UAI 2014 Probabilistic Inference Competition

Vibhav Gogate

The University of Texas at Dallas

How it works?

- **Four Tasks**
  - PR: Partition function estimation (sum-product problem)
  - MAR: Marginal estimation (Ratio of two sum-product problems)
  - MAP: Maximum-a-posterior inference (optimization)
  - **MMAP: Marginal MAP (optimization + sum-product problem)**
    - **NEW**
- **Three time bounds**
  - 20 seconds
  - 20 minutes
  - 1 hour
- **12 Total Categories**

Participation

- 8 Countries
- 12 teams

Submit your solver using the form below

Email*

Solver Name*

Passcode*

Upload your Solver (Linux 64-bit binary only)*

Choose File | No file chosen

Submit

- You can submit statically linked Linux 64-bit binaries
- Automatically run on Intel i7 machines with 16GB RAM
  - 120 machines (Courtesy of University of Texas at Dallas and the DARPA PPAML program)

What happens after you submit?

Click here

MAR

Solver	0 secs	20 mins	1 hour
bp	2380884 (0/136)	29.6825611571 (0/136)	29.073690003 (0/136)
ai	36354462598 (0/136)	0.270294708679 (0/136)	0.110956812986 (0/136)
incbp	3.73903084034 (3/136)	2.06104839502 (1/136)	1.74589115437 (0/136)
edbp	3.27285560022 (0/136)	1.36757343164 (0/136)	0.612348461643 (0/136)

You see your results

MAR

Problem	20 secs	20 mins	1 hour
1	5.55296832448e-06	6.17422748948e-06	6.17422748948e-06
2	0.0249092492619	0.0	0.0
3	0.0	0.0	0.0
4	0.0	0.0	0.0
5	0.00421080674536	0.0	0.0
6	0.0723015289956	0.0	0.0
7	0.00491858423473	0.0	0.0
8	1.72104203283e-05	0.0	0.0
9	7.59596769663e-05	0.0	0.0
10	0.0	0.0	0.0
11	0.000181550907783	0.0	0.0
12	3.48659756934e-07	3.56760114935e-07	3.56760114935e-07
13	6.9468517587e-06	0.0	0.0

Benchmarks

- Approximately 130 problems for each category
- Diverse Domains
  - Grid (Ising models)
  - Medical Diagnosis
  - Protein-Protein Interaction
  - Hard SAT problems
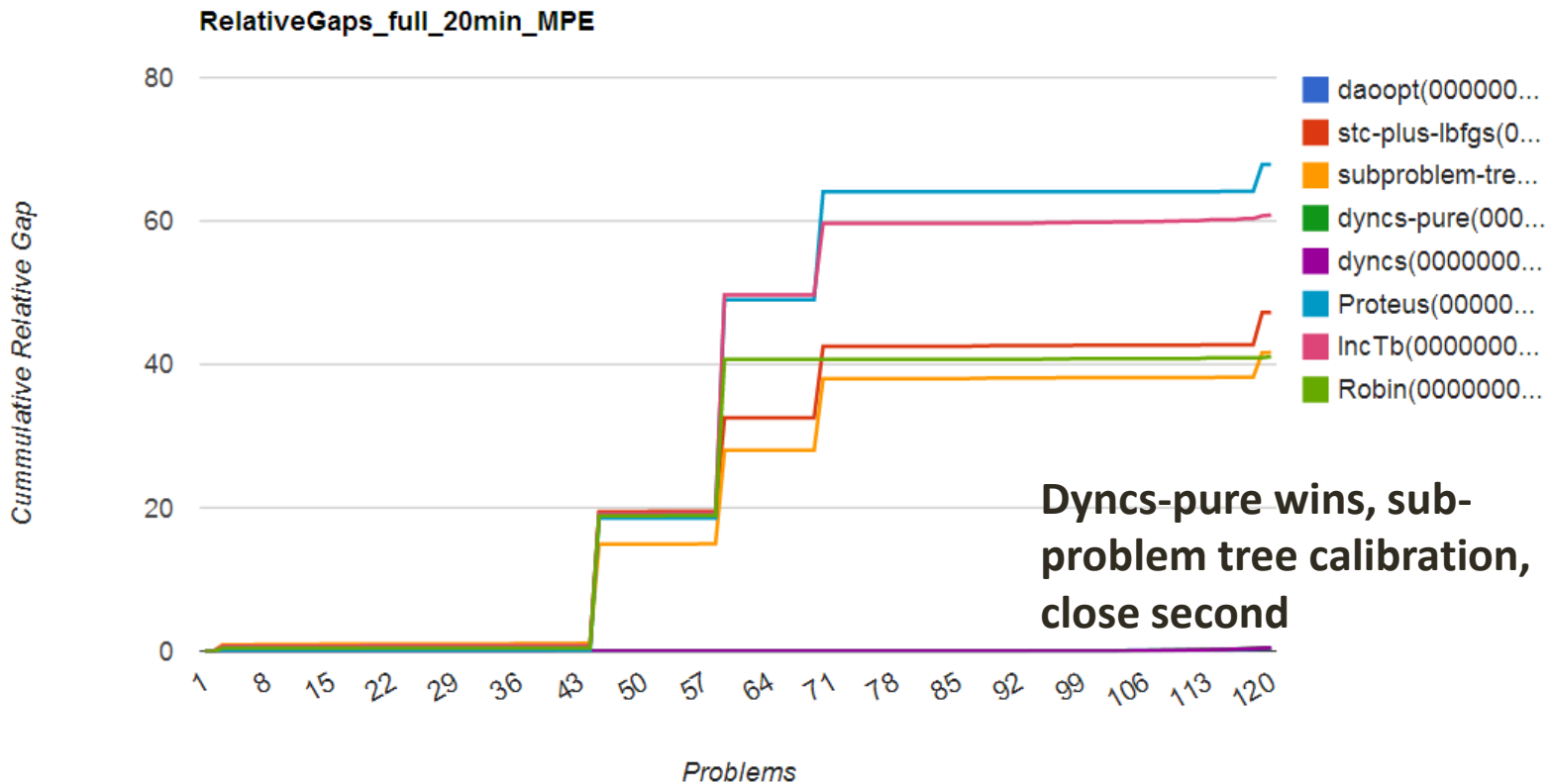 - Radio-frequency assignment problems
  - And so on

Evaluation Metrics used

- PR: $|\log(Z_{solver}) - \log(Zex_{act})|$
- MAR:
  - Max-absolute error
  - Hellinger Distance
- MAP (and MMAP)
  - Weight of the assignment output by the solver = $\sum_i \log(\phi_i(x))$
  - Relative Gap = abs(best-solver)/abs(best)
- MMAP evaluation is harder
  - MMAP Requires exact inference to compute the weight
- Four MMAP problem types
  - Hard Sum, Hard Optimization
  - Easy Sum, Hard Optimization
  - **Hard Sum, Easy Optimization (we ran only these types of problems)**
  - Easy Sum, Easy Optimization
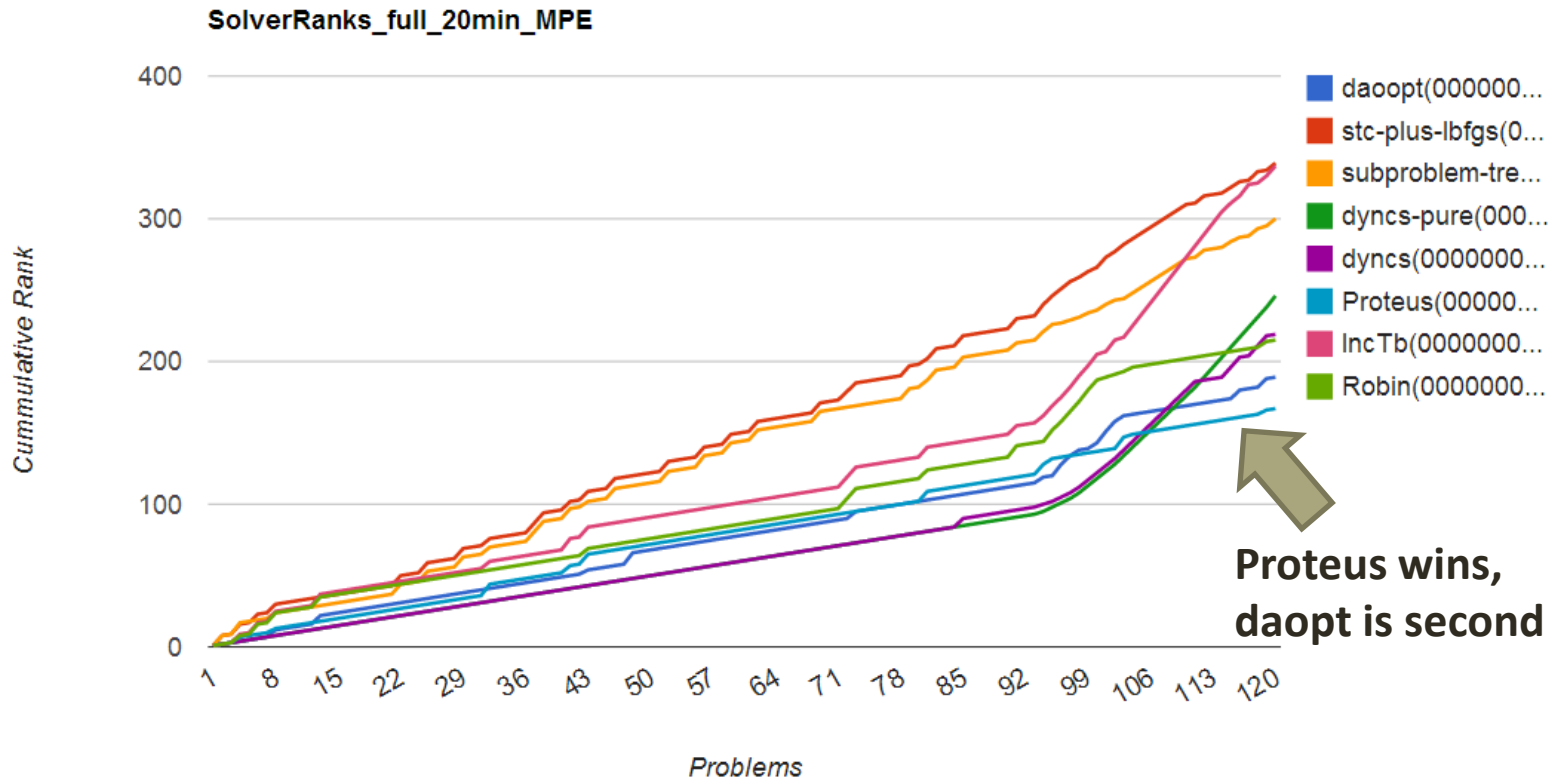
Evaluation: Deciding Winners!

- Easy if you just have one graphical model and many solvers
  - The solver with the best score/smallest error wins
- Aggregating results and declaring a winner over many problems is hard!
  - Decision theory.
- Borda counts
  - For each problem
    - Best solvers get 1 point
    - Second best solvers get 2 points and so on
  - Solver with the minimum points wins
- Many problems with this scoring
  - **Future research**
  - **For MAP Inference, I really don't know who won!!!**
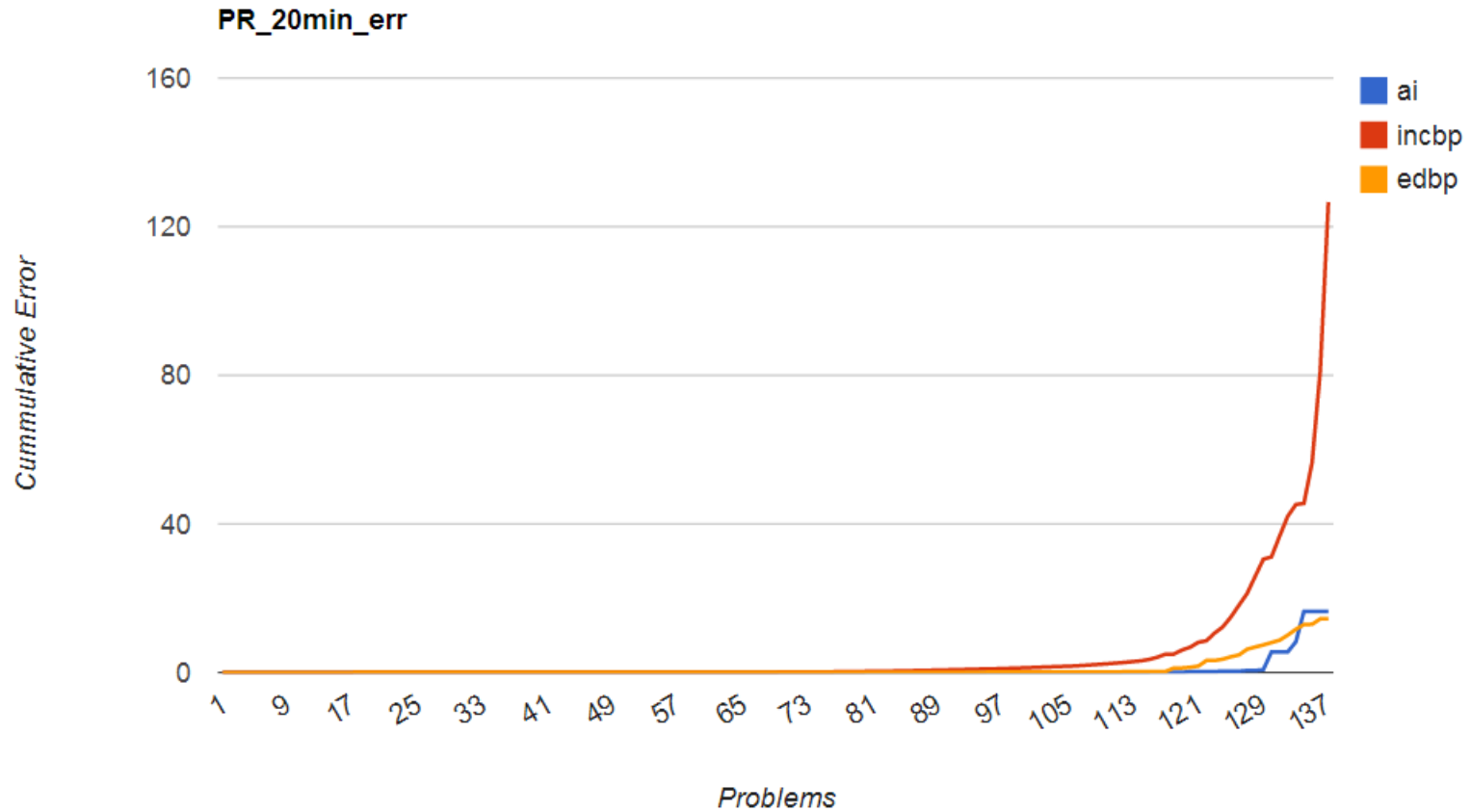
Picking Winners is hard

RelativeGaps_full_20min_MPE

Legend:
- daoopt(000000...
- stc-plus-lbfgs(0...
- subproblem-tre...
- dyncs-pure(000...
- dyncs(0000000...
- Proteus(00000...
- IncTb(0000000...
- Robin(0000000...

Dyncs-pure wins, sub-problem tree calibration, close second

Evaluation Criteria: Relative Gap = $\dfrac{Abs(best - solver)}{Abs(best)}$

Picking Winners is hard

SolverRanks_full_20min_MPE

Legend:
- daoopt(000000...
- stc-plus-lbfgs(0...
- subproblem-tre...
- dyncs-pure(000...
- dyncs(0000000...
- Proteus(00000...
- IncTb(0000000...
- Robin(0000000...

Y-axis: Cummulative Rank (0, 100, 200, 300, 400)

X-axis: Problems (1, 8, 15, 22, 29, 36, 43, 50, 57, 64, 71, 78, 85, 92, 99, 106, 113, 120)

Proteus wins, daopt is second

Borda count

PR Results

PR_20min_err

MAR Results

MAR_20min_maxabs

MMAP Results

SolverRanks_full_20min_MMAP
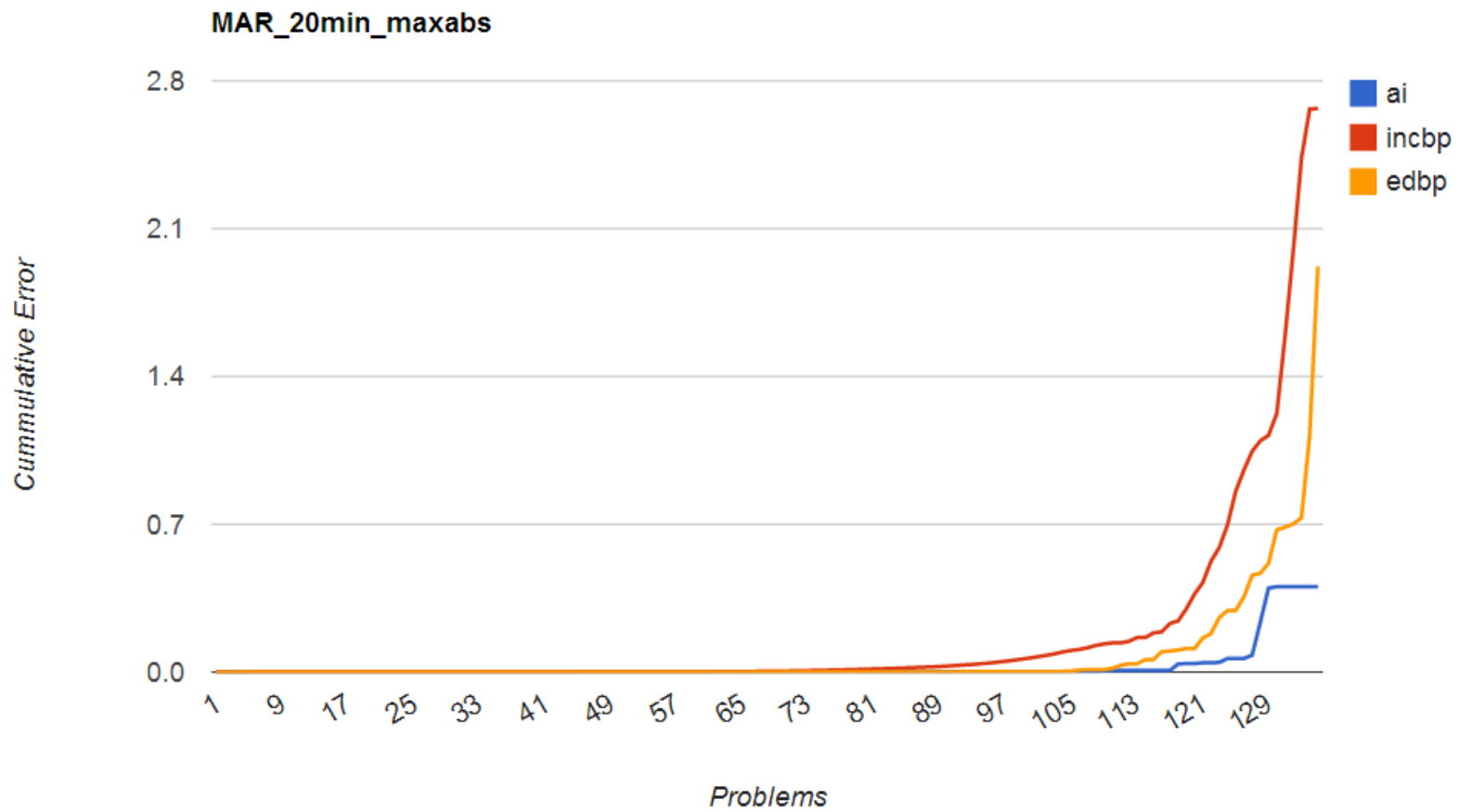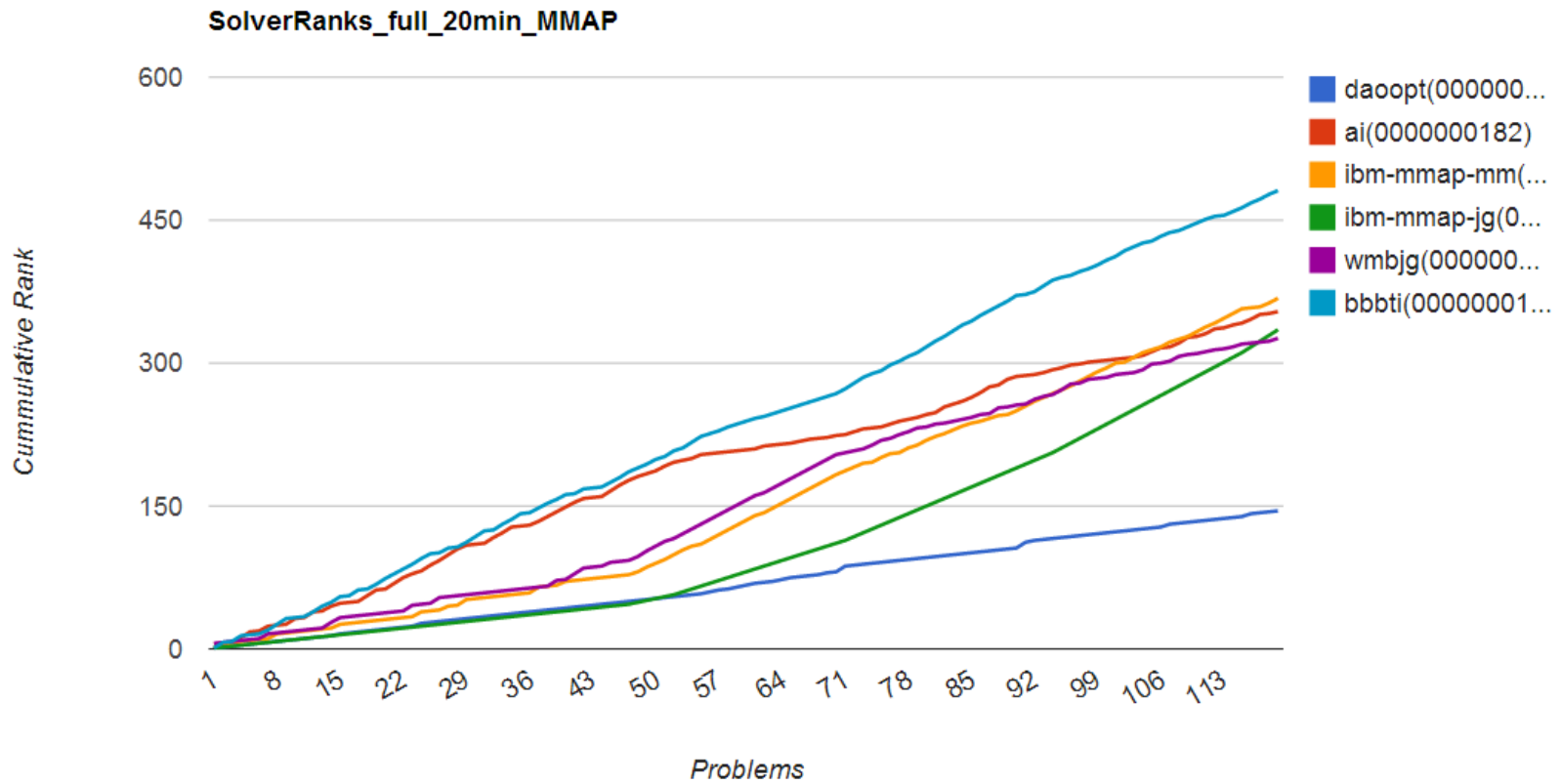
Winners

	20 seconds	20 minutes	1 hour
PR			
First-prize	Alex Ihler	Arthur Choi	Alex Ihler
Second-prize	Arthur choi	Alex Ihler	Arthur choi
MAR			
First-prize	Alex Ihler	Alex Ihler	Alex Ihler
Second-prize	Arthur choi	Arthur choi	Arthur choi
MAP			
First-prize	David Allouche	Barry Hurley	Barry Hurley
Second-prize	Barry Hurley	Lars Otten	Lars Otten
Third-prize	David Allouche	David Allouche	David Allouche
MMAP			
First-prize	Lars Otten	Lars Otten	Lars Otten
Second-prize	Radu Marinescu	Radu Marinescu	Radu Marinescu

Solver Types

- PR and MAR: Generalized Belief propagation based approaches
- MAP:
  - Branch&Bound search
  - Weighted CSP approaches
  - Subtree-calibration/LBFGS
- MMAP
  - Branch&Bound search

Post-Competition

- We need hard instances for PR and MAR
  - Recently added a few for PR

PR   **OLD/Easier Problems**

Solver	20 secs	20 mins	1 hour
ai	8.32295 (0/138)	16.5076 (0/138)	14.1457 (0/138)
incbp	204.813197306 (1/138)	126.635230766 (0/138)	85.0328975405 (0/138)
edbp	40.6246051 (0/138)	14.4681677 (1/138)	19.1625978 (1/138)
aomdd-pr	1.34165 (110/138)	9.534046 (99/138)	20.79951 (100/138)

PR   **Hard Problems**

Solver	20 secs	20 mins	1 hour
edbp	7303.29922458 (14/32)	9620.24741961 (9/32)	12008.9214312 (12/32)
aomdd-pr	3.12093055939 (29/32)	20.3789675843 (29/32)	7.392223089 (29/32)
ai	2207.30842613 (1/32)	718.347240185 (1/32)	662.59748887 (1/32)
incbp	4576.12378072 (9/32)	2075.39520114 (6/32)	987.872324127 (6/32)

Rest of the Session

- Winner Certificates
- Huayan Wang's MAP Inference solver
- PR and MAR Inference
    - EDBP
- MAP Inference
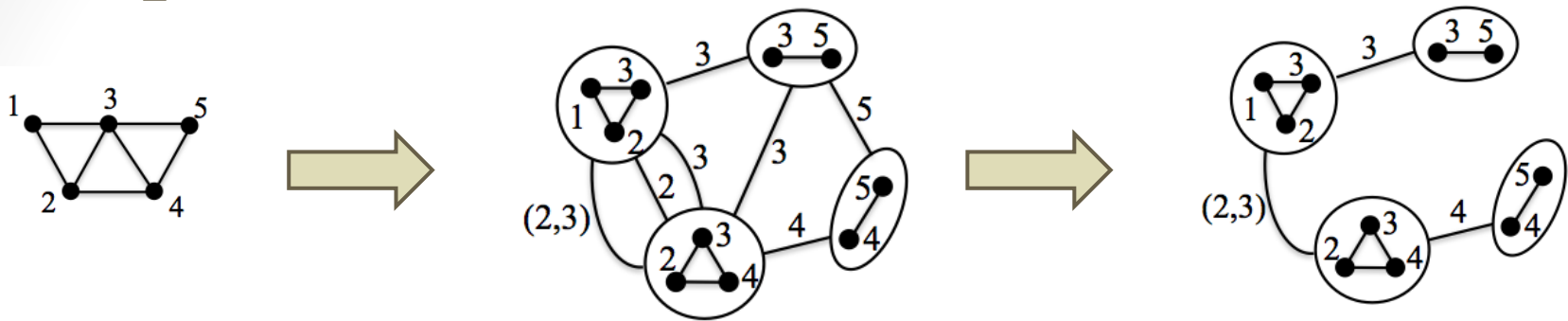   - Radu Marinescu
    - Barry Hurley

Thank you!!
Amir Globerson, Gal Elidan, Rina Dechter, and Alex Ihler
My students who ran the competition
- David Smith, Li Chou, Somdeb Sarkhel, Tahrima Rahman, Srikanth Doss, and Deepak Venugopal.

Huayan Wang and Daphne Koller

Subproblem tree calibration

Algorithm

Given MRF (left figure)

Split into subproblems (dual decomposition)

Build a multi-graph with a node for each subproblem (middle figure)

Repeat

      Randomly choose a subproblem-tree (right figure)

      "Calibrate" the tree by max-product / min-sum message passing

Performed very well; very close to winning

Properties

- Each tree calibration is a block coordinate descent step for the dual problem.
- The "block" corresponds to all edges in the subproblem-tree.
- Subsumes MPLP, TRW-S, and max-sum diffusion as special cases.
- Handles larger and more flexible "blocks" than these methods.

Subproblem tree calibration

- References

  - Huayan Wang and Daphne Koller: **Subproblem-Tree Calibration: A Unified Approach to Max-Product Message Passing**, *The 30th International Conference on Machine Learning* (**ICML 2013**)

  - Huayan Wang and Daphne Koller: **A Fast and Exact Energy Minimization Algorithm for Cycle MRFs**, *The 30th International Conference on Machine Learning* (**ICML 2013**)

- Download code at
  - http://ai.stanford.edu/~huayanw/

The Solver

PR and MAR inference based on the **RCR framework**:

- **RELAX** edges until model is tractable (until it is a tree)
- **COMPENSATE** for the relaxation
- **RECOVER** edges and improve the approximation

Anytime, Generalized BP Algorithm. See Choi & Darwiche 11.

The Results

this year: won UAI-14 MAR-20m
previously: won UAI-10 PR-20s, MAR-20s.
In SamIam system: `http://reasoning.cs.ucla.edu/samiam`

Beyond The Solver (Future Competitions?)

LIFTED approximate inference (lifted RCR)

- see Van den Broeck, Choi & Darwiche UAI-12
- `http://dtai.cs.kuleuven.be/ml/systems/wfomc`

INCREMENTAL COMPILATION for approximate inference

- using Sentential Decision Diagrams (new class of ACs) (Darwiche IJCAI-11, C,K&D ECSQARU-13)