Results of the Fourth International Competition on Computational Models of Argumentation

Jean-Marie Lagniez\textsuperscript{1}, Emmanuel Lonca\textsuperscript{1}, Jean-Guy Mailly\textsuperscript{2} and Julien Rossit\textsuperscript{2}

\textsuperscript{1}: CRIL, CNRS and Université d’Artois, Lens
\textsuperscript{2}: LIPADE - Distributed Artificial Intelligence

ICCMA@IJCAI2021
August 24th, 2021
• 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

1 Background: Abstract Argumentation

2 Competition Rules

3 Participants and Results
   - Participants
   - Results: Exact Solvers
   - Results: Approximate Solvers

4 Conclusion
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. \( \forall S' \in \text{pr}(F), S \subseteq S' \), 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 $F$?
1 Background: Abstract Argumentation

2 Competition Rules

3 Participants and Results
   - Participants
   - Results: Exact Solvers
   - Results: Approximate Solvers

4 Conclusion
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: $\sigma \in \{\text{co, pr, stb, sst, stg, id}\}$
  • we choose to remove the grounded semantics (not challenging enough)

• Tasks: Given an AF $F = \langle A, R \rangle$
  • $\text{CE-}\sigma$: give the number of $\sigma$-extensions of $F$
  • $\text{SE-}\sigma$: give one $\sigma$-extension of $F$
  • $\text{DC-}\sigma$: for $a \in A$ an argument, is $a$ credulously accepted in $F$?
  • $\text{DS-}\sigma$: $a \in A$ an argument, is $a$ skeptically accepted in $F$?

• Four problems for each subtrack except $\sigma = \text{id}$ ($\text{CE-}\text{id} = 1$, and $\text{DC-}\text{id} = \text{DS-}\text{id}$)
New Track: Approximate Algorithms

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

\[
\text{Score}(Solver, Task) = \sum_{i \in \text{Task}} \text{Score}(Solver, i)
\]

\[
\text{Score}(Solver, Subtrack) = \sum_{\text{Task} \in \text{Subtrack}} \text{Score}(Solver, \text{Task})
\]
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 ($\textbf{DS}$, $\textbf{DC}$)

• For each AF, one argument is randomly chosen
• The same argument is used for all the $\textbf{DS}$ and $\textbf{DC}$ queries on the same AF
Outline

1 Background: Abstract Argumentation

2 Competition Rules

3 Participants and Results
   • Participants
   • Results: Exact Solvers
   • Results: Approximate Solvers

4 Conclusion
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)
- $\mu$-toksia (Niskanen and Järvisalo)
- PYGLAF (Alviano)

Approximate solvers:

- AFGCN (Malmqvist)
- HARPER++ (Thimm)
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)
- $\mu$-toksia (Niskanen and Järvisalo)
- PYGLAF (Alviano)

Approximate solvers:
- AFGCN (Malmqvist)
- HARPER++ (Thimm)

5 new solvers and 4 updated solvers from previous ICCMA
### Exact Solvers - Complete Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A-Folio DPDB</td>
<td>1838</td>
</tr>
<tr>
<td>2</td>
<td>PYGLAF</td>
<td>1835</td>
</tr>
<tr>
<td>3</td>
<td>(\mu)-toksia</td>
<td>1803</td>
</tr>
<tr>
<td>4</td>
<td>ASPARTIX-V21</td>
<td>1787</td>
</tr>
<tr>
<td>5</td>
<td>FUDGE</td>
<td>1695</td>
</tr>
<tr>
<td>6</td>
<td>MatrixX</td>
<td>759</td>
</tr>
<tr>
<td>7</td>
<td>ConArg</td>
<td>428</td>
</tr>
</tbody>
</table>
## Exact Solvers - Preferred Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PYGLAF</td>
<td>1299</td>
</tr>
<tr>
<td>2</td>
<td>$\mu$-toksia</td>
<td>1210</td>
</tr>
<tr>
<td>3</td>
<td>FUDGE</td>
<td>1190</td>
</tr>
<tr>
<td>4</td>
<td>ASPARTIX-V21</td>
<td>1052</td>
</tr>
<tr>
<td>5</td>
<td>ConArg</td>
<td>429</td>
</tr>
<tr>
<td>Rank</td>
<td>Solver</td>
<td>Score</td>
</tr>
<tr>
<td>------</td>
<td>--------------</td>
<td>-------</td>
</tr>
<tr>
<td>1</td>
<td>PYGLAF</td>
<td>1515</td>
</tr>
<tr>
<td>2</td>
<td>$\mu$-toksia</td>
<td>1103</td>
</tr>
<tr>
<td>3</td>
<td>ASPARTIX-V21</td>
<td>744</td>
</tr>
<tr>
<td>4</td>
<td>ConArg</td>
<td>428</td>
</tr>
</tbody>
</table>
### Exact Solvers - Stable Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A-Folio-DPDB</td>
<td>1862</td>
</tr>
<tr>
<td>2</td>
<td>PYGLAF</td>
<td>1743</td>
</tr>
<tr>
<td>3</td>
<td>FUDGE</td>
<td>1585</td>
</tr>
<tr>
<td>4</td>
<td>$\mu$-toksia</td>
<td>1441</td>
</tr>
<tr>
<td>5</td>
<td>ASPARTIX-V21</td>
<td>1429</td>
</tr>
<tr>
<td>6</td>
<td>ConArg</td>
<td>429</td>
</tr>
<tr>
<td>7</td>
<td>MatrixX</td>
<td>259</td>
</tr>
<tr>
<td>Rank</td>
<td>Solver</td>
<td>Score</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>-------</td>
</tr>
<tr>
<td>1</td>
<td>ASPARTIX-V21</td>
<td>879</td>
</tr>
<tr>
<td>2</td>
<td>$\mu$-toksia</td>
<td>788</td>
</tr>
<tr>
<td>3</td>
<td>ConArg</td>
<td>425</td>
</tr>
</tbody>
</table>

PYGLAF was removed from this track because of incorrect results on CE-STG.
Exact Solvers - Stage Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ASPARTIX-V21</td>
<td>879</td>
</tr>
<tr>
<td>2</td>
<td>$\mu$-toksia</td>
<td>788</td>
</tr>
<tr>
<td>3</td>
<td>ConArg</td>
<td>425</td>
</tr>
</tbody>
</table>

- PYGLAF was removed from this track because of incorrect results on CE-STG
## Exact Solvers - Ideal Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FUDGE</td>
<td>492</td>
</tr>
<tr>
<td>2</td>
<td>ASPARTIX-V21</td>
<td>306</td>
</tr>
<tr>
<td>3</td>
<td>PYGLAF</td>
<td>238</td>
</tr>
<tr>
<td>4</td>
<td>$\mu$-toksia</td>
<td>216</td>
</tr>
<tr>
<td>5</td>
<td>ConArg</td>
<td>214</td>
</tr>
</tbody>
</table>
• \(\mu\)-toksia was submitted in two versions: single thread and multi-thread (four threads with different configurations of the underlying SAT solver)

- 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
<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HARPER++</td>
<td>747</td>
</tr>
<tr>
<td>2</td>
<td>AFGCN</td>
<td>668</td>
</tr>
<tr>
<td>Rank</td>
<td>Solver</td>
<td>Score</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>1</td>
<td>AFGCN</td>
<td>567</td>
</tr>
<tr>
<td>2</td>
<td>HARPER++</td>
<td>438</td>
</tr>
<tr>
<td>Rank</td>
<td>Solver</td>
<td>Score</td>
</tr>
<tr>
<td>------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>1</td>
<td>AFGCN</td>
<td>522</td>
</tr>
<tr>
<td>2</td>
<td>HARPER++</td>
<td>351</td>
</tr>
</tbody>
</table>
Approximate Solvers - Stage Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AFGCN</td>
<td>392</td>
</tr>
<tr>
<td>2</td>
<td>HARPER++</td>
<td>349</td>
</tr>
</tbody>
</table>
### Approximate Solvers - Stable Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AFGCN</td>
<td>637</td>
</tr>
<tr>
<td>2</td>
<td>HARPER+++</td>
<td>457</td>
</tr>
</tbody>
</table>
## Approximate Solvers - Ideal Subtrack

<table>
<thead>
<tr>
<th>Rank</th>
<th>Solver</th>
<th>Score</th>
<th>Cumulated Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>HARPER++</td>
<td>108</td>
<td>9.848397</td>
</tr>
<tr>
<td>2</td>
<td>AFGCN</td>
<td>108</td>
<td>470.655630</td>
</tr>
</tbody>
</table>
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
Background: Abstract Argumentation

Competition Rules

Participants and Results
- Participants
- Results: Exact Solvers
- Results: Approximate Solvers

Conclusion
## Results Summary

<table>
<thead>
<tr>
<th>Subtrack</th>
<th>Exact Winner</th>
<th>Approximate Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>A-Folio-DPDB</td>
<td>HARPER++</td>
</tr>
<tr>
<td>Preferred</td>
<td>PYGLAF</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Semi-Stable</td>
<td>PYGLAF</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Stage</td>
<td>ASPARTIX-V21</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Stable</td>
<td>A-Folio-DPDB</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Ideal</td>
<td>FUDGE</td>
<td>HARPER++</td>
</tr>
</tbody>
</table>
Results Summary

<table>
<thead>
<tr>
<th>Subtrack</th>
<th>Exact Winner</th>
<th>Approximate Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>A-Folio-DPDB</td>
<td>HARPER++</td>
</tr>
<tr>
<td>Preferred</td>
<td>PYGLAF</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Semi-Stable</td>
<td>PYGLAF</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Stage</td>
<td>ASPARTIX-V21</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Stable</td>
<td>A-Folio-DPDB</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Ideal</td>
<td>FUDGE</td>
<td>HARPER++</td>
</tr>
</tbody>
</table>

• Exact algorithms: **3 subtracks won by updated solvers** from previous ICCMA, and **3 subtracks won by new solvers**
Results Summary

<table>
<thead>
<tr>
<th>Subtrack</th>
<th>Exact Winner</th>
<th>Approximate Winner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete</td>
<td>A-Folio-DPDB</td>
<td>HARPER++</td>
</tr>
<tr>
<td>Preferred</td>
<td>PYGLAF</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Semi-Stable</td>
<td>PYGLAF</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Stage</td>
<td>ASPARTIX-V21</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Stable</td>
<td>A-Folio-DPDB</td>
<td>AFGCN</td>
</tr>
<tr>
<td>Ideal</td>
<td>FUDGE</td>
<td>HARPER++</td>
</tr>
</tbody>
</table>

- Exact algorithms: 3 subtracks won by updated solvers from previous ICCMA, and 3 subtracks won by new solvers
- Approximate algorithm: entirely new
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 ($\text{CE-}\sigma$, $\text{SE-}\sigma$)
  - 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