HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY

UNIVERSITY OF HELSINKI
FACULTY OF SCIENCE
Department of Computer Science

Synthesizing Argumentation Frameworks from Examples

Andreas Niskanen
joint work with Johannes P. Wallner and Matti Järvisalo

Motivation: The study of **computational aspects of argumentation** is an active area of modern AI research. Recent studies on the problem of **realizability** in the context of Dung's **argumentation frameworks**.

Contributions:
- **Introduce the problem** of AF synthesis as a natural generalization of realizability
- **Complexity analysis** for multiple AF semantics
- **Algorithms** based on constraint optimization
- **Implementation** and empirical evaluation

AF SYNTHESIS: DEFINITIONS AND COMPLEXITY

ARGUMENTATION FRAMEWORKS

A directed graph $F = (A, R)$, where

- A is the set of **arguments**
- $R \subseteq A \times A$ is the **attack relation**
  - $(a, b) \in R$: a attacks b

Semantics σ define sets of jointly accepted arguments or **extensions**

- Independent sets with specific properties, e.g. self-defence

Realizability: given σ and extensions, is there an AF representing **exactly** these?

AF SYNTHESIS

Given sets of extensions and non-extensions as weighted **positive** and **negative** examples, construct closest AF representing them.

Cost of an AF F: the sum of the weights of examples **not satisfied**

Input: (A, E^+, E^-, σ), where

- A is a non-empty set of **arguments**
- E^+, E^- are sets of **examples**
- σ is an AF **semantics**

Task: Find the closest AF in terms of minimizing the cost

COMPUTATIONAL COMPLEXITY

	general	$E^+ = \emptyset$	$E^- = \emptyset$
Conflict-free	NP-c	trivial	trivial
Admissible	NP-c	trivial	trivial
Stable	NP-c	trivial	NP-c

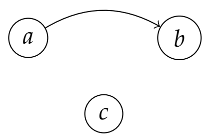
- Complete digraph satisfies all negative examples $\Rightarrow E^+ = \emptyset$ trivial
- Empty digraph satisfies all positive examples under conflict-free and admissible $\Rightarrow E^- = \emptyset$ trivial
- NP-hardness follows by a reduction from the Boolean satisfiability problem
- Note: Under stable semantics even the case $E^- = \emptyset$ is NP-complete!

Example: AF synthesis under different semantics

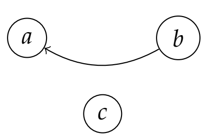
Input:

positive E^+	negative E^-
$e_1 = (\{a,b\}, 1)$	$e_4 = (\{a\}, 1)$
$e_2 = (\{a,c\}, 1)$	$e_5 = (\{a,b,c\}, 5)$
$e_3 = (\{b,c\}, 5)$	

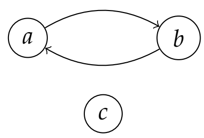
Solutions:



Conflict-free



Admissible



Stable

AF SYNTHESIS VIA MAXIMUM SATISFIABILITY

(Weighted partial) MaxSAT: a Boolean **optimization paradigm**.

Input: Hard clauses and weighted soft clauses
Task: Find a truth assignment that satisfies all hard clauses and maximizes the sum of the weights of satisfied soft clauses.

Encoding AF synthesis as MaxSAT: declare Boolean **variables**

- $r_{a,b}$ for all $a, b \in A$, true iff attack (a, b) included
- Ext^e_σ for all $e \in E^+ \cup E^-$, true iff e is a σ-extension

Hard clauses encode the problem structure: for all examples e,

$$\mathrm{Ext}^e_\sigma \leftrightarrow \varphi_\sigma(e)$$

where $\varphi_\sigma(e)$ encodes that e is a σ-extension.
Soft clauses encode the objective function:

- for all positive examples e, Ext^σ_e
- for all negative examples e, $\neg\mathrm{Ext}^\sigma_e$

Weights of soft clauses according to weights of examples.

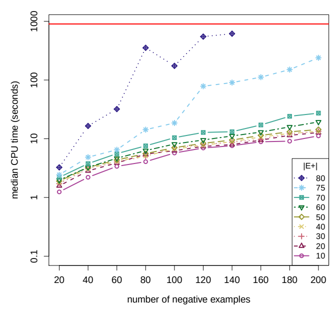
$$\varphi_{cf}(e) = \bigwedge_{a,b \in e} \neg r_{a,b}$$

$$\varphi_{adm}(e) = \varphi_{cf}(e) \wedge \bigwedge_{a \in e} \bigwedge_{b \in A \setminus e} \left(r_{b,a} \rightarrow \bigvee_{c \in e} r_{c,b} \right)$$

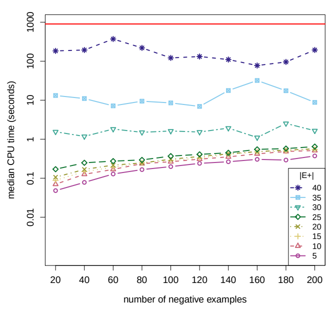
$$\varphi_{stb}(e) = \varphi_{cf}(e) \wedge \bigwedge_{a \in A \setminus e} \left(\bigvee_{b \in e} r_{b,a} \right)$$

System AFSynth and benchmarks available at:
`cs.helsinki.fi/group/coreo/afsynth`

EMPIRICAL EVALUATION



Admissible semantics.



Stable semantics.