Université de Paris

Results of the Fourth International Competition on Computational Models of Argumentation

Jean-Marie Lagniez[1], Emmanuel Lonca[1], Jean-Guy Mailly[2] and Julien Rossit[2]

1: CRIL, CNRS and Université d'Artois, Lens
2: LIPADE - Distributed Artificial Intelligence

ICCMA@IJCAI2021
August 24th, 2021

Université de Paris

About ICCMA

- *The competition aims at nurturing research and development of implementations for computational models of argumentation.*
  `http://argumentationcompetition.org/`
- Current steering committee: S. Gaggl (Pres.), N. Oren (Vice-Pres.), J.-G. Mailly (Secr.), F. Cerutti, M. Thimm, M. Vallati, S. Villata
- ICCMA 2015: M. Thimm and S. Villata
  - 18 solvers
- ICCMA 2017: S. Gaggl, T. Linsbichler, M. Maratea and S. Woltran
  - 16 solvers/6 benchmarks
- ICCMA 2019: S. Bistarelli, F. Santini, L. Kotthoff, T. Mantadelis and C. Taticchi
  - 9 solvers/2 benchmarks
- ICCMA 2021: J.-M. Lagniez, E. Lonca, J.-G. Mailly and J. Rossit

Université de Paris

Outline

Université de Paris

Abstract Argumentation [Dung 95]

Argumentation Framework (AF) and Extension Semantics

$F = (A, R)$ where A is a set of arguments and $R \subseteq A \times A$ represents attacks between arguments. $S \subseteq A$ is

- *conflict-free* (**cf**) if there is no $a, b \in S$ s.t. $(a, b) \in R$
- *admissible* (**ad**) if $S \in \text{cf}(F)$ and S defends all its elements
- *stable* (**stb**) if $S \in \text{cf}(F)$ and S attacks each argument in $A \setminus S$
- *complete* (**co**) if $S \in \text{ad}(F)$ and S doesn't defend any argument in $A \setminus S$
- *preferred* (**pr**) if S is $\subseteq$-maximal in $\text{ad}(F)$
- *semi-stable* (**sst**) if $S \in \text{co}(F)$ and S is range-maximal in $\text{co}(F)$
- *stage* (**stg**) if if $S \in \text{cf}(F)$ and S is range-maximal in $\text{cf}(F)$
- *ideal* (**id**) if $S \in \text{ad}(F)$ s.t. $\forall S' \in pr(F)$, $S \subseteq S'$, and S is $\subseteq$-maximal among those sets

Reasoning Tasks

- **CE**-σ: Given an AF F, how many σ-extensions has F?
- **SE**-σ: Given an AF F, provide one σ-extension of F (if it exists).
- **DS**-σ: Given an AF F and an argument a, is a in each σ-extension of F?
- **DC**-σ: Given an AF F and an argument a, is a in some σ-extension of F?

Université de Paris

Outline

Université de Paris

Tracks

- Classical track: exact algorithms
- New track: approximate algorithms
- In each track, one sub-track for each semantics
- In each sub-track, several reasoning tasks

Classical Track: Exact Algorithms

- Semantics under consideration: $\sigma \in \{\text{co}, \text{pr}, \text{stb}, \text{sst}, \text{stg}, \text{id}\}$
  - we choose to remove the grounded semantics (not challenging enough)
- Tasks: Given an AF $F = \langle A, R \rangle$
  - **CE**-σ: give the number of σ-extensions of F
  - **SE**-σ: give one σ-extension of F
  - **DC**-σ: for $a \in A$ an argument, is a credulously accepted in F?
  - **DS**-σ: $a \in A$ an argument, is a skeptically accepted in F?
- Four problems for each subtrack except $\sigma = \text{id}$ (**CE**-id $= 1$, and **DC**-id $=$ **DS**-id)

New Track: Approximate Algorithms

- Semantics under consideration: $\sigma \in \{\text{co}, \text{pr}, \text{stb}, \text{sst}, \text{stg}, \text{id}\}$
- Tasks: Given an AF $F = \langle A, R \rangle$
  - **DC**-σ: for $a \in A$ an argument, is a credulously accepted in F?
  - **DS**-σ: $a \in A$ an argument, is a skeptically accepted in F?
- Two problems for each subtrack except $\sigma = \text{id}$ (**DC**-id $=$ **DS**-id)

I/O and Environment

- Input and output from 2019 edition
  - New problem **CE**: simply print the number of extensions
- Environment:
  - Intel Xeon E5-2637 v4 CPU/128GB RAM
  - Time limit: 600s for the “exact" track, 60s for the “approximate" track
  - Memory limit: 128GB

Université de Paris

Scoring Rules

- One ranking for each sub-track
  - six rankings for the "exact" track
  - six rankings for the "approximate" track
  - To be ranked, a solver must participate to the full sub-track
  - No requirement to participate to all the (sub-)tracks
- Scoring: "exact" track
  - Any wrong result: exclusion from the sub-track
  - Correct answer in the runtime limit: 1 point
  - Timeout or non-parsable output: 0 point
  - Tie-break: cumulated runtime over the instances correctly solved
- Scoring: "approximate" track
  - Correct answer in the runtime limit: 1 point
  - Timeout, non-parsable output or wrong result: 0 point
  - Tie-break: cumulated runtime over the instances correctly solved
- $Score(Solver, Task) = \sum_{i \in Task} Score(Solver, i)$
- $Score(Solver, Subtrack) = \sum_{Task \in Subtrack} Score(Solver, Task)$

Université de Paris

Benchmark Selection

ICCMA 2019 instances

- 165 hardest instances from ICCMA 2019
- Goal: check the evolution of solvers during two years

New instances

- 422 new instances:
  - Generate a (meta-)graph G following a classical generation pattern (*e.g.* Erdos-Renyi, Barabasi-Albert,. . .)
  - For each node n_i in this graph, generate a new graph F_i
  - For each edge (n_1, n_2) in G, pick some arguments a_1 in F_1 and a_2 in F_2, and add an edge (a_1, a_2)
- Intuition: create graphs with “communities of arguments"

Query argument selection (**DS**, **DC**)

- For each AF, one argument is randomly chosen
- The same argument is used for all the **DS** and **DC** queries on the same AF

Université de Paris

Outline

Participants

Exact solvers:

- A-Folio DPDB (Fichte, Hecher, Gorczyca and Dewoprabowo)
- ASPARTIX-V21 (Dvorák, König, Wallner and Woltran)
- ConArg (Bistarelli, Rossi, Santini and Taticchi)
- FUDGE (Thimm, Cerutti, Vallati)
- MatrixX (Heinrich)
- μ-toksia (Niskanen and Järvisalo)
- PYGLAF (Alviano)

Approximate solvers:

- AFGCN (Malmqvist)
- HARPER++ (Thimm)

Université de Paris

Participants

Exact solvers:

- A-Folio DPDB (Fichte, Hecher, Gorczyca and Dewoprabowo)
- ASPARTIX-V21 (Dvorák, König, Wallner and Woltran)
- ConArg (Bistarelli, Rossi, Santini and Taticchi)
- FUDGE (Thimm, Cerutti, Vallati)
- MatrixX (Heinrich)
- μ-toksia (Niskanen and Järvisalo)
- PYGLAF (Alviano)

Approximate solvers:

- AFGCN (Malmqvist)
- HARPER++ (Thimm)

- 5 new solvers and 4 updated solvers from previous ICCMA

Université de Paris

Exact Solvers - Complete Subtrack

Rank	Solver	Score
1	A-Folio DPDB	1838
2	PYGLAF	1835
3	μ-toksia	1803
4	ASPARTIX-V21	1787
5	FUDGE	1695
6	MatrixX	759
7	ConArg	428

Université de Paris

Exact Solvers - Preferred Subtrack

Rank	Solver	Score
1	PYGLAF	1299
2	μ-toksia	1210
3	FUDGE	1190
4	ASPARTIX-V21	1052
5	ConArg	429

Exact Solvers - Semi-Stable Subtrack

Rank	Solver	Score
1	PYGLAF	1515
2	μ-toksia	1103
3	ASPARTIX-V21	744
4	ConArg	428

Université de Paris

Exact Solvers - Stable Subtrack

Rank	Solver	Score
1	A-Folio-DPDB	1862
2	PYGLAF	1743
3	FUDGE	1585
4	μ-toksia	1441
5	ASPARTIX-V21	1429
6	ConArg	429
7	MatrixX	259

Exact Solvers - Stage Subtrack

Rank	Solver	Score
1	ASPARTIX-V21	879
2	μ-toksia	788
3	ConArg	425

Université de Paris

Exact Solvers - Stage Subtrack

Rank	Solver	Score
1	ASPARTIX-V21	879
2	μ-toksia	788
3	ConArg	425

- PYGLAF was removed from this track because of incorrect results on CE-STG

Exact Solvers - Ideal Subtrack

Rank	Solver	Score
1	FUDGE	492
2	ASPARTIX-V21	306
3	PYGLAF	238
4	μ-toksia	216
5	ConArg	214

About Parallel Computing

- μ-toksia was submitted in two versions: single thread and multi-thread (four threads with different configurations of the underlying SAT solver)

Complete

Rank	Solver	Score
	μ-toksia-parallel	1866
1	A-Folio DPDB	1838
3	μ-toksia	1803

Preferred

Rank	Solver	Score
2	μ-toksia	1210
	μ-toksia-parallel	1195

Semi-Stable

Rank	Solver	Score
2	μ-toksia	1103
	μ-toksia-parallel	1008

Stable

Rank	Solver	Score
4	μ-toksia	1441
	μ-toksia-parallel	1366

Stage

Rank	Solver	Score
2	μ-toksia	788
	μ-toksia-parallel	627

Ideal

Rank	Solver	Score
2	ASPARTIX-V21	306
	μ-toksia-parallel	300
4	μ-toksia	216

- Multi-threading does not seem have a significant impact on a global level
- A more fine grained analysis of the results might provide a better insight of the question

Université de Paris

Approximate Solvers - Complete Subtrack

Rank	Solver	Score
1	HARPER++	747
2	AFGCN	668

Approximate Solvers - Preferred Subtrack

Rank	Solver	Score
1	AFGCN	567
2	HARPER++	438

Approximate Solvers - Semi-Stable Subtrack

Rank	Solver	Score
1	AFGCN	522
2	HARPER++	351

Approximate Solvers - Stage Subtrack

Rank	Solver	Score
1	AFGCN	392
2	HARPER++	349

Université de Paris

Approximate Solvers - Stable Subtrack

Rank	Solver	Score
1	AFGCN	637
2	HARPER++	457

Université de Paris

Approximate Solvers - Ideal Subtrack

Rank	Solver	Score	Cumulated Runtime
1	HARPER++	108	9.848397
2	AFGCN	108	470.655630

Some Thoughts on the Results

- Breaking open doors: no scoring system is perfect, and other measures would provide other results
- The best solver may differ, depending on applications, constraints,...
  - *E.g.*, for approximate reasoning, AFGCN wins when accuracy matters, but HARPER++ wins when time constraints must be fulfilled
- Detailed results and their analysis will be available ASAP

Outline

Université de Paris

Results Summary

Subtrack	Exact Winner	Approximate Winner
Complete	A-Folio-DPDB	HARPER++
Preferred	PYGLAF	AFGCN
Semi-Stable	PYGLAF	AFGCN
Stage	ASPARTIX-V21	AFGCN
Stable	A-Folio-DPDB	AFGCN
Ideal	FUDGE	HARPER++

Results Summary

Subtrack	Exact Winner	Approximate Winner
Complete	A-Folio-DPDB	HARPER++
Preferred	PYGLAF	AFGCN
Semi-Stable	PYGLAF	AFGCN
Stage	ASPARTIX-V21	AFGCN
Stable	A-Folio-DPDB	AFGCN
Ideal	FUDGE	HARPER++

- Exact algorithms: 3 subtracks won by updated solvers from previous ICCMA, and 3 subtracks won by new solvers

Université de Paris

Results Summary

Subtrack	Exact Winner	Approximate Winner
Complete	A-Folio-DPDB	HARPER++
Preferred	PYGLAF	AFGCN
Semi-Stable	PYGLAF	AFGCN
Stage	ASPARTIX-V21	AFGCN
Stable	A-Folio-DPDB	AFGCN
Ideal	FUDGE	HARPER++

- Exact algorithms: 3 subtracks won by updated solvers from previous ICCMA, and 3 subtracks won by new solvers
- Approximate algorithm: entirely new

Université de Paris

Conclusion

- Thanks and congratulations to all the participants
- Thanks to the ICCMA steering committee
- Thanks to the French Ministry of Research and the *Région Hauts de France* for funding the CRIL cluster through CPER DATA
- Ideas for the future:
  - Revive the dynamic argumentation track
  - Structured argumentation
  - New metrics for approximate solvers (**CE**-σ, **SE**-σ)
  - Parallel computing
- Detailed results and benchmark descriptions will be available soon at `http://argumentationcompetition.org/2021/index.html`
- See `http://argumentationcompetition.org` or `https://twitter.com/argcompetition` for information on the future of ICCMA