Multi-Language Evaluation of Exact Solvers in Graphical Model Discrete Optimization (Summary)

Barry Hurley[1], Barry O'Sullivan[1], David Allouche[2], George Katsirelos[2], Thomas Schiex[2], Matthias Zytnicki[2], and Simon de Givry[2⋆]

[1] Insight Centre for Data Analytics, University College Cork, Ireland
firstname.lastname@insight-centre.org
[2] MIAT, UR-875, INRA, F-31320 Castanet Tolosan, France
firstname.lastname@toulouse.inra.fr

By representing the constraints and objective function in factorized form, graphical models can concisely define various NP-hard optimization problems. They are therefore extensively used in several areas of computer science and artificial intelligence. Graphical models can be deterministic, *e.g.,* Constraint Networks (in Minizinc mzn format) and weighted variants such as Cost Function Networks, aka Weighted Constraint Satisfaction Problems (wcsp), or stochastic, *e.g.,* Bayesian Networks and Markov Random Fields (uai). They optimize a sum *or* product of local functions (constraints being represented as functions with values in $\{0, \infty\}$ or $\{0, 1\}$ *resp.*), defining a joint cost or probability distribution for discrete variables. Simple transformations exist between these two types of models, but also with MaxSAT (wcnf) and linear programming (lp).

We report on a large comparison of exact solvers which are all state-of-the-art for their own target language. These solvers are all evaluated on deterministic and probabilistic graphical models coming from the Probabilistic Inference Challenge 2011, the Computer Vision and Pattern Recognition OpenGM2 benchmark, the Weighted Partial MaxSAT Evaluation 2013, the MaxCSP 2008 Competition, the MiniZinc Challenge 2012 & 2013, and the CFN-Lib, a library of Cost Function Networks.

3026 problems divided into 43 categories	DAOOPT	TOULBAR2	CPLEX	CPLEX$_{tuple}$	MAXHS	MAXHS$_{tuple}$	GECODE
Nb. of problems solved in less than 1 hour	1832	**2433**	1273	1862	1417	1567	202.
Borda-score (see Minizinc Chal., norm. by nb. of applicable categories)	2.08 [5]	**4.24 [1]**	3.01 [2]	2.86 [3]	2.66 [4]	1.65 [7]	1.84 [6]

All 3026 instances are made publicly available in five different formats (mzn, wcsp, uai, wcnf, lp) and seven formulations (two encodings for wcnf and lp)[3]. The results show that a small number of evaluated solvers are able to perform well on multiple areas. By exploiting the variability and complementarity of solver performances, we show that a simple portfolio approach can be very effective. This portfolio, based on TOULBAR2[4], MPLP2[5], and CPLEX 12.6, won the last Uncertainty in Artificial Intelligence (UAI) Evaluation 2014[6][7]. We hope that our proposed collection of benchmarks, readily available in many formats, will enrich the various competitions in CP, AI, and OR, leading to more robust solvers and new solving strategies.

⋆ Supported by grants SFI/10/IN.1/I3032, SFI/12/RC/2289, and the GenoToul Bioinfo. platform.

[3] http://genoweb.toulouse.inra.fr/~degivry/evalgm

[4] http://www.inra.fr/mia/T/toulbar2 (version 0.9.8, parameters *-A -V -dee -hbfs*)

[5] http://cs.nyu.edu/~dsontag/ (version 2)

[6] http://www.hlt.utdallas.edu/~vgogate/uai14-competition/leaders.html (MAP/Proteus)

[7] https://github.com/9thbit/uai-proteus