Centre for
Data Analytics

Insight

UAI Inference Competition 2014 MPE submissions: Proteus and Robin

Barry Hurley (Proteus) with support from
Simon de Givry, Thomas Schiex, David Allouche, and
George Katsileros (INRA/Robin)

25 July 2014

This work is supported by Science Foundation Ireland (SFI) Grant 10/IN.1/I3032 and FP7 FET-Open Grant 284715.
The Insight Centre for Data Analytics is supported by SFI Grant SFI/12/RC/2289.
This work has been funded by the Agence Nationale de la Recherche (ANR-10-BLA-0214).
We thank the GenoToul Bioinformatics Platform of INRA-Toulouse for providing computing resources and support.

DCU NUI Galway OÉ Gaillimh UCC University College Cork, Ireland Coláiste na hOllscoile Corcaigh UCD DUBLIN SFI Science Foundation Ireland

Submissions

Robin

Simon de Givry, Thomas Schiex, David Allouche, and George Katsileros – INRA Toulouse.

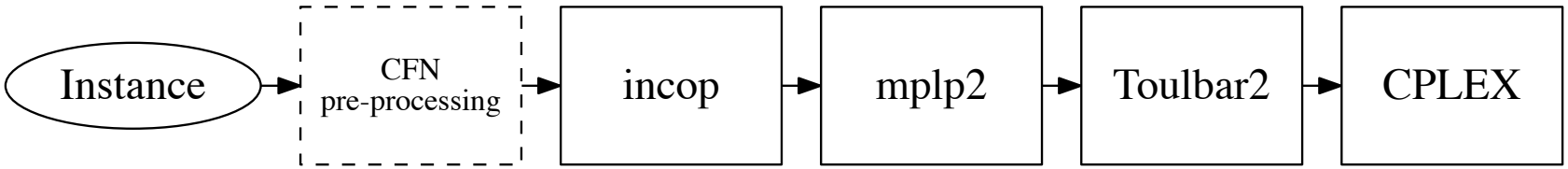
- Fixed schedule of four solvers: local search (incop), MRF (MPLP2), CFN (Toulbar2), ILP (CPLEX).

Proteus

Barry Hurley – Insight Centre for Data Analytics, UCC.

- builds upon a large experimental study by INRA on a large collection of graphical model instances from multiple paradigms.
- Portfolio of five complementary solvers
- Machine learning techniques select the solver on a per-instance basis

Robin



- CFN pre-processing by Toulbar2: local consistencies reformulation, dominance analysis, elimination of variables with small degree or unary domain.
- The best solution found is passed on to next.
- Fixed ordering and equal time share between solvers.

Experimental study of graphical model instances

	Solver	Num. solved to optimality	Num. Best	Misclass. Penalty Solved (Tot.Time)
1	Virtual Best (9)	2114		
2	Toulbar2	1983	1551	33 (12419.1)
3	CPLEX (tuple)	1819	53	6 (2480.8)
4	daoopt	1777	2	1 (275.2)
5	mplp2	1496	102	4 (729.1)
6	maxhs (tuple)	1425	1	0 (0.3)
7	CPLEX (direct)	964	254	40 (5791.1)
8	maxhs (direct)	880	135	8 (5202.4)
9	Mistral	208	11	0 (1.4)
10	Opturion CPX	189	0	0 (0.0)
11	Gecode	136	5	0 (0.1)

2556 instances from CFN (WCSP), CP, MRF, CVPR, Max CSP, WPMS.

`http://genoweb.toulouse.inra.fr/~degivry/evalgm/`

Proteus Details

Training Phase

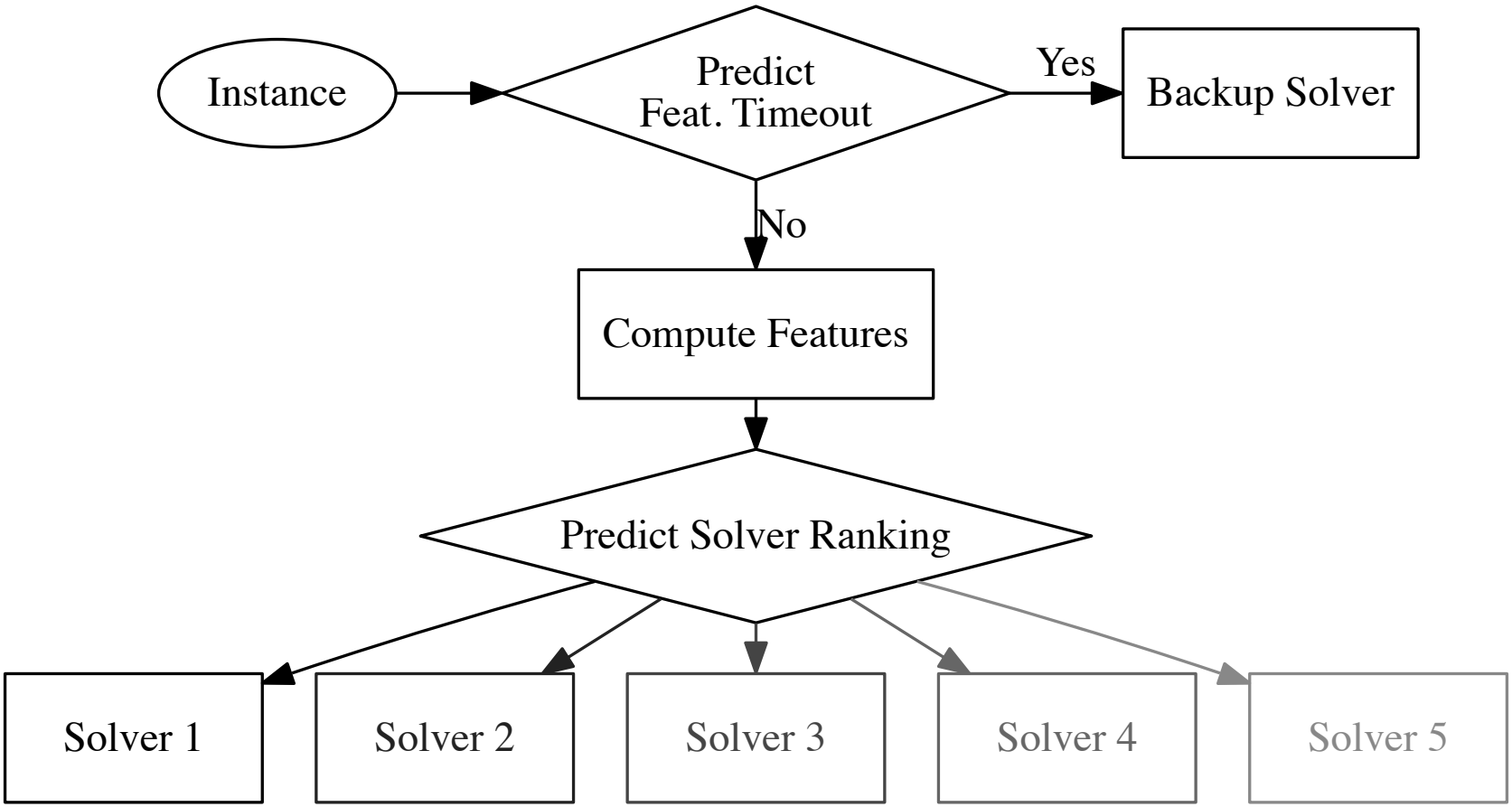
- Solvers: Toulbar2, incop + Toulbar2, mplp2, CPLEX (direct), CPLEX (tuple).
- Collect performance matrix of solvers on training instances
- Collect feature data of instances
- Learn a simple model to predict feature computation error
- Learn an empirical hardness model (random forest)

Runtime

- Static instance features: $|X|$, $|\Phi|$, statistics about domain size and arities, density of unary, binary, and ternary functions, ...
- Run the solvers in order of the predicted ranking, portfolio exits if the solver produced a valid output.

`https://github.com/9thbit/uai-proteus`

Proteus Portfolio



Features Importances

Feature	Gini Importance
Initial upper-bound	0.12575
Arity - coefficient of variation	0.09898
Density of binary functions	0.09639
Arity - std. deviation	0.09494
Filesize	0.08124
Time to read problem	0.07833
Time to compute UB	0.07760
Number of functions	0.06977
Number of variables	0.05658
Mean arity	0.04400
Mean domain size	0.04274
Maximum domain size	0.03714
Minimum domain size	0.02858
...	

Table : Mean feature importance across all decision trees.

Conclusions

Conclusions

- Complementarities among solvers is what brought Proteus and Robin to the podium.
- There is very little co-operation between solvers currently, lots of room for improvement.

Suggestions

- Account for proving optimality in the score (false claims detected by comparing optimum to other solvers, invalid solvers ruled out).
- All zeros is often the optimal solution, so values should be shuffled in the instances in future.
- Possibly start one month earlier next year?
- Continuous evaluation could be very useful for the community.

Bibliography I


Cooper, M. C., de Givry, S., Sanchez, M., Schiex, T., Zytnicki, M., and Werner, T. (2010).
Soft arc consistency revisited.
Artificial Intelligence, 174(7-8):449–478.


Hurley, B., Kotthoff, L., Malitsky, Y., and O'Sullivan, B. (2014).
Proteus: A Hierarchical Portfolio of Solvers and Transformations.
In *CPAIOR*, pages 301–317.


Kotthoff, L. (2014).
Algorithm selection for combinatorial search problems: A survey.
AI Magazine.


Neveu, B., Trombettoni, G., and Glover, F. (2004).
ID Walk: A Candidate List Strategy with a Simple Diversification Device.
In *Principles and Practice of Constraint Programming – CP 2004*.


Sontag, D., Choe, D. K., and Li, Y. (2012).
Efficiently Searching for Frustrated Cycles in MAP Inference.