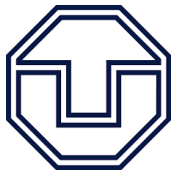
The 2nd International Competition on Computational Models of Argumentation

S. Gaggl, T. Linsbichler, M. Maratea, S. Woltran



The 2017 International Workshop on Theory and Applications of Formal Argument (TAFA 2017)

Outline

The 2nd ICCMA Competition

Organization

- Sarah A. Gaggl, TU Dresden, Germany
- Thomas Linsbichler, TU Wien, Austria
- Marco Maratea, University of Genova, Italy
- Stefan Woltran, TU Wien, Austria

ICCMA Steering Committee

Matthias Thimm (President)
Hannes Strass (Vice-President)
Federico Cerutti (Secretary)
Sarah A. Gaggl
Nir Oren
Mauro Vallati
Serena Villata

Webpage: `http://www.dbai.tuwien.ac.at/iccma17`

The 2nd ICCMA Competition

Two years after the 1st competition

- Hosted again by TAFA
- Continuing the work along the lines of the first event

The 2nd ICCMA Competition

Two years after the 1st competition

- Hosted again by TAFA
- Continuing the work along the lines of the first event

Goals

- Measure the progress of the state of the art in AF solving
- Improve benchmark suite with meaningful benchmarks
- Study the behavior of different solving techniques

The 2nd ICCMA Competition

Novelties

- Introduce new semantics and “Dung’s Triathlon” track
- Dedicated Call for Benchmarks
- Hardness-based classification of instances
  - Inspired by SAT- and ASP-Competitions
  - Exploiting best solvers from ICCMA 2015
- Introduce a new scoring scheme

Outline

System Competition Format

Semantics, Problems, Tasks, Tracks

7 semantics: complete, preferred, stable, **semi-stable**, **stage**, grounded, **ideal**

4 reasoning problems: Given an AF (and some argument):

- SE: determine some extension;
- EE: determine all extensions;
- DC: decide whether the argument is credulously accepted;
- DS: decide whether the argument is skeptically accepted.

Task: a reasoning problem under a particular semantics

Track: all tasks for a particular semantics

+ a special track (Dung's Triathlon - D3)

- EE-grounded + EE-stable + EE-preferred in one call
- goal is to test the solvers' capability of exploiting interrelationships between semantics

Setup

System Inputs

- Input- and output-format adopted from 1st edition
- Fixed input in TGF or APX format
- Scripts run with fixed parameters

System Environment

- Bull HPC-Cluster (Taurus)
  - Intel Xeon (Haswell) CPU (E5-2670) with 2.60GHz
  - from 16 cores we used every 4th
- Time limits (CPU time) for each instance
  - all track except Dung's Triathlon: 10 minutes
  - Dung's Triathlon track : 30 minutes
- Memory Limit: 6.5 GB for D3, 4 GB for all other tasks

Scoring

ICCMA 2017 Scoring Schema

For each instance I, a solver gets $Score(Solver, I)$ as follows:

- 1 point, if it delivers the correct result;
- -5 points, if it delivers an incorrect result;
- 0 points otherwise.

Task

$$Score(Solver, Task) = \sum_{I \in Task} Score(Solver, I)$$

Track

$$Score(Solver, Track) = \sum_{Task \in Track} Score(Solver, Task)$$

- All ties are broken by the total time spent on correct results.

Benchmarks Selection

Goal is to select the set of instances to be run, such that they correspond to an (expected) wide range of hardness.

The classification has been implemented through the following steps:

1. Grouping tasks according to “compatible complexity”.
2. Instance collection.
3. Instance classification.
4. Instance selection.

Task Grouping

Tasks are grouped according to “compatible complexity” of the respective tasks. The employed grouping is the following:

A. DS-PR, EE-PR, EE-CO
B. DC-ST, DS-ST, EE-ST, SE-ST, DC-PR, SE-PR, DC-CO
C. DS-CO, SE-CO, DC-GR, SE-GR
D. DC-ID, SE-ID
E. *-SST, *-STG

Groups **D** and **E** include the newly employed semantics.

Instance Collection

Overview

- Considered the domains from 1st edition: **GroundedGenerator**, **SccGenerator**, **StableGenerator**.
  - Instances generated with different parameters.
- Dedicated call for benchmarks
  - Received 6 submissions (5 generators, 3 sets)
  - Generators employed to obtain instance sets

$\Rightarrow$ 11 domains in total

Instance Collection: New Domains

- **AdmBuster: a benchmark example for (strong) admissibility**, by M. Caminada (Prifysgol Caerdydd, UK) and M. Podlaszewski (Talkwalker).
- **AFBenchGen2: A Generator for Random Argumentation Frameworks**, by F. Cerutti (Cardiff Univ., UK), M. Vallati (Univ. of Huddersfield, UK), and M. Giacomin (Univ. of Brescia, Italy).
- **Assumption-Based Argumentation Translated to Argumentation Frameworks**, by T. Lehtonen (Univ. of Helsinki, Finland), J.P. Wallner (TU Wien, Austria), and M. Järvisalo (Univ. of Helsinki, Finland).
- **Planning2AF: Exploiting Planning Problems for Generating Challenging Abstract Argumentation Frameworks**, by F. Cerutti (Cardiff Univ., UK), M. Giacomin (Univ. of Brescia, Italy), and M. Vallati (Univ. of Huddersfield, UK).
- **SemBuster: a benchmark example for semi-stable semantics**, by M. Caminada (Cardiff Univ., UK) and B. Verheij (Rijksuniversiteit Groningen, Netherlands).
- **Traffic Networks Become Argumentation Frameworks**, by M. Diller (TU Wien, Austria).

Instance classification

Classify the collected instances w.r.t. their expected level of difficulty.

- selection of a representative task:

  A: EE-PR; **B:** EE-ST; **C:** SE-GR.

- select "representative" solvers from ICCMA 2015:

  A: Cegartix, CoQuiAAS, Aspartix-V
  B: Aspartix-D, ArgSemSAT, ConArg
  C: CoQuiAAS, LabSATSolver, ArgSemSAT

  D,E: no reference solvers

- definition of hardness categories: Instances solved ...

  (very easy) by all repr. solvers in less than 6 seconds.
  (easy) by all repr. solvers in less than 60 seconds.
  (medium) by all repr. solvers within the timeout (600 sec.).
  (hard) by at least one repr. solver within 1200 seconds.
  (too hard) by none of the repr. solvers within 1200 seconds.

Instance selection

- Final benchmark set for each group contains 350 instances:
  - 50 very easy,
  - 50 easy,
  - 100 medium,
  - 100 hard,
  - 50 too hard.
- Distibution among domains as uniform as possible
- Groups **D** and **E**: same benchmark set as group **A**.
- No “very hard” instances for group **C**

  $\Rightarrow$ number of “hard” instances increased to 150
- One query argument selected for each instance (DC-*, DS-*)
  - none for “very easy”, two for “too hard”
  - guided for ideal semantics; otherwise randomly

Outline

Participants

The competition featured 16 systems

- **argmat-clpb**
- **argmat-dvisat**
- **argmat-mpg**
- **argmat-sat**
- **ArgSemSAT**
- **ArgTools**
- **ASPrMin**
- **cegartix**
- **Chimaerarg**
- **ConArg**
- **CoQuiAAS**
- **EqArgSolver**
- **gg-sts**
- **goDIAMOND**
- **heureka**
- **pyglaf**

- Compared to ICCMA 2015: 9 new, 7 updated
- At least 9 solvers for each task

Results

Current State (UPDATED 26-09-2017)

- All Tasks have been run
- 400000 CPU hours used for competition runs
- 11 TB of output produced
- Final results for all regular tasks and tracks published
- Validation of D3 still ongoing

GR Tasks

DC			SE		
Solver	**Points**	Time	**Solver**	**Points**	Time
CoQuiAAS	350	201.0	heureka	345	72.2
cegartix	350	660.7	CoQuiAAS	345	134.8
heureka	345		goDIAMOND	345	322.3
goDIAMOND	343		cegartix	345	491.8
argmat-dvisat	340	78.3	pyglaf	343	
argmat-clpb	340	128.6	argmat-dvisat	342	85.5
EqArgSolver	340	224.5	argmat-clpb	342	134.7
argmat-sat	340	253.2	argmat-sat	342	251.6
pyglaf	340	5825.9	EqArgSolver	342	278.3
ArgTools	336		ArgTools	338	
argmat-mpg	330		argmat-mpg	332	
ConArg	294		ConArg	294	
ArgSemSAT	272		ArgSemSAT	289	
gg-sts	-1624		gg-sts	-247	

GR Track

Solver	**Points**	Time
CoQuiAAS	695	335.8
cegartix	695	1152.5
heureka	690	
goDIAMOND	688	
pyglaf	683	
argmat-dvisat	682	163.8
argmat-clpb	682	263.2
EqArgSolver	682	502.8
argmat-sat	682	504.8
ArgTools	674	
argmat-mpg	662	
ConArg	588	
ArgSemSAT	561	
gg-sts	-1871	

Congratulations to the winner:

CoQuiAAS v2.0: Taking Benefit from Constraint Programming to Solve Argumentation Problems. **Jean-Marie Lagniez, Emmanuel Lonca, and Jean-Guy Mailly.**

ID Tasks

DC			SE	
Solver	**Points**	Time	**Solver**	**Points**
pyglaf	283		pyglaf	302
argmat-dvisat	228	9199.9	argmat-dvisat	265
cegartix	228	13935.3	argmat-sat	260
argmat-sat	217		goDIAMOND	218
goDIAMOND	196		ArgTools	157
ArgTools	111		cegartix	140
argmat-mpg	85		argmat-mpg	132
ConArg	67		ConArg	114
CoQuiAAS	-313		gg-sts	-313
gg-sts	-737		CoQuiAAS	-481

ID Track

Solver	Points
pyglaf	585
argmat-dvisat	493
argmat-sat	477
goDIAMOND	414
cegartix	368
ArgTools	268
argmat-mpg	217
ConArg	181
CoQuiAAS	-794
gg-sts	-1050

Congratulations to the winner:

The **pyglaf** argumentation reasoner.
Mario Alviano.

CO Tasks

DC		DS		SE		EE	
Solver	**Points**	**Solver**	**Points**	**Solver**	**Points**	**Solver**	**Points**
pyglaf	301	cegartix	350	cegartix	350	pyglaf	238
CoQuiAAS	300	heureka	345	heureka	347	argmat-dvisat	221
cegartix	299	pyglaf	344	goDIAMOND	347	argmat-mpg	198
argmat-sat	298	goDIAMOND	343	pyglaf	346	cegartix	189
goDIAMOND	295	argmat-dvisat	340	argmat-dvisat	342	argmat-sat	187
ArgSemSAT	294	argmat-sat	340	argmat-sat	342	ArgSemSAT	182
ArgTools	282	ArgTools	336	argmat-mpg	332	ConArg	182
argmat-mpg	266	argmat-mpg	330	CoQuiAAS	331	heureka	178
argmat-dvisat	248	ArgSemSAT	330	ConArg	298	CoQuiAAS	174
ConArg	231	CoQuiAAS	327	ArgTools	160	goDIAMOND	171
heureka	148	ConArg	296	ArgSemSAT	94	ArgTools	157
EqArgSolver	137	gg-sts	258	EqArgSolver	72	EqArgSolver	124
argmat-clpb	8	EqArgSolver	68	argmat-clpb	12	argmat-clpb	20
gg-sts	-892	argmat-clpb	0	gg-sts	-247	gg-sts	-295

CO Track

Solver	Points
pyglaf	1229
cegartix	1188
argmat-sat	1167
goDIAMOND	1156
argmat-dvisat	1151
CoQuiAAS	1132
argmat-mpg	1126
heureka	1018
ConArg	1007
ArgTools	935
ArgSemSAT	900
EqArgSolver	401
argmat-clpb	30
gg-sts	-1176

Congratulations to the winner:

The **pyglaf** argumentation reasoner.
Mario Alviano.

PR Tasks

DC		DS		SE		EE	
Solver	**Points**	**Solver**	**Points**	**Solver**	**Points**	**Solver**	**Points**
ArgSemSAT	304	ArgSemSAT	303	ArgSemSAT	311	ASPrMin	285
argmat-sat	303	argmat-sat	280	argmat-dvisat	299	pyglaf	273
cegartix	299	goDIAMOND	269	argmat-sat	296	argmat-dvisat	265
pyglaf	297	argmat-dvisat	263	pyglaf	293	argmat-sat	260
goDIAMOND	288	cegartix	260	goDIAMOND	284	cegartix	240
CoQuiAAS	286	pyglaf	259	cegartix	276	ArgSemSAT	228
ArgTools	282	heureka	215	ArgTools	273	heureka	179
argmat-dvisat	248	gg-sts	201	heureka	209	EqArgSolver	177
ConArg	233	ArgTools	188	argmat-mpg	203	goDIAMOND	173
argmat-mpg	228	EqArgSolver	179	ConArg	183	ConArg	168
EqArgSolver	148	ConArg	173	EqArgSolver	148	ArgTools	155
heureka	142	argmat-mpg	173	gg-sts	-172	argmat-mpg	141
gg-sts	-706	CoQuiAAS	-103	CoQuiAAS	-618	Chimaerarg	92
						CoQuiAAS	-428
						gg-sts	-430

PR Track

Solver	**Points**	Time
ArgSemSAT	1146	
argmat-sat	1139	
pyglaf	1122	
argmat-dvisat	1075	28597.16
cegartix	1075	58263.31
goDIAMOND	1014	
ArgTools	898	
ConArg	757	
heureka	745	19691.87
argmat-mpg	745	30744.76
EqArgSolver	652	
ASPrMin	285	
Chimaerarg	92	
CoQuiAAS	-863	
gg-sts	-1107	

Congratulations to the winner:

ArgSemSAT-2017.
Federico Cerutti, Mauro Vallati, Massimiliano Giacomin, and Tobia Zanetti.

ST Tasks

DC		DS		SE		EE	
Solver	**Points**	**Solver**	**Points**	**Solver**	**Points**	**Solver**	**Points**
pyglaf	307	pyglaf	287	goDIAMOND	328	pyglaf	280
argmat-sat	303	argmat-sat	271	pyglaf	309	argmat-dvisat	263
cegartix	298	cegartix	267	argmat-sat	297	argmat-sat	258
goDIAMOND	297	goDIAMOND	263	cegartix	290	argmat-mpg	257
ArgTools	286	argmat-mpg	261	argmat-mpg	280	goDIAMOND	255
argmat-mpg	275	ConArg	238	argmat-dvisat	273	ArgTools	249
ConArg	269	argmat-dvisat	234	ConArg	259	cegartix	247
argmat-dvisat	269	heureka	225	heureka	244	ConArg	236
ArgSemSAT	254	ArgSemSAT	222	ArgSemSAT	206	heureka	217
heureka	252	EqArgSolver	117	ArgTools	164	ArgSemSAT	206
EqArgSolver	117	argmat-clpb	5	EqArgSolver	162	EqArgSolver	162
argmat-clpb	8	ArgTools	-12	argmat-clpb	62	CoQuiAAS	111
CoQuiAAS	-303	CoQuiAAS	-75	CoQuiAAS	-32	argmat-clpb	60
gg-sts	-356	gg-sts	-105	gg-sts	-337	Chimaerarg	-220
						gg-sts	-395

ST Track

Solver	Points
pyglaf	1183
goDIAMOND	1143
argmat-sat	1129
cegartix	1102
argmat-mpg	1073
argmat-dvisat	1039
ConArg	1002
heureka	938
ArgSemSAT	888
ArgTools	687
EqArgSolver	558
argmat-clpb	135
Chimaerarg	-220
CoQuiAAS	-299
gg-sts	-1193

Congratulations to the winner:

The **pyglaf** argumentation reasoner.
Mario Alviano.

SST Tasks

DC		DS		SE		EE	
Solver	**Points**	**Solver**	**Points**	**Solver**	**Points**	**Solver**	**Points**
cegartix	300	argmat-sat	289	argmat-sat	316	argmat-sat	273
ArgSemSAT	287	ArgSemSAT	284	ArgSemSAT	313	pyglaf	256
argmat-sat	286	cegartix	260	pyglaf	292	cegartix	241
goDIAMOND	270	goDIAMOND	257	cegartix	290	ArgSemSAT	229
pyglaf	244	pyglaf	255	goDIAMOND	282	goDIAMOND	223
argmat-mpg	185	argmat-mpg	183	argmat-mpg	214	argmat-mpg	173
ArgTools	147	ConArg	163	ConArg	198	ConArg	163
ConArg	144	gg-sts	66	ArgTools	149	ArgTools	149
CoQuiAAS	39	ArgTools	-177	gg-sts	-343	gg-sts	-304
gg-sts	-740	CoQuiAAS	-447	CoQuiAAS	-737	CoQuiAAS	-497

SST Track

Solver	Points
argmat-sat	1164
ArgSemSAT	1113
cegartix	1091
goDIAMOND	1032
pyglaf	1047
argmat-mpg	755
ConArg	668
ArgTools	268
gg-sts	-1321
CoQuiAAS	-1642

Congratulations to the winner:

argmat-sat: Applying SAT Solvers for Argumentation Problems based on Boolean Matrix Algebra.
Fuan Pu, Guiming Luo, and Ya Hang.

STG Tasks

DC		DS		SE		EE	
Solver	**Points**	**Solver**	**Points**	**Solver**	**Points**	**Solver**	**Points**
argmat-sat	250	argmat-sat	291	argmat-sat	328	pyglaf	197
cegartix	238	pyglaf	230	cegartix	271	argmat-sat	196
pyglaf	213	cegartix	226	pyglaf	269	argmat-mpg	179
argmat-mpg	204	goDIAMOND	179	CoQuiAAS	245	cegartix	163
goDIAMOND	203	ConArg	156	argmat-mpg	235	ConArg	152
ConArg	146	gg-sts	38	goDIAMOND	217	goDIAMOND	125
ArgTools	26	ArgTools	14	ConArg	195	ArgTools	-34
CoQuiAAS	-2	argmat-mpg	0	ArgTools	61	CoQuiAAS	-312
gg-sts	-310	CoQuiAAS	-236	gg-sts	-556	gg-sts	-497

STG Track

Solver	Points
argmat-sat	1065
pyglaf	909
cegartix	898
goDIAMOND	724
ConArg	649
argmat-mpg	618
ArgTools	67
CoQuiAAS	-305
gg-sts	-1325

Congratulations to the winner:

argmat-sat: Applying SAT Solvers for Argumentation Problems based on Boolean Matrix Algebra.
Fuan Pu, Guiming Luo, and Ya Hang.

D3 Track

Solver	Points
argmat-dvisat	276
pyglaf	275
argmat-sat	271
ConArg	264
cegartix	259
EqArgSolver	192
goDIAMOND	178
argmat-mpg	164
gg-sts	-326
CoQuiAAS	-498

Congratulations to the winner:

argmat-dvisat: A division-based algorithm framework for solving argumentation problems using SAT. **Fuan Pu, Guiming Luo, and Ya Hang.**

Suggestions for future ICCMA events

Output Format

- Use linebreaks in EE-solutions (e.g. after each extension)
- Standardized exit codes

Benchmark Suite

- Maintain classification by hardness
- Maximal benchmark size?

Validation

- Reserve more time for validation
- How to handle uniquely solved instances

Thanks . . .

. . . to all participants

. . . to all who submitted new benchmarks

. . . to Matthias Thimm for presenting the current results

. . . to TAFA for hosting ICCMA

. . . to the Center for Information Services and High Performance Computing (ZIH) at TU Dresden for generous allocations of computer time.

`http://www.dbai.tuwien.ac.at/iccma17/results.html`

`iccma17@dbai.tuwien.ac.at`