Doing Argumentation Theory in Modal Logic

Davide Grossi

ILLC, University of Amsterdam

"Model-Theoretic Foundations of Argumentation Networks"

Davide Grossi

ILLC, University of Amsterdam

d.grossi@uva.nl   Institute of Logic, Language and Computation

UNIVERSITEIT VAN AMSTERDAM

Aim

- Study modal languages that *talk about* argumentation frameworks (argumentation frameworks as structures for logical semantics)
- Why?
- Import techniques (e.g., calculi, logical games) and results (e.g., axiomatizations, complexity)
- ... *for free!*

Outline

- **PART I**: Dung Frameworks = Kripke Frames
- **PART II**: Dung Frameworks + Labellings = Kripke Models
- **PART III**: Argumentation in Modal Logic
  - Axiomatizations, completeness, complexity
- **PART IV**: Dialogue Games via Semantic Games
  - Model-checking games
- **PART V**: *"When are two arguments the same?"*
  - Bisimulation, bisimulation games

Part I

Dung Frameworks = *Kripke* Frames

… just a relational structure (i)

$$\mathcal{A} = (A, \to)$$

- Arguments = States (or points, possible worlds, etc.)
- Attack = Accessibility relation

... just a relational structure (ii)

$$\mathcal{A}, a \models \langle \rightarrow \rangle \top \iff \exists b \in A, a \rightarrow b$$

$$\mathcal{A}, a \models \langle \leftarrow \rangle \top \iff \exists b \in A, a \rightarrow^{-1} b$$

- “there exists an argument b attacked by (or defeated by) a”
- “there exists an argument b attacking (or defeating) a”

Part II

Dung Fr. + Labellings = Kripke Models

d.grossi@uva.nl Institute of Logic, Language and Computation

Universiteit van Amsterdam

... just a labelled relational structure (i)

$$\mathcal{M} = (\mathcal{A}, \mathcal{I})$$

- ☐ Arguments = States (or points, possible worlds, etc.)
- ☐ Attack = Accessibility relation
- ☐ Valuation = function from a vocabulary **P** to sets of arguments

... just a labelled relational structure (ii)

Definition 1 (Argumentation models) *Let* $\mathbf{P}$ *be a set of propositional atoms. An argumentation model* $\mathcal{M} = (\mathcal{A}, \mathcal{I})$ *is a structure such that:*

- $\mathcal{A} = (A, \rightarrow)$ *is an argumentation framework;*
- $\mathcal{I} : \mathbf{P} \longrightarrow 2^A$ *is an assignment from* $\mathbf{P}$ *to subsets of* A*.*

The set of all argumentation models is called $\mathfrak{A}$*. A pointed argumentation model is a pair* $(\mathcal{M}, a)$ *where* $\mathcal{M}$ *is an argumentation model and* a *an argument.*

- ☐ Arguments = States (or points, possible worlds, etc.)
- ☐ Attack = Accessibility relation
- ☐ Valuation = function from a vocabulary **P** to sets of arguments

... just a labelled relational structure (iii)

Example 1 *(Argument labelings as argumentation models)* If argumentation frameworks can be studied as Kripke frames, then an argumentation framework together with a labelling function [Caminada, 2006] from the set $\{\mathtt{1}, \mathtt{0}, ?\}$ is nothing but a Kripke model on the alphabet $\{\mathtt{1}, \mathtt{0}, ?\}$:

- $\mathcal{A} = (A, \rightarrow)$ is an argumentation framework;
- $\mathcal{I}$ is a valuation function from the set of atoms $\mathbf{P} = \{\mathtt{1}, \mathtt{0}, ?\}$ to the set 2^A;
- $\mathcal{M} \models \mathtt{Fct}$, where $\mathtt{Fct} := (\mathtt{1} \wedge \neg\mathtt{0} \wedge \neg?) \vee (\neg\mathtt{1} \wedge \mathtt{0} \wedge \neg?) \vee (\neg\mathtt{1} \wedge \neg\mathtt{0} \wedge ?)$

A logic for "local" argumentation (i)

$$\mathcal{L}^{\mathsf{K}^{-1}} : \varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\rightarrow\rangle\varphi \mid \langle\leftarrow\rangle\varphi$$

$$(\mathcal{A}, \mathcal{I}), a \models \langle\rightarrow\rangle\varphi \iff \exists b \in A, a \rightarrow b \;\&\; (\mathcal{A}, \mathcal{I}), b \in ||\varphi||$$

$$(\mathcal{A}, \mathcal{I}), a \models \langle\leftarrow\rangle\varphi \iff \exists b \in A, a \rightarrow^{-1} b \;\&\; (\mathcal{A}, \mathcal{I}), b \in ||\varphi||$$

- ☐ "existence of *attackers* with a specific label"
- ☐ "existence of *attacked arguments* with a specific label"

A logic for "local" argumentation (ii)

$(\texttt{Prop})$	propositional schemata
$(\texttt{K})$	$[i](\varphi_1 \longrightarrow \varphi_2) \longrightarrow ([i]\varphi_1 \longrightarrow [i]\varphi_2)$
$(\texttt{Conv})$	$\varphi \longrightarrow [i]\neg[j]\neg\varphi$
$(\texttt{Dual})$	$\langle i \rangle \leftrightarrow \neg[i]\neg\varphi$
$(\texttt{MP})$	IF $\vdash \varphi_1 \longrightarrow \varphi_2$ AND $\vdash \varphi_1$ THEN φ_2
$(\texttt{N})$	IF $\vdash \varphi$ THEN $\vdash [i]\varphi$

with $i \neq j \in \{\rightarrow, \leftarrow\}$.

- This axiomatics is sound and strongly complete w.r.t. the class of all argumentation frameworks

Argumentation notions as global validities (i)

$$
\begin{array}{rcl}
\mathit{Acceptable}(\varphi, \psi, \mathcal{M}) & \Longleftrightarrow & \mathcal{M} \models \varphi \rightarrow [\leftarrow]\langle\leftarrow\rangle\psi \\
\mathit{SelfAcceptable}(\varphi, \mathcal{M}) & \Longleftrightarrow & \mathcal{M} \models \varphi \rightarrow [\leftarrow]\langle\leftarrow\rangle\varphi \\
\mathit{CFree}(\varphi, \mathcal{M}) & \Longleftrightarrow & \mathcal{M} \models \varphi \rightarrow \neg\langle\rightarrow\rangle\varphi \\
\mathit{Adm}(\varphi, \mathcal{M}) & \Longleftrightarrow & \mathcal{M} \models \varphi \rightarrow ([\rightarrow]\neg\varphi \wedge [\leftarrow]\langle\leftarrow\rangle\varphi) \\
\mathit{Complete}(\varphi, \mathcal{M}) & \Longleftrightarrow & \mathcal{M} \models (\varphi \rightarrow [\rightarrow]\neg\varphi) \wedge (\varphi \leftrightarrow [\leftarrow]\langle\leftarrow\rangle\varphi) \\
\mathit{Stable}(\varphi, \mathcal{M}) & \Longleftrightarrow & \mathcal{M} \models \varphi \leftrightarrow \neg\langle\leftarrow\rangle\varphi
\end{array}
$$

- These are all meta-language expressions!

d.grossi@uva.nl Institute of Logic, Language and Computation

UNIVERSITEIT VAN AMSTERDAM

Argumentation notions as global validities (ii)

Fact 1 (Equivalence of $\rightarrow$ and $\leftarrow$ for conflict-freeness) *Let $\mathcal{M}$ be an argumentation model. It holds that:*

$$\mathcal{M} \models \varphi \rightarrow \neg\langle\rightarrow\rangle\varphi \iff \mathcal{M} \models \varphi \rightarrow \neg\langle\leftarrow\rangle\varphi$$

- "Ask not what you cannot attack, but what cannot attack you!"
- We can restrict our logic to the the logic **K** interpreted on converse of the attack relation!

Argumentation notions as global validities (iii)

$$
\begin{aligned}
Acceptable(\varphi, \psi, \mathcal{M}) &\iff \mathcal{M} \models \varphi \rightarrow [\leftarrow]\langle\leftarrow\rangle\psi \\
SelfAcceptable(\varphi, \mathcal{M}) &\iff \mathcal{M} \models \varphi \rightarrow [\leftarrow]\langle\leftarrow\rangle\varphi \\
CFree(\varphi, \mathcal{M}) &\iff \mathcal{M} \models \varphi \rightarrow \neg\langle\leftarrow\rangle\varphi \\
Adm(\varphi, \mathcal{M}) &\iff \mathcal{M} \models \varphi \rightarrow ([\leftarrow]\neg\varphi \wedge [\leftarrow]\langle\leftarrow\rangle\varphi) \\
Complete(\varphi, \mathcal{M}) &\iff \mathcal{M} \models (\varphi \rightarrow [\leftarrow]\neg\varphi) \wedge (\varphi \leftrightarrow [\leftarrow]\langle\leftarrow\rangle\varphi) \\
Stable(\varphi, \mathcal{M}) &\iff \mathcal{M} \models \varphi \leftrightarrow \neg\langle\leftarrow\rangle\varphi
\end{aligned}
$$

- These are all meta-language expressions!

d.grossi@uva.nl Institute of Logic, Language and Computation

Universiteit van Amsterdam

Part III

Argumentation in Modal Disguise

d.grossi@uva.nl Institute of Logic, Language and Computation

Universiteit van Amsterdam

K + Global modality (i)

$$\mathcal{L}^{\mathsf{K}^{\mathsf{U}}} : \varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\leftarrow\rangle\varphi \mid \langle\mathsf{U}\rangle\varphi$$

Definition 2 (Satisfaction for $\mathcal{L}^{\mathsf{K}^{\mathsf{U}}}$ in argumentation models) *Let $\varphi \in \mathcal{L}^{\mathsf{K}^{\mathsf{U}}}$. The satisfaction of φ by a pointed argumentation model $(\mathcal{M}, a)$ is inductively defined as follows (Boolean clauses are omitted):*

$$\begin{array}{lll} \mathcal{M}, a \models \langle\leftarrow\rangle\varphi & \textit{iff} & \exists b \in A : (a,b) \in \rightarrow^{-1} \text{ AND } \mathcal{M}, b \models \varphi \\ \mathcal{M}, a \models \langle\mathsf{U}\rangle\varphi & \textit{iff} & \exists b \in A : \mathcal{M}, b \models \varphi \end{array}$$

- The global modality allows to access arguments that are not related via the attack relation (cf. *relevance*)

K + Global modality (ii)

The logic K^{U} is axiomatized as follows:

(Prop)	propositional tautologies
(K)	$[i](\varphi_1 \rightarrow \varphi_2) \rightarrow ([i]\varphi_1 \rightarrow [i]\varphi_2)$
(T)	$[\mathsf{U}]\varphi \rightarrow \varphi$
(4)	$[\mathsf{U}]\varphi \rightarrow [\mathsf{U}][\mathsf{U}]\varphi$
(5)	$\neg[\mathsf{U}]\varphi \rightarrow [\mathsf{U}]\neg[\mathsf{U}]\varphi$
(Incl)	$[\mathsf{U}]\varphi \rightarrow [i]\varphi$
(Dual)	$\langle i\rangle\varphi \leftrightarrow \neg[i]\neg\varphi$

with $i \in \{\leftarrow, \mathsf{U}\}$.

- This axiomatics is sound and strongly complete w.r.t. the class of argumentation frameworks under the given semantics

K + Global modality (iii)

We list the following known results, which are relevant for our purposes.

- The complexity of deciding whether a formula of $\mathcal{L}^{\mathsf{K}^{\mathsf{U}}}$ is satisfiable is EXP-complete [Hemaspaandra, 1996].
- The complexity of checking whether a formula of $\mathcal{L}^{\mathsf{K}^{\mathsf{U}}}$ is satisfied by a pointed model $\mathcal{M}$ is P-complete [Graedel and Otto, 1999].

- If we can express extensions as modal formulae in this logic we can import these results for free to argumentation theory.

Doing argumentation in Modal Logic (i)

$$\begin{array}{rcl} Acc(\varphi,\psi) & := & [\mathsf{U}](\varphi \rightarrow [\leftarrow]\langle\leftarrow\rangle\psi) \\ CFree(\varphi) & := & [\mathsf{U}](\varphi \rightarrow \neg\langle\leftarrow\rangle\varphi) \\ Adm(\varphi) & := & [\mathsf{U}](\varphi \rightarrow ([\leftarrow]\neg\varphi \wedge [\leftarrow]\langle\leftarrow\rangle\varphi)) \\ Complete(\varphi) & := & [\mathsf{U}]((\varphi \rightarrow [\leftarrow]\neg\varphi) \wedge (\varphi \leftrightarrow [\leftarrow]\langle\leftarrow\rangle\varphi)) \\ Stable(\varphi) & := & [\mathsf{U}](\varphi \leftrightarrow \neg\langle\leftarrow\rangle\varphi) \end{array}$$

- Now we can express the meta-language formulation of the argumentation notions in the object-language!

d.grossi@uva.nl Institute of Logic, Language and Computation
UNIVERSITEIT VAN AMSTERDAM

Doing argumentation in Modal Logic (ii)

Theorem 1 (Fundamental Lemma) *The following formula is a theorem of* K^{U}*:*

$$Adm(\varphi) \wedge Acc(\psi \vee \xi, \varphi) \rightarrow Adm(\varphi \vee \psi) \wedge Acc(\xi, \varphi \vee \psi)$$

- We can state theorems of argumentation as formulae!

Doing argumentation in Modal Logic (iii)

1.	$((\alpha \to \gamma) \wedge (\beta \to \gamma)) \to (\alpha \vee \beta \to \gamma)$	**Prop**
2.	$([\mathsf{U}](\alpha \to \gamma) \wedge [\mathsf{U}](\beta \to \gamma)) \to [\mathsf{U}](\alpha \vee \beta \to \gamma)$	$2, \mathbf{N}, \mathbf{K}, \mathbf{MP}$
3.	$([\mathsf{U}](\varphi \to [\leftarrow]\langle\leftarrow\rangle\varphi) \wedge [\mathsf{U}](\psi \to [\leftarrow]\langle\leftarrow\rangle\varphi)) \to$	
	$[\mathsf{U}](\varphi \vee \psi \to [\leftarrow]\langle\leftarrow\rangle\varphi)$	Instance of 3
4.	$[\leftarrow]\langle\leftarrow\rangle\varphi \to [\leftarrow]\langle\leftarrow\rangle(\varphi \vee \psi)$	$\mathbf{Prop}, \mathbf{K}, \mathbf{N}$
5.	$([\mathsf{U}](\varphi \to [\leftarrow]\langle\leftarrow\rangle\varphi) \wedge [\mathsf{U}](\psi \to [\leftarrow]\langle\leftarrow\rangle\varphi)) \to$	
	$[\mathsf{U}](\varphi \vee \psi \to [\leftarrow]\langle\leftarrow\rangle\varphi \vee \psi)$	$4, \mathbf{Prop}, \mathbf{K}, \mathbf{N}$
6.	$Acc(\varphi, \varphi) \wedge Acc(\psi, \varphi) \to Acc(\varphi \vee \psi, \varphi \vee \psi)$	5, definition

- **And prove them via formal derivations!**

Doing argumentation in Modal Logic (iv)

An argumentation labeling $\mathcal{M} = (\mathcal{A}, \mathcal{I})$ is a *complete labeling* if and only if for each $a \in A$:

1. $\mathcal{M}, a \models \mathtt{1}$ if and only if for all b s.t. $a \leftarrow b$, $\mathcal{M}, b \models \mathtt{0}$;
2. $\mathcal{M}, a \models \mathtt{0}$ if and only if there exists b s.t. $a \leftarrow b$ and $\mathcal{M}, b \models \mathtt{1}$
3. $\mathcal{M} \models \mathtt{Fct}$.

Fact 2 *Let $\mathcal{M} = (\mathcal{A}, \mathcal{I})$ be an argumentation model for the set of atoms $\mathbf{P} = \{\mathtt{1}, \mathtt{0}, ?\}$. It holds that:*

$$Complete(\mathcal{M}) \iff \mathcal{M} \models Complete(\mathtt{1}) \wedge [\mathsf{U}]\mathtt{Fct} \wedge [\mathsf{U}](? \leftrightarrow (\langle\leftarrow\rangle\neg\mathtt{0} \wedge \neg\langle\leftarrow\rangle\mathtt{1}))$$

- We can characterize Caminada Labelings by modal formulae and prove similar results for the other types of extensions

What about Grounded Extension?

- This is known to be the least fixpoint of the characteristic function of an argumentation framework
- The characteristic function corresponds, in modal terms, to the operator:

$$[\leftarrow]\langle\leftarrow\rangle$$

- So the grounded extension of an argumentation framework is just the smallest proposition p for which the following formula is globally true in the model:

$$p \leftrightarrow [\leftarrow]\langle\leftarrow\rangle p$$

mu-calculus (i)

$$\mathcal{L}^{\mathsf{K}^{\mu}} : \varphi ::= p \mid \bot \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\leftarrow\rangle\varphi \mid \mu p.\varphi(p)$$

Definition 3 (Satisfaction for $\mathcal{L}^{\mathsf{K}^{\mu}}$ in argumentation models) *Let $\varphi \in \mathcal{L}^{\mathsf{K}^{\mu}}$. The satisfaction of φ by a pointed argumentation model $(\mathcal{M}, a)$ is inductively defined as follows:*

$$\mathcal{M}, a \models \mu p.\varphi(p) \quad \textit{iff} \quad a \in \bigcap\{X \in 2^{A} \mid ||\varphi||_{\mathcal{M}[p:=X]} \subseteq X\}$$

- The mu operator allows us to express the least fixpoint of a formula viewed as set-transformer
- So the grounded extension of an argumentation framework is denoted by the formula:

$$\mu p.[\leftarrow]\langle\leftarrow\rangle p$$

mu-calculus (ii)

$(\texttt{Prop})$	propositional schemata
$(\texttt{K})$	$[\leftarrow](\varphi_1 \rightarrow \varphi_2) \rightarrow ([\leftarrow]\varphi_1 \rightarrow [\leftarrow]\varphi_2)$
$(\texttt{Fixpoint})$	$\varphi(\mu p.\varphi(p)) \leftrightarrow \mu p.\varphi(p)$
$(\texttt{MP})$	IF $\vdash \varphi_1 \rightarrow \varphi_2$ AND $\vdash \varphi_1$ THEN φ_2
$(\texttt{N})$	IF $\vdash \varphi$ THEN $\vdash [\leftarrow]\varphi$
$(\texttt{Least})$	IF $\vdash \varphi_1(\varphi_2) \rightarrow \varphi_2$ THEN $\vdash \mu p.\varphi_1(p) \rightarrow \varphi_2$

- This axiomatics is sound and complete for argumentation frameworks [Walukiewicz, 2000]

mu-calculus (iii)

We list some relevant known results.

- The satisfiability problem of K^{μ} is decidable [Streett, 1989].
- The complexity of the model-checking problem for K^{μ} is known to be in NP $\cap$ co-NP [Graedel, 1999], however, it is still an open question whether it is in P.
- The complexity of the model-checking problem for a formula of size m and alternation depth d on a system of size n is $O(m \cdot n^{d+1})$ [Emerson, 1986].

☐ **We can tractably model-check grounded extensions!**

d.grossi@uva.nl Institute of Logic, Language and Computation

Universiteit van Amsterdam

What about Preferred Extensions?

- They are the maximal, conflict-free, post-fixpoints of the characteristic function of an argumentation framework
- This is not expressible in the mu-calculus and it is a MSO formula (which I refrain from writing)
- Reasoning about PE goes beyond modal logic
- This might give a hint on why PE are typically hard to handle algorithmically

Part IV

Dialogue games = Evaluation Games

Dialogue games

- The proof-theory of argumentation is commonly given in terms of dialogue games
- The semantics of modal logic offers a unified framework for systematizing games that check the membership of arguments to admissible sets, complete, grounded and stable extensions

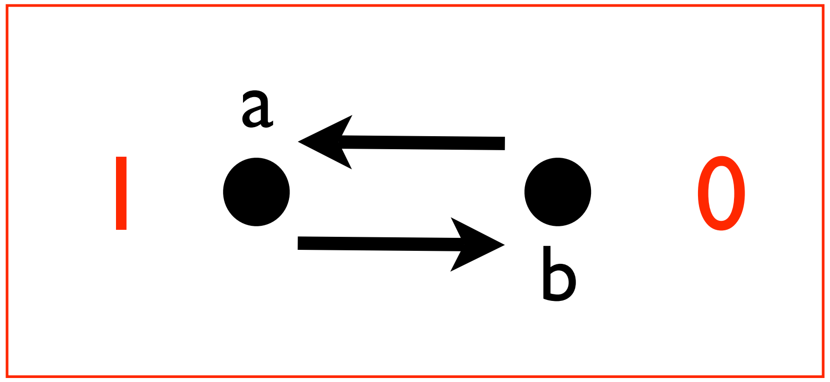
Evaluation games (i)

- Eve (the proponent) tries to prove that an argument belongs to a given set which enjoys a specific property in the argumentation model
- Adam (the opponent) tries to falsify Eve's claim
- Positions consists of pairs "(*formula, argument*)"
- Who plays depends on the formula in the position
- A player wins iff its adversary runs out of available moves

Evaluation games (ii)

Position	Turn	Available moves
$(\varphi_1 \vee \varphi_2, a)$	$\exists$	$\{(\varphi_1, a), (\varphi_2, a)\}$
$(\varphi_1 \wedge \varphi_2, a)$	$\forall$	$\{(\varphi_1, a), (\varphi_2, a)\}$
$(\langle\leftarrow\rangle\varphi, a)$	$\exists$	$\{(\varphi, b) \mid (a, b) \in \rightarrow^{-1}\}$
$([\leftarrow]\varphi, a)$	$\forall$	$\{(\varphi, b) \mid (a, b) \in \rightarrow^{-1}\}$
$(\langle \mathsf{U}\rangle\varphi, a)$	$\exists$	$\{(\varphi, b) \mid b \in A\}$
$([\mathsf{U}]\varphi, a)$	$\forall$	$\{(\varphi, b) \mid b \in A\}$
$(\bot, a)$	$\exists$	$\emptyset$
$(\top, a)$	$\forall$	$\emptyset$
(p, a) & $a \notin \mathcal{I}(p)$	$\exists$	$\emptyset$
(p, a) & $a \in \mathcal{I}(p)$	$\forall$	$\emptyset$
$(\neg p, a)$ & $a \in \mathcal{I}(p)$	$\exists$	$\emptyset$
$(\neg p, a)$ & $a \notin \mathcal{I}(p)$	$\forall$	$\emptyset$

Evaluation games (iii)



$(1 \wedge [\mathsf{U}](1 \leftrightarrow \neg\langle\leftarrow\rangle 1), a)$ $\forall$

$(1, a)$

$\exists$ve wins

$([\mathsf{U}](1 \leftrightarrow \neg\langle\leftarrow\rangle 1), a)$ $\forall$

$(1 \leftrightarrow \neg\langle\leftarrow\rangle 1, a)$ $\forall$

$(1 \leftrightarrow \neg\langle\leftarrow\rangle 1, b)$ $\forall$

$(\neg 1 \vee \neg\langle\leftarrow\rangle 1, a)$ $\exists$

$(1 \vee \langle\leftarrow\rangle 1, a)$ $\exists$

$(\neg\langle\leftarrow\rangle 1, a)$ $\forall$

$(\neg 1, a)$

$(1, b)$

$\forall$dam wins

$\exists$ve wins

Evaluation games (iv)

Theorem 2 (Adequacy of the evaluation game for K^{U}) *Let* $\varphi \in \mathcal{L}^{\mathsf{K}^{\mathsf{U}}}$*, and let* $\mathcal{M} = (\mathcal{A}, \mathcal{I})$ *be an argumentation model. Then, for any argument* $a \in A$*, it holds that:*

$$(\varphi, a) \in Win_{\exists}(\mathcal{E}(\varphi, \mathcal{M})) \Longleftrightarrow \mathcal{M}, a \models \varphi.$$

- So, in the previous game, Adam could not possibly force Eve to loose!

Evaluation games for argumentation

$$\begin{aligned}
Adm: &\quad \mathcal{E}(\varphi \wedge [\mathsf{U}](\varphi \rightarrow ([\rightarrow]\neg\varphi \wedge [\leftarrow]\langle\leftarrow\rangle\varphi)), \mathcal{M})@(\varphi \wedge [\mathsf{U}](\varphi \rightarrow ([\rightarrow]\neg\varphi \wedge [\leftarrow]\langle\leftarrow\rangle\varphi), a) \\
Complete: &\quad \mathcal{E}(\varphi \wedge [\mathsf{U}](\varphi \leftrightarrow [\leftarrow]\langle\leftarrow\rangle\varphi)), \mathcal{M})@(\varphi \wedge [\mathsf{U}](\varphi \leftrightarrow [\leftarrow]\langle\leftarrow\rangle\varphi), a) \\
Stable: &\quad \boxed{\mathcal{E}(\varphi \wedge [\mathsf{U}](\varphi \leftrightarrow \neg\langle\leftarrow\rangle\varphi)), \mathcal{M})@(\varphi \wedge [\mathsf{U}](\varphi \leftrightarrow \neg\langle\leftarrow\rangle\varphi), a)} \\
Grounded: &\quad \mathcal{E}(\mu p.[\leftarrow]\langle\leftarrow\rangle p, \mathcal{M})@(\mu p.[\leftarrow]\langle\leftarrow\rangle p, a)
\end{aligned}$$

- Evaluation games provide a comprehensive framework for a game-theoretical "proof-theory" of argumentation!

Evaluation games vs. Dialogue games

Given: $(\mathcal{A}, \mathcal{I}), a, \varphi$ $\quad$ $(\mathcal{A}, \mathcal{I}), a \models \varphi?$

Given: $\mathcal{A}, a, \varphi$ $\quad$ $\exists \mathcal{I} : (\mathcal{A}, \mathcal{I}), a \models \varphi?$

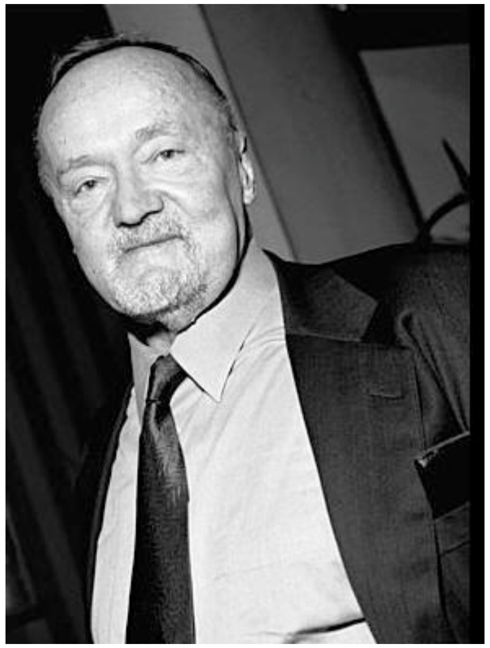
$$\mathcal{A} \models \forall p_1 \ldots p_n ST_a(\neg\varphi(p_1 \ldots p_n))?$$

- Evaluation games are algorithms for modal model-checking
- Dialogue games as defined in argumentation theory seems to be inherently more complex!

Part V

When are two arguments the *same*?

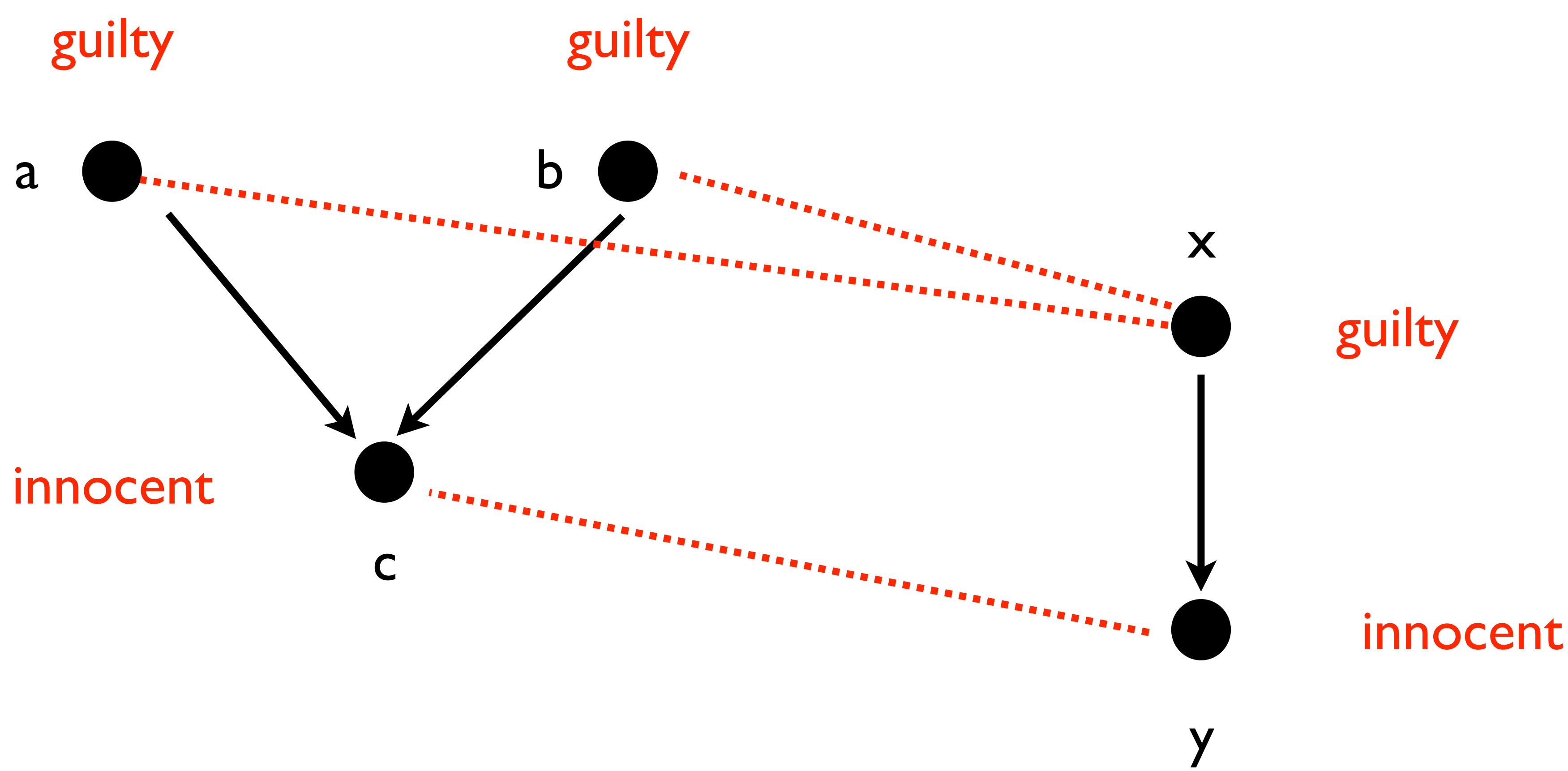
The Uses of Argument (1958)



“What things about the form and merits of our arguments are *field-invariant* and what things about them are field-dependent? [...] The *force* of the conclusion [...] is the same regardless of fields: the criteria or sorts of grounds required to justify such a conclusion vary from field to field” [Toulmin,1958]

- In argumentation theory the “invariance” theme has not yet been address
- When are two arguments/frameworks “the same”?
- E.g. principle of *Stare Decisis* in common-law

"Sameness" = "Behavioral equivalence"



- Are c and y different from the point of view of abstract argumentation? What about the rest?
- E.g., if "guilty" denotes the grounded extension on the left, so should it on the right

Bisimulation (i)

Definition 3 (Bisimulation) *Let* $\mathcal{M} = (A, \rightarrow, \mathcal{I})$ *and* $\mathcal{M}' = (A', \rightarrow', \mathcal{I}')$ *be two argumentation models. A bisimulation between* $\mathcal{M}$ *and* $\mathcal{M}'$ *is a non-empty relation* $Z \subseteq A \times A'$ *such that for any* aZa'*:*

Atom: *a and a' are propositionally equivalent;*

Zig: *if $a \leftarrow b$ for some $b \in A$, then $a' \leftarrow b'$ for some $b' \in A'$ and bZb';*

Zag: *if $a' \leftarrow b'$ for some $b' \in A$ then $a \leftarrow b$ for some $\underline{i}nA$ and aZa'.*

A total bisimulation *is a bisimulation* $Z \subseteq A \times A'$ *such that its left projection covers A and its right projection covers A'.*

- Two arguments are the same iff they are labelled in the same way, they are attacked by arguments with same labels (*bisimulation*) and this holds for all arguments in the framework (*total bisimulation*)

Bisimulation (ii)

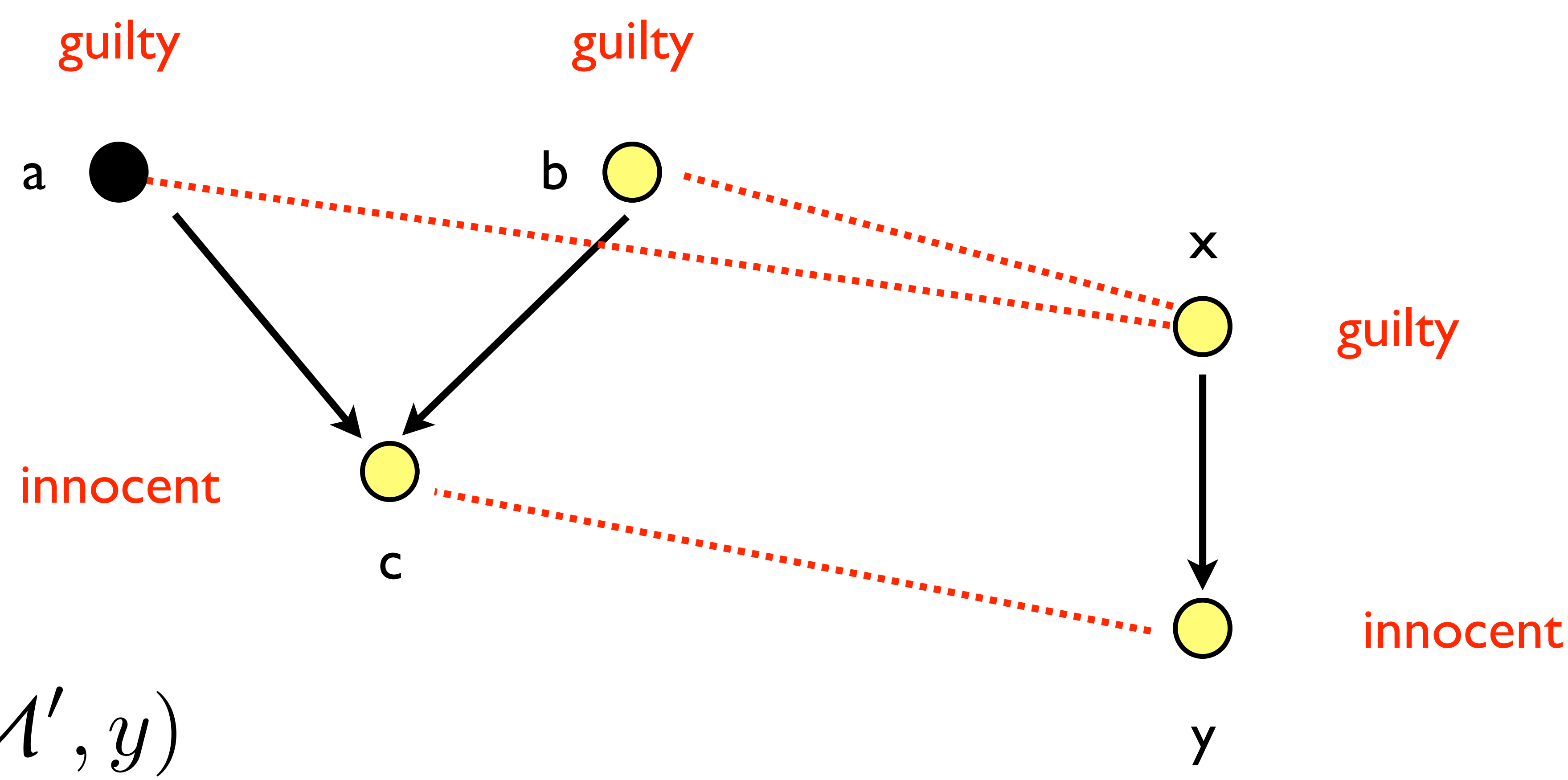
Theorem 3 (Bisimilar arguments) *Let $(\mathcal{M}, a)$ and $(\mathcal{M}', a')$ be two pointed argumentation models, and let Z be a total bisimulation between $\mathcal{M}$ and $\mathcal{M}'$. It holds that a belongs to an admissible set (complete extension, stable extension, grounded extension) if and only if a' belongs to an admissible set (complete extension, stable extension, grounded extension).*

- Follows directly from the fact that the logics expressing those concepts are invariant under (total) bisimulation

Bisimulation games (i)

- The game is played by a **S**poiler who tries to show that two given pointed models are not bisimilar, and a **D**uplicator who tries to show the contrary
- A position in a game consists of a pair (*pointed model*, *pointed model*)
- **S**poiler starts, **D**uplicator responds
- **S**poiler wins iff a position is reached where the two pointed models do not satisfy the same labels, or when **D**uplicator is out of moves

"Sameness" = "Behavioral equivalence"



$$(\mathcal{M}, c)(\mathcal{M}', y)$$

$$(\mathcal{M}, a)(\mathcal{M}', y) \qquad (\mathcal{M}, b)(\mathcal{M}', y) \qquad (\mathcal{M}, c)(\mathcal{M}', x)$$

$$(\mathcal{M}, b)(\mathcal{M}', x)$$

Duplicator wins!

d.grossi@uva.nl Institute of Logic, Language and Computation

UNIVERSITEIT VAN AMSTERDAM

Bisimulation games (ii)

Theorem 4 (Adequacy of bisimulation games) *Let $(\mathcal{M}, a)$ and $(\mathcal{M}', a')$ be two argumentation models. Duplicator has a winning strategy in the (total) bisimulation game $\mathcal{B}(\mathcal{M}, \mathcal{M}')@(a, a')$ if and only if $\mathcal{M}, a$ and $\mathcal{M}', a'$ are (totally) bisimilar.*

- Bisimulation games are an adequate "proof procedure" for checking whether two labelled argumentation frameworks behave in the same way from the point of view of argumentation theory

Part VI

Conclusions

A logical landscape for argumentation theory

$$\begin{tikzcd}
K^{-1} \arrow[r] \arrow[d] & K^{\mu} \arrow[d] \\
K^{U} \arrow[r] & MSO
\end{tikzcd}$$

K^{-1}:

- complete, stable
- meta-language

K^{μ}:

- grounded
- calculus,
- evaluation games,
- P complexity
- bisimulation games

K^{U}:

- complete, stable
- calculus,
- evaluation games,
- P complexity,
- bisimulation games

MSO:

- preferred
- model-checking ?
- games?

Related work

- G. Boella et. al. (2005) *"A Logic of Abstract Argumentation"*
  - Non-standard logic (no axiomatics, no complexity, no games)
  - Dung's notions are treated as primitives
- D. Gabbay (draft) *"Modal Provability Foundations for Argumentation Networks"*
  - Arguments as propositions
  - Modal provability logic (on finite trees)
  - Formulae encode argumentation frameworks (finiteness of attackers assumed)

Future work

- Apply results and techniques of MSO to study preferred and semi-stable
- Study the dynamics of argumentation in Dynamic Logic
- Study the robustness of the membership of an argument to a given extension in Sabotage Logic (van Benthem, 2005)
- Study *accrual* in Graded Modal Logic (de Rijke, 2000)

d.grossi@uva.nl Institute of Logic, Language and Computation

Universiteit van Amsterdam