FAKULTÄT
FÜR !NFORMATIK

Faculty of Informatics

# Investigating Subclasses of Abstract Dialectical Frameworks

Atefeh Keshavarzi

February 20, 2018

Advisor: Stefan Woltran

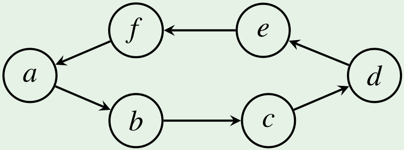
# Introduction

- Argumentation has recently become a main topic within artificial intelligence
  - crucial importance in other fields of science
  - connection to other areas of AI
- Argumentation: negotiating beliefs among agents
  - Argument presents beliefs and reasonable justifications

# Introduction

- Abstract Argumentation frameworks (AFs ) [Dung, 1995]
  - A set of arguments
  - A binary relation between arguments which represents conflicts (attacks)

## Example

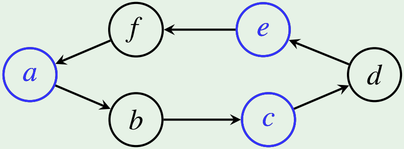



- Semantics: Methods used to clarify the acceptance of arguments
- Extension: set of jointly accepted arguments

# Introduction

- Abstract Argumentation frameworks (AFs ) [Dung, 1995]
  - A set of arguments
  - A binary relation between arguments which represents conflicts (attacks)

**Example**



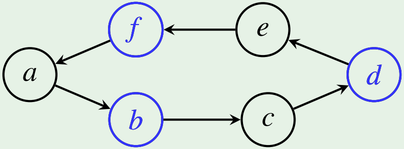

$stb(F) = \{\{a, c, e\},$

- Semantics: Methods used to clarify the acceptance of arguments
- Extension: set of jointly accepted arguments

# Introduction

- Abstract Argumentation frameworks (AFs ) [Dung, 1995]
  - A set of arguments
  - A binary relation between arguments which represents conflicts (attacks)

**Example**



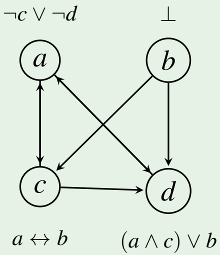

$stb(F) = \{\{a, c, e\}, \{b, d, f\}\}$

- Semantics: Methods used to clarify the acceptance of arguments
- Extension: set of jointly accepted arguments

# Introduction

- Abstract Dialectical Frameworks (ADFs) [Brewka and Woltran, 2010, Brewka et al., 2013]
  - Unify several generalizations of AFs
  - Express relations between arguments beyond simple attack [group attack, support, group support,...]
  - This is modeled via acceptance conditions

## Example



# Motivation

## Vibrant research on AFs

- Interesting subclasses of AFs are identified
- Properties of these subclasses are studied
- Coincidence of semantics for a subclass of AFs is clarified
- Expressivity of different semantics of AFs is investigated

## Question

How do these carry over to ADFs?

# Motivation

## Main Contributions

- Defining subclasses of ADFs
  - Coincidence of semantics
  - Properties
- Studying the expressiveness of different argumentation formalisms from the perspective of realizability
- Analyzing the effect of cycles on the performance of solvers for ADFs

# Outline

# Background

### Definition

An *argumentation framework* (AF) is a pair $(A, R)$ s.t.

- $A$ is a finite set of arguments
- $R \subseteq A \times A$ is a binary relation representing attacks between arguments

- $S \in \mathit{cf}(F)$ if there is no $a, b \in S$ s.t. $(a, b) \in R$
- An argument $a \in A$ is *defended* by $S \subseteq A$ (in $F$) if $\forall c \in A$: if $(c, a) \in R$ then $\exists b \in S$ s.t. $(b, c) \in R$
- $\Gamma_F(S) = \{a \in A \mid a \text{ is defended by } S \text{ in } F\}$

# Background

## Semantics of AFs

Given an AF $F = (A, R)$. A set $S \in cf(F)$ is

- $S \in adm(F)$ if $S \subseteq \Gamma_F(S)$
- $S \in pref(F)$ if $S$ is $\subseteq$-maximal admissible
- $S \in comp(F)$ if $S = \Gamma_F(S)$
- $S \in grd(F)$ if $S$ is the $\subseteq$-least fixed point of $\Gamma_F(S)$
- $S \in stb(F)$ if $\forall a \in A$: $\exists b \in S$ s.t. $(b, a) \in R$
- $S \in naive(F)$ if $S$ is $\subseteq$-maximal conflict-free

# Background

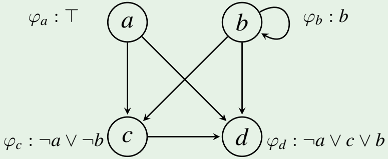
## Definition

An *abstract dialectical framework* (ADF) is a tuple $D = (S, L, C)$ where

- $S$ is a set of nodes (argument, statement, positions)
- $L \subseteq S \times S$ is a set of links
- $C = \{\varphi_s\}_{s \in S}$ is a set of propositional formulas (acceptance conditions)

## Example

$D = (\{a, b, c, d\}, \{(a, c), (a, d), (c, d), (b, b), (b, c), (b, d)\}, \{\varphi_a : \top, \varphi_b : b, \varphi_c : \neg a \vee \neg b, \varphi_d : \neg a \vee c \vee b\})$



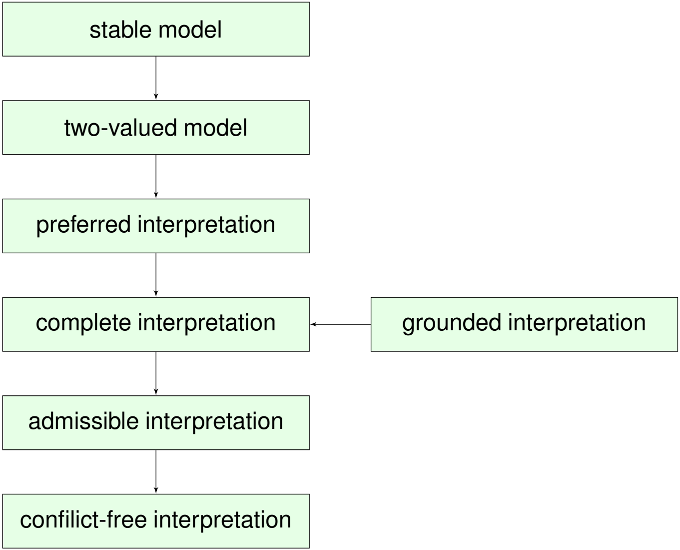
# Background

## Semantics of ADFs

Given an ADF $D$. An interpretation $v$ is

- $v \in \mathit{adm}(D)$ if $v \leq_i \Gamma_D(v)$
- $v \in \mathit{pref}(D)$ if $v$ is $\leq_i$-maximal admissible
- $v \in \mathit{comp}(D)$ if $v = \Gamma_D(v)$
- $v$ is $\mathit{grd}(D)$ if $v$ is the $\leq_i$-least fixed point of $\Gamma_D(v)$
- $v \in \mathit{mod}(D)$ if $v$ is a two-valued interpretation and $v = \Gamma_D(v)$
- $v \in \mathit{stb}(D)$ if $V$ is a model of $D$ and $v^t = w^t$, in which $w$ is the grounded interpretation of $D^t = (v^t, L \cap (v^t \times v^t), \{\varphi_s[p/\bot \ : \ v(p) = f]\}_{s \in v^t})$
- $v \in \mathit{cf}(D)$ if for each $s \in S$; $v(s) = t$ implies $\varphi_s^v$ is satisfiable and $v(s) = f$ implies $\varphi_s^v$ is unsatisfiable

# Background


stable model
two-valued model
preferred interpretation
complete interpretation
grounded interpretation
admissible interpretation
confilict-free interpretation

# Background

### Definition

Given an ADF $D = (S, L, C)$. A link $(r, s) \in L$ is

- attacking in $D$ if $\forall v \in V_2$, $v(\varphi_s) = f$ implies $v|_t^r(\varphi_s) = f$
- supporting in $D$ if $\forall v \in V_2$, $v(\varphi_s) = t$ implies $v|_t^r(\varphi_s) = t$
- dependent if it is neither attacking nor supporting
- redundant if it is both attacking and supporting

### Definition

Given an ADF $D = (S, L, C)$. Let $L^+$ be the set of all support links and $L^-$ be the set of all attack links of $L$. $D$ is named a bipolar ADF (BADF) iff $L = L^+ \cup L^-$.

# Investigating Subclasses of ADFs

**Theorem [Dung, 1995]**

Every well-founded AF has exactly one complete extension which is grounded, preferred and stable.

**Question**

Does there exist a subclass of ADFs in which different semantics collapse to the same set of interpretations?

# Investigating Subclasses of ADFs

## Theorem [Dung, 1995]

Every well-founded AF has exactly one complete extension which is grounded, preferred and stable.

## Question

Does there exist a subclass of ADFs in which different semantics collapse to the same set of interpretations?

## Definition

A finite ADF $D = (S, L, C)$ is named *acyclic* if its corresponding directed graph does not contain any directed cycle.

## Theorem

An acyclic ADF $D = (S, L, C)$ has exactly one complete interpretation which is grounded, two-valued model, preferred and stable.

# Investigating Subclasses of ADFs

**Definition**

Given an AF $F = (A, L)$.

- It is named *coherent* whenever *pref*$(F)$ = *stb*$(F)$
- It is called *relatively grounded* if *grd*$(F) = \bigcap$ *pref*$(F)$

**Theorem [Coste-Marquis et al., 2005]**

Given a symmetric AF $F$. $F$ is coherent and relatively grounded.

# Investigating Subclasses of ADFs

**Question**

Is every symmetric ADF coherent?

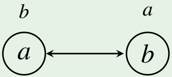
# Investigating Subclasses of ADFs

## Question

Is every symmetric ADF coherent?

## Example

Given an ADF, $D = (\{a, b\}, \{\varphi_a : b, \varphi_b : a\})$



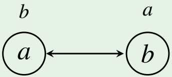
# Investigating Subclasses of ADFs

### Question

Is every symmetric ADF coherent?

### Example

Given an ADF, $D = (\{a, b\}, \{\varphi_a : b, \varphi_b : a\})$




$$\textit{pref}(D) = \{\{a \mapsto t, b \mapsto t\}, \{a \mapsto f, b \mapsto f\}\}$$

$$\textit{mod}(D) = \{\{a \mapsto t, b \mapsto t\}, \{a \mapsto f, b \mapsto f\}\}$$

$$\textit{stb}(D) = \{\{a \mapsto f, b \mapsto f\}\}$$

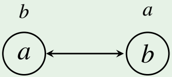
# Investigating Subclasses of ADFs

## Question

Is every symmetric ADF coherent?

## Example

Given an ADF, $D = (\{a, b\}, \{\varphi_a : b, \varphi_b : a\})$




$$pref(D) = \{\{a \mapsto t, b \mapsto t\}, \{a \mapsto f, b \mapsto f\}\}$$

$$mod(D) = \{\{a \mapsto t, b \mapsto t\}, \{a \mapsto f, b \mapsto f\}\}$$

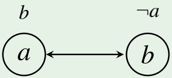
$$stb(D) = \{\{a \mapsto f, b \mapsto f\}\}$$

$$pref(D) = mod(D) \wedge mod(D) \neq stb(D)$$

# Investigating Subclasses of ADFs

## Example

Given an ADF, $D = (\{a, b\}, \{\varphi_a : b, \varphi_b : \neg a\})$




$$pref(D) = \{\{a \mapsto u, b \mapsto u\}\}$$

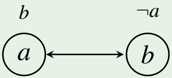
$$mod(D) = \{\}$$

$$stb(D) = \{\}$$

# Investigating Subclasses of ADFs

## Example

Given an ADF, $D = (\{a, b\}, \{\varphi_a : b, \varphi_b : \neg a\})$




$$pref(D) = \{\{a \mapsto u, b \mapsto u\}\}$$

$$mod(D) = \{\}$$

$$stb(D) = \{\}$$

$$pref(D) \neq mod(D) \wedge mod(D) = stb(D)$$

# Investigating Subclasses of ADFs

**Definition**

Given an ADF $D = (S, L, C)$.

- It is called semi-coherent if $\mathit{pref}(D) = \mathit{mod}(D)$
- It is called weak-coherent if $\mathit{mod}(D) = \mathit{stb}(D)$

# Investigating Subclasses of ADFs

**Definition**

Given an ADF $D = (S, L, C)$.

- It is called semi-coherent if *pref*$(D)$ = *mod*$(D)$
- It is called weak-coherent if *mod*$(D)$ = *stb*$(D)$

**Theorem**

Symmetric ADFs are neither weak-coherent nor semi-coherent.

# Investigating Subclasses of ADFs

### Definition

Given a symmetric ADF $D = (S, L, C)$.

- Attack symmetric ADF (ASADF): all links are attacking ($L = L^+$)
- Acyclic support symmetric ADF (ASSADF): $D$ is a symmetric BADF and does not contain any support cycle

# Investigating Subclasses of ADFs

## Definition

Given a symmetric ADF $D = (S, L, C)$.

- Attack symmetric ADF (ASADF): all links are attacking ($L = L^+$)
- Acyclic support symmetric ADF (ASSADF): $D$ is a symmetric BADF and does not contain any support cycle

## Theorem

ASADFs and ASSADFs are weak-coherent, but not semi-coherent.

# Investigating Subclasses of ADFs

| | Acyclic ADFs | Symmetric ADFs | ASADFs | ASSADFs | CADFs |
|---|---|---|---|---|---|
| *cf* $=$ *adm* | – | – | – | – | – |
| *adm* $=$ *comp* | – | – | – | – | $\surd$ |
| *pref* $=$ *mod* (semi-coherent) | $\surd$ | – | – | – | – |
| *mod* $=$ *stb* (weak-coherent) | $\surd$ | – | $\surd$ | $\surd$ | – |
| *comp* $=$ *pref* | $\surd$ | – | – | – | – |
| *stb* $= \{$*grd*$\}$ | $\surd$ | – | – | – | – |
| relatively grounded | $\surd$ | – | – | – | – |
| coherent | $\surd$ | – | – | – | – |

# Realizability and Expressivity

**Question**

Which set of sets can be the outcome of the evaluation of a formalism under semantics $\sigma$?

# Realizability and Expressivity

## Question

Which set of sets can be the outcome of the evaluation of a formalism under semantics $\sigma$?

## Example

Given $\mathbb{S} = \{\{a, c, e\}, \{b, d, f\}\}$.

- $\exists F \in AFs$ s.t. $\mathbb{S} = stb(F)$?

# Realizability and Expressivity

### Question

Which set of sets can be the outcome of the evaluation of a formalism under semantics $\sigma$?

### Example

Given $\mathbb{S} = \{\{a, c, e\}, \{b, d, f\}\}$.

- $\exists F \in AFs$ s.t. $\mathbb{S} =$ *stb*$(F)$? Yes

# Realizability and Expressivity

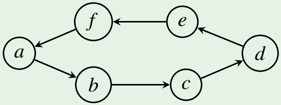
## Question

Which set of sets can be the outcome of the evaluation of a formalism under semantics $\sigma$?

## Example

Given $\mathbb{S} = \{\{a, c, e\}, \{b, d, f\}\}$.

- $\exists F \in AFs$ s.t. $\mathbb{S} =$ *stb*$(F)$? Yes



# Realizability and Expressivity

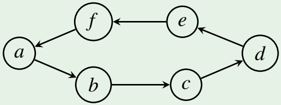
## Question

Which set of sets can be the outcome of the evaluation of a formalism under semantics $\sigma$?

## Example

Given $\mathbb{S} = \{\{a, c, e\}, \{b, d, f\}\}$.

- $\exists F \in AFs$ s.t. $\mathbb{S} = stb(F)$? Yes




- $\exists F \in AFs$ s.t. $\mathbb{S} = adm(F)$?

# Realizability and Expressivity

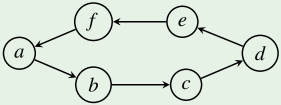
### Question

Which set of sets can be the outcome of the evaluation of a formalism under semantics $\sigma$?

### Example

Given $\mathbb{S} = \{\{a, c, e\}, \{b, d, f\}\}$.

- $\exists F \in AFs$ s.t. $\mathbb{S} =$ *stb*$(F)$? Yes




- $\exists F \in AFs$ s.t. $\mathbb{S} =$ *adm*$(F)$? No

# Realizability and Expressivity

### Definition

The *signature* $\Sigma^{\sigma}_{\mathcal{F}}$ of a formalism $\mathcal{F}$ w.r.t. semantics $\sigma$ is defined as:

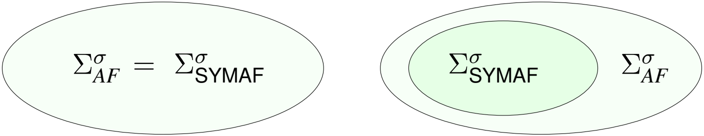
$$\Sigma^{\sigma}_{\mathcal{F}} = \{\sigma(kb) \mid kb \in \mathcal{F}\}$$

### Question

What is the expressive power of semantics of different formalisms via realizability?

# Expressivity of SYMAFs and AFs

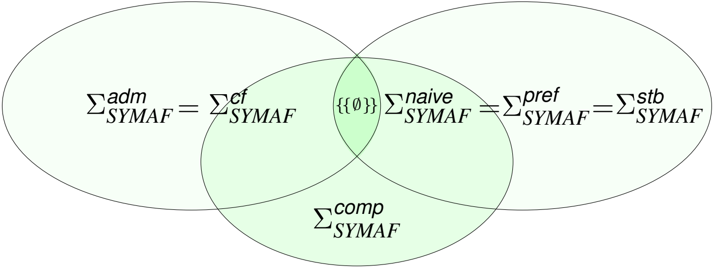
- Relations between signatures of SYMAFs and AFs for $\sigma \in \{adm, cf, naive, pref, stb, comp, grd\}$:




for $\sigma \in \{cf, naive, grd\}$ — for $\sigma \in \{adm, pref, stb, comp\}$

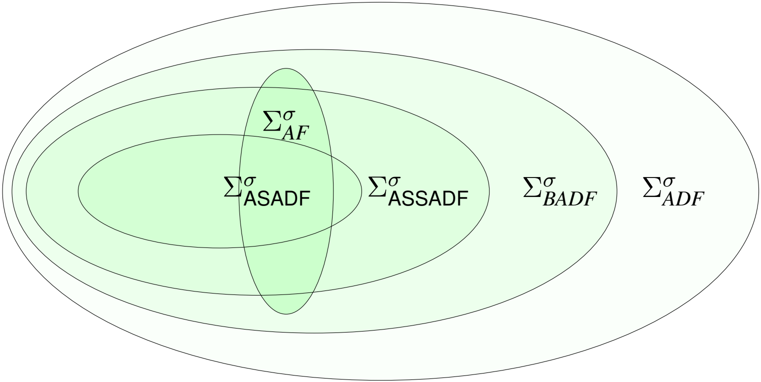
# Expressivity of Symmetric AFs

- Relations between signatures of SYMAFs for $\sigma \in \{adm, cf, naive, pref, comp, stb\}$:



# Expressivity of Subclasses of ADFs

- Expressivity of subclasses of ADFs for $\sigma \in \{adm, pref, comp\}$:




- $\Sigma^{stb}_{AF} \subsetneq \Sigma^{stb}_{\text{ASADF}} \subseteq^{stb}_{\text{ASSADF}} \subseteq \Sigma^{stb}_{BADF} = \Sigma^{stb}_{ADF}$

# Experiments on Subclasses of ADFs

- Generator by which acyclic ADFs, ASADFs and ASSADFs are produced
- The effect of cycles on the performance of solvers for ADFs
  - Credulous acceptance
  - Skeptical acceptance
- Solvers: DIAMOND, goDIAMOND, YADF and QADF

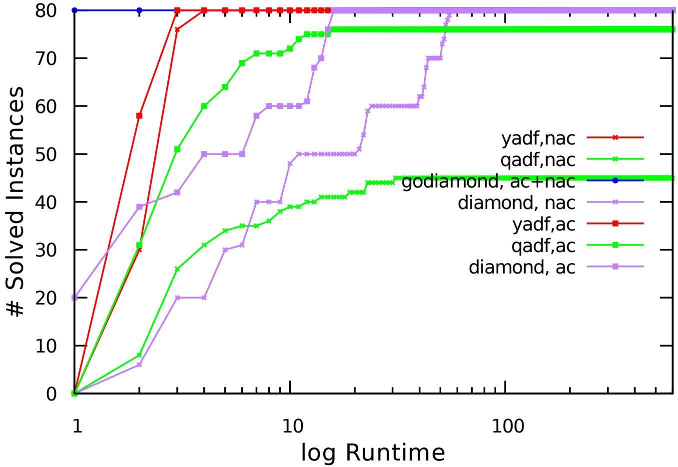
# Experiments on Subclasses of ADFs

- Credulous acceptance under admissible semantics

| | | Time-outs | Mean |
|---|---|---|---|
| *adm*-acyclic | DIAMOND | 0 | 5.2813 |
| | goDIAMOND | 0 | 0.0667 |
| | YADF | 0 | 1.5727 |
| | QADF | 4 | 3.1635 |
| *adm*-non-acyclic | DIAMOND | 0 | 17.641 |
| | goDIAMOND | 0 | 0.202 |
| | YADF | 0 | 2.1679 |
| | QADF | 35 | 5.5060 |

# Experiments on Subclasses of ADFs

- Number of acyclic and non-acyclic ADF-credulous-acceptance problems solved by solvers:



# Experiments on Subclasses of ADFs

- Skeptical acceptance under preferred semantics:

| | | Time-outs | Mean |
|---|---|---|---|
| *pref*-acyclic | DIAMOND | 80 | – |
| | goDIAMOND | 0 | 0.144 |
| | YADF | 57 | 90.696 |
| | QADF | 80 | – |
| *pref*-non-acyclic | DIAMOND | 80 | – |
| | goDIAMOND | 8 | 1.2838 |
| | YADF | 40 | 126.128 |
| | QADF | 80 | – |

# Summary

- Reformulate and prove Dung's Fundamental lemma for ADFs
- Specify a subclass of ADFs in which different semantics can collapse to the same set of interpretations
- Clarify whether symmetric ADFs are coherent and relatively grounded
- Study properties of acyclic ADFs, symmetric ADFs, ASADFs, ASSADFs and CADFs
- Expressiveness of different argumentation formalisms from the perspective of realizability
- Provide a generator to produce subclasses of ADFs
- Use this generator to analyze the effect of cycles on the performance of solvers for ADFs

# Future Work

- Further ADF semantics are introduced by [Polberg, 2016]. How these semantics behave in the subclasses studied in this thesis?
- Studying the computational complexity of the subclasses of ADFs
- Answering open questions in regard to expressiveness

$$\Sigma^{stb}_{\text{ASADF}} \subsetneq^{?} \Sigma^{stb}_{\text{ASSADF}} \subsetneq^{?} \Sigma^{stb}_{BADF}$$

- The results of our experiments suggest that solvers for ADFs can/should be optimized

Thank you for your attention!

# References I


- Brewka, G., Strass, H., Ellmauthaler, S., Wallner, J. P., and Woltran, S. (2013).
  Abstract dialectical frameworks revisited.
  In Rossi, F., editor, IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013, pages 803–809. IJCAI/AAAI.
- Brewka, G. and Woltran, S. (2010).
  Abstract dialectical frameworks.
  In Lin, F., Sattler, U., and Truszczynski, M., editors, Principles of Knowledge Representation and Reasoning: Proceedings of the Twelfth International Conference, KR 2010, Toronto, Ontario, Canada, May 9-13, 2010. AAAI Press.
- Coste-Marquis, S., Devred, C., and Marquis, P. (2005).
  Symmetric argumentation frameworks.
  In Godo, L., editor, Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 8th European Conference, ECSQARU 2005, Barcelona, Spain, July 6-8, 2005, Proceedings, volume 3571 of Lecture Notes in Computer Science, pages 317–328. Springer.

# References II


Dung, P. M. (1995).
On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games.
Artif. Intell., 77(2):321–358.

Polberg, S. (2016).
Understanding the abstract dialectical framework.
In Michael, L. and Kakas, A. C., editors, Logics in Artificial Intelligence - 15th European Conference, JELIA 2016, Larnaca, Cyprus, November 9-11, 2016, Proceedings, volume 10021 of Lecture Notes in Computer Science, pages 430–446.

Polberg, S. and Doder, D. (2014).
Probabilistic abstract dialectical frameworks.
In Fermé, E. and Leite, J., editors, Logics in Artificial Intelligence - 14th European Conference, JELIA 2014, Funchal, Madeira, Portugal, September 24-26, 2014. Proceedings, volume 8761 of Lecture Notes in Computer Science, pages 591–599. Springer.

# References III

Strass, H. and Wallner, J. P. (2015).
Analyzing the computational complexity of abstract dialectical frameworks via approximation fixpoint theory.
Artif. Intell., 226:34–74.