c", "d"] },
    functions: {
     "w0": {scope: [], costs: [-6.0]},
     "w1": {scope: ["X1"], costs: [1.0, 0.5]},
     "w2": {scope: ["X2"], costs: [1.0, 0.5]},
     "w12": {scope: ["X1", "X2"], costs: [-1.0, 0.5, -2.0, 5.5]}
    }
}
```



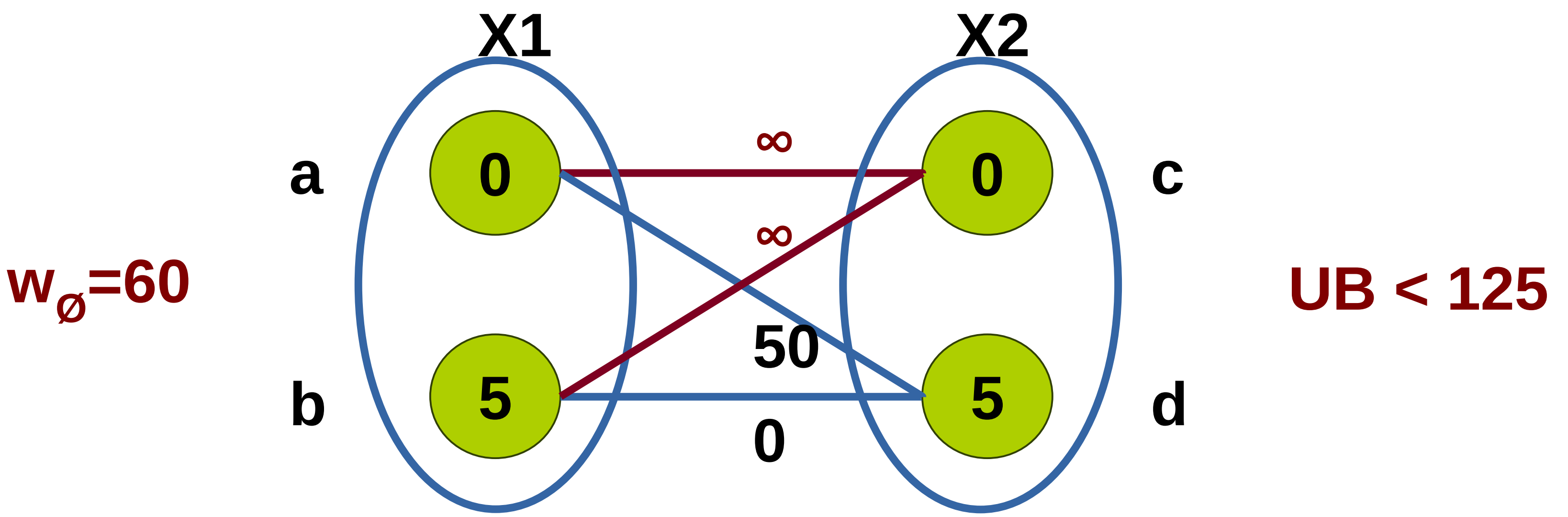
Constraints are Cost Functions

```
{
    problem: { name: "maximization", mustbe: ">-5.0"},
    variables: { "X1": ["a", "b"], "X2": ["c", "d"] },
    functions: {
     "w0": {scope: [], costs: [-6.0]},
     "w1": {scope: ["X1"], costs: [1.0, 0.5]},
     "w2": {scope: ["X2"], costs: [1.0, 0.5]},
     "w12": {scope: ["X1", "X2"], costs: [-1.0, 0.5, -2.0, 5.5]}
    }
}
```



Constraints are Cost Functions

```
{
   problem: { name: "maximization", mustbe: ">-5.0"},
   variables: { "X1": ["a", "b"], "X2": ["c", "d"] },
   functions: {
    "w0": {scope: [], costs: [-6.0]},
    "w1": {scope: ["X1"], costs: [1.0, 0.5]},
    "w2": {scope: ["X2"], costs: [1.0, 0.5]},
    "w12": {scope: ["X1", "X2"], costs: [-1.0, 0.5, -2.0, 5.5]}
   }
}
```



Other equivalent formulations

Input formats available in toulbar2

- WCSP

```
wcsp 2 2 4 125
2 2
2 0 1 0 4
0 0 65
0 1 50
1 0 75
1 1 0
1 0 125 2
0 0
1 5
1 1 125 2
0 0
1 5
0 60 0
```

MRF(.uai)

```
MARKOV
2
2 2
4
2 0 1
1 0
1 1
1 0
4
0.000341454887383
0.00215443469003
0.0001
1.0
2
1.0
0.541169526546
2
1.0
0.541169526546
2
0.00063095734448
0.00063095734448
```

- Max-SAT

```
p wcnf 2 7 125
65 1 2 0
50 1 -2 0
75 -1 2 0
5 -1 0
5 -2 0
60 1 0
60 -1 0
```

- QPBO

```
4 13
1 3 32.5
1 4 25
2 3 37.5
2 2 5
4 4 5
1 1 60
2 2 60
1 1 -1000
2 2 -1000
1 2 1000
3 3 -1000
4 4 -1000
3 4 1000
```