The ASPIC+ Framework for Structured Argumentation: Motivation, Theory and Developments

Sanjay Modgil

Department of Informatics, King's College London

sanjay.modgil@kcl.ac.uk

Outline of Tutorial

Part 1

- **Motivation**: Argumentative Characterisations of Non-monotonic Reasoning – The Dialogical Turn in Logic

  - From Non-monotonic Logics to Reasoning via Dialogue
  - Scaffolding human-human and human-AI reasoning

- **The ASPIC+ Framework** for Argumentative Characterisations of Non-monotonic Reasoning

Outline of Tutorial

Part 2

- **The ASPIC+ Framework and Rationality**

  - Guidelines for satisfying Rationality Postulates
  - *Dialectical* ASPIC+ : **Full** Rationality under **Resource Bounds**

Argumentative Characterisations of Non-monotonic Reasoning – The Dialogical Turn

Logic and Dialectic

- Logic initially (from the Greeks through to medieval period) conceived of as an inherently dialectical/ dialogical enterprise in which agents engage in questioning and argument

  ➔ more solipsistic emphasis on individual agents reasoning using logic (due in large part to Avicenna, Descartes, Kant …) to reason about mental attitudes (beliefs, desires …)

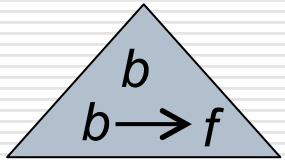
- Lorenzen and Lorenz, Keith Stenning, Johan van Benthem, Catarina D. Novaes … reviving dialectical/dynamic accounts of (typically deductive monotonic) logics

- A deductive proof of α from β_1 …. β_n characterised as a game in which PRO aims to "persuade" OPP that α follows deductively from β_1 …. β_n

Non-monotonic Logic and Argumentation

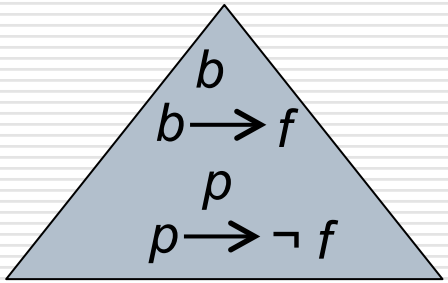
- But normative logical characterisations of rational reasoning of particular value when when formalising epistemic reasoning and decision making, i.e., when reasoning ***non-monotonically***

  Conclusions withdrawn when new information conflicts with previous conclusions or the assumptions made in drawing previous conclusions.



$\mid\!\sim f$

Non-monotonic Logic and Argumentation



$\not\vdash\!\sim f$

$\vdash\!\sim \neg f$

- Essentially, *nm* reasoning concerned with arbitrating amongst conflicts; it is this insight that is substantiated by *argumentative characterisations of non-monotonic inference …*

Dung's Argumentation Theory [1]

- Given a set Δ of (ordered) wff in some language $\mathcal{L}$:

1) Construct arguments ($\mathcal{A}rgs$) from Δ

Eg (Δ,≤) = ordered set of classical wff and X = $(\Gamma,\alpha) \in \mathcal{A}rgs$ iff Γ is a *consistent minimal* subset of Δ that classically entails α

2) Define conflict based defeat relation ($\mathcal{D}ef$) amongst $\mathcal{A}rgs$

X = (Γ,α) *defeats* Y = (Π,β) if $\neg\alpha \in \Pi$ and X $\not\prec$ $(\{\neg\alpha\},\neg\alpha)$ where $\prec$ lifts ≤ to sets of formulae

3) Evaluate justified (winning) arguments in directed graph ($\mathcal{A}rgs$, $\mathcal{D}ef$)

1. P. M. Dung. **On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games**. *Artificial Intelligence*, 77(2):321–358, 1995

Logic Programming Example

Δ =

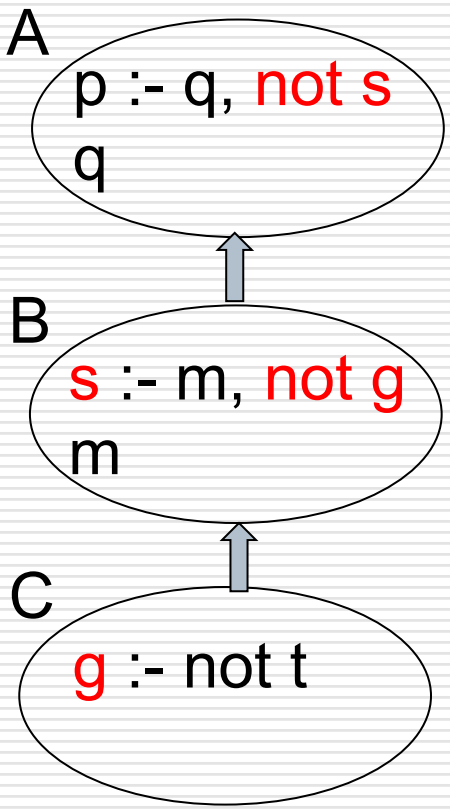
p :- q, not s

q

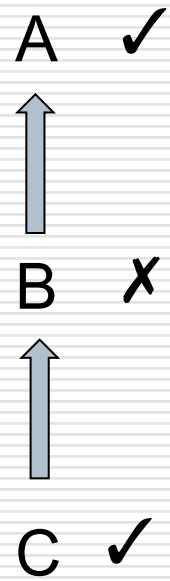
s :- m, not g

m

g :- not t



$(\mathcal{A}rgs, \mathcal{D}ef)$



C defends A

Argumentation-based characterisations of non-monotonic consequence relations

$(\mathcal{A}rgs, \mathcal{D}ef)_{\Delta} \mid\!\sim \alpha$ (the claim of a credulously/sceptically justified argument)

$\Updownarrow$

$\Delta \mid\!\sim_{nml} \alpha$

E.g. *nml* = *Logic Programming*, *Default Logic*, *Prioritised Default Logic*, *Defeasible Logic*, *Preferred Subtheories,…*

Logic and Dialectic

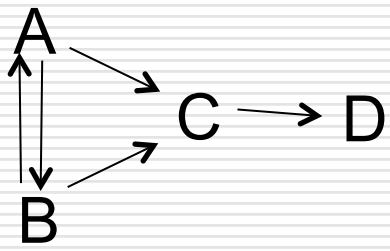
- But normative logical characterisations of rational reasoning of particular value when arbitrating amongst decision options and contentious/conflicting beliefs i.e., when reasoning non-monotonically ....

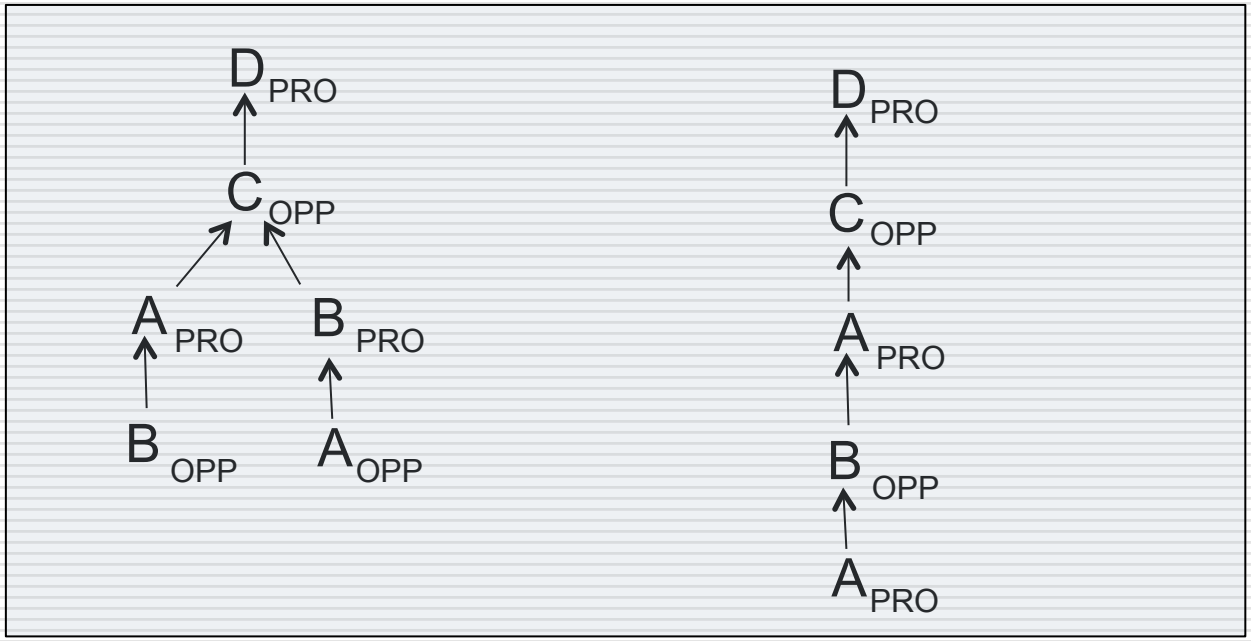
- .... where reasoning is explicit and external (on 'public' view) rather than internal and confined to **black box** *human* **and** *machine learning* minds

- And where reasoning leverages inputs from multiple 'minds' – complex reasoning is first and foremost a communicative/dialogical activity (as also evidenced by evolutionary accounts of how System 2 reasoning evolved)

- Argumentative characterisations of non-monotonic logics can be generalised to obtain models of distributed non-monotonic reasoning in the form of dialogues that take place in public view

Generalisation proceeds via Argument Game Proof Theories for Deciding Membership of Extensions *





In *grounded* game (PRO loses)

In *preferred* game PRO wins and is said to have a *winning strategy*

* S. Modgil and M. Caminada. *Proof Theories and Algorithms for Abstract Argumentation Frameworks*. In: *Argumentation in AI,* I. Rahwan and G. Simari (eds), 105-132, 2009.

Argument Game Proof theories for Non-monotonic Logics

- Argument game proof theories for non-monotonic logics

$$\Delta \mid\!\sim_{nml} \alpha \iff (\mathcal{A}rgs, \mathcal{D}ef)_{\Delta} \mid\!\sim \alpha \iff \exists x \in \mathcal{A}rgs, \text{ claim}(x) = \alpha$$

and there is a winning strategy for *x*

Argument Game Proof theories for Non-monotonic Logics

- Argument game proof theories for non-monotonic logics

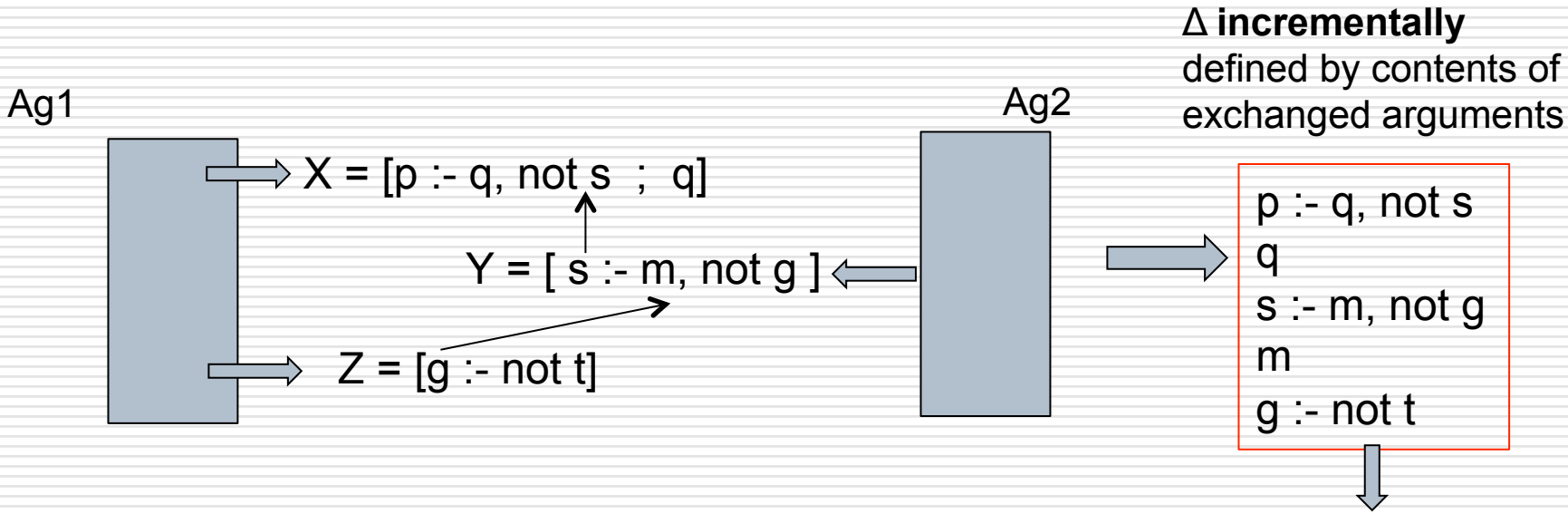
$$\Delta \mid\!\sim_{nml} \alpha \iff (\mathcal{A}rgs, \mathcal{D}ef)_{\Delta} \mid\!\sim \alpha \iff \exists x \in \mathcal{A}rgs, \text{ claim}(x) = \alpha$$

and there is a winning strategy for *x*

- Suppose we don't assume a *given Δ*: reasoning is a dynamic activity in which agent(s) assert statements and arguments whose contents *incrementally* define Δ (beliefs incrementally reveal themselves both in monological agent reasoning and dialogical multi-agent reasoning)

From single agent reasoning to distributed (non-monotonic) reasoning via dialogue

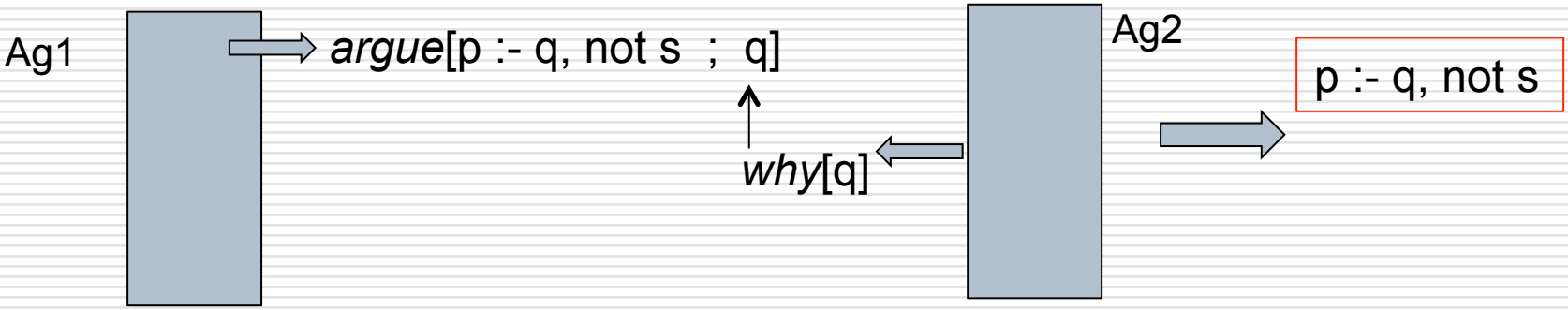
"The lonesome thinker in an armchair is as marginal as he looks: most of our logical skills are displayed in interaction" – J. Van Bentham



Ag1 wins dialogue ***iff*** $(\mathcal{A}rgs, \mathcal{D}ef)_{\Delta} \mid\sim p$

Distributed non-monotonic reasoning via dialogue

Locutions *attack* or *surrender* to other locutions *



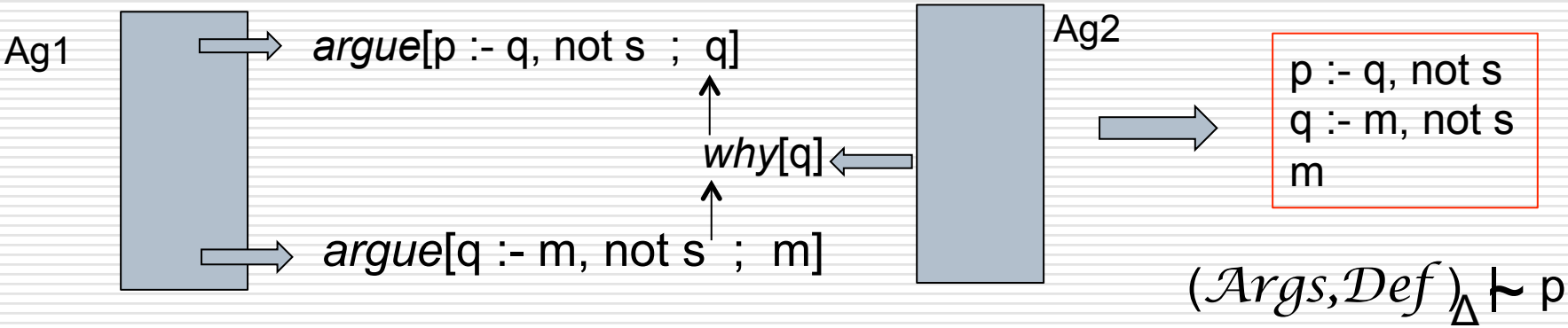
Ag1 losing dialogue $(\mathcal{A}rgs, \mathcal{D}ef)_{\Delta} \mid\!\sim p$

why(q) **attacks** *argue*[p :- q, not s ; q]

- H. Prakken. **Coherence and flexibility in dialogue games for argumentation**. In: *Journal of Logic and Computation, 15(6)*, 1009-1040, 2005.

Distributed non-monotonic reasoning via dialogue

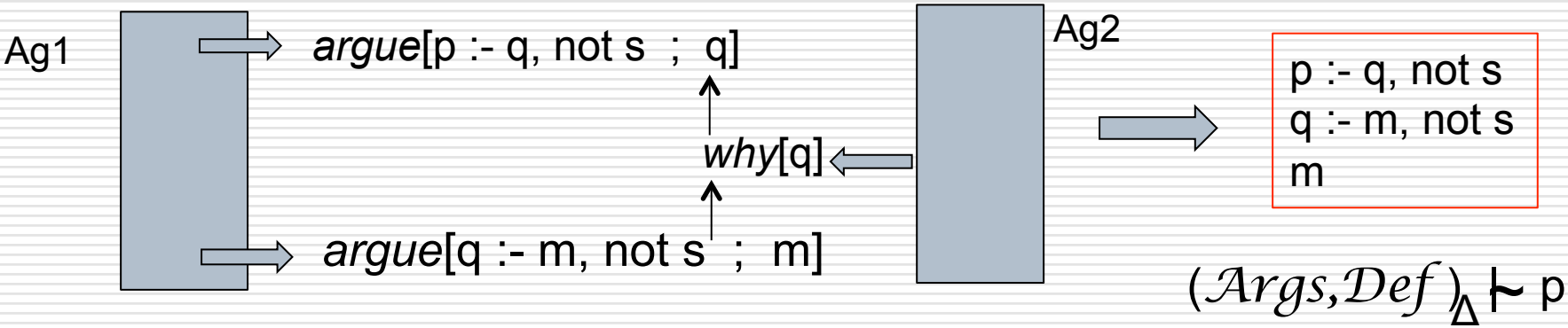
Winning strategy for *A* in **dialogue** under semantics *s* ⟺ *A* justified under *s* semantics in AF defined by assertional contents of **locutions** moved



Ag1 winning dialogue *argue*[q :- m, not s ; m] **attacks** *why*(q)

Distributed non-monotonic reasoning via dialogue

Winning strategy for *A* in **dialogue** under semantics *s* ⟺ *A* justified under *s* semantics in AF defined by assertional contents of **locutions** moved



Ag1 winning dialogue *argue*[q :- m, not s ; m] **attacks** *why*(q)

Other locution types – *concede*, *retract*, *prefer*, "*but*" (indirect illocutionary force) …

Dialectical Semantics for Non-monotonic logics

Statement true (decision option most preferred) to the extent that all attempts to thus far prove otherwise have failed

In practice, engaging with 'all attempts to prove otherwise' is an activity that involves *many agents/sources of information*

- A more pragmatic *Popper-ian* epistemology *cf* appealing to some inaccessible objective standard of truth (distinct model theoretic structure_

A belief is adjudged to have the status of knowledge (true and justified) not in view of some distinct model theoretic/metaphysical realm that validates its truth conditions, but in view of having successfully refuted challenges based on other beliefs at hand

Proof theory = dialogue ⟺ Semantics = Extensions of AF defined by contents of locutions

Two Application Areas *

Dialogical ‘Scaffolding’ for human and AI reasoning

- Dialogical support for enhancing the *quality* of human reasoning (pedagogical applications)
- Dialogical support for enhancing the *scope* of AI reasoning (addressing the value loading/alignment problem)

* S, Modgil *Many Kinds of Minds are Better than One: Value Alignment Through Dialogue .* In: Workshop on Argumentation and Philosophy (co-located with COMMA'18), 2018.

* S, Modgil. *Dialogical Scaffolding for Human and Artificial Agent Reasoning.* In: 5th International Workshop on AI and Cognition (AIC 2017), 73-86, November 2017.

The Argumentative theory of Reasoning

- System 2 (logical) reasoning evolved for communication (social brain hypothesis) when recipient needs to exercise *epistemic vigilance* *

  Recipient disposed to evaluate arguments and seeks counter-arguments to avoid being manipulated/exploited *implies*
  Sender preferentially disposed to construct arguments supporting communicated claims

- ➔ *Lone reasoner* seeks reasons in support of, and overlooks reasons contrary to, beliefs (*confirmation bias is a normal feature of reasoning*)
  ➔ *Lone decision makers* harness reasoning in anticipation of communicating decisions: evidence that we favour easily justified decision options that are less subject to criticism, rather than satisfying rationality criteria

- Social media filtering algorithms are digital ***amplified*** incarnations of these evolutionary dispositions ➔ entrenched ideological positions as in groupthink

* H. Mercier and. Sperber. *The Enigma of Reason: A New Theory of Human Understanding* Harvard Press, 2017.

Many minds are better than one

- Argumentative theory also implies (empirically supported claim) that reasoning serves us better when performed in group dialogue (*assuming intention to get to the truth/make right decision*) – division of cognitive labour

- Assumption does not apply to social media context *but does* in educational settings

- **Vision**: Socratic dialogue engines in school and university teaching

  E.g. *E-Clinic* – dialogue engine that plays role of consultant on ward round

  ➔ enhanced instilling of medial reasoning skills.

The AI *Value Loading/Alignment* Problem *

Two (related) problems for Artificial General Intelligence

1. *Specifying*
   - *rule based axiomatisations of deontic reasoning in symbolic AI*
   - *utility functions in ML*

   perfectly aligned with human values/preferences in open and changing environments

2. *Unintended consequences misaligned with human values*
   - recalling concerns about symbolic rule-based encodings of ethical theories e.g. Asimov's three laws, a sadist is a masochist who follows the golden rule

Problem 2 acquires renewed urgency given that a **feature** of machine learning is finding unforseen ways of achieving goals, and that ***any*** final goal incentivises pursuit of *instrumental goals* (including AI-**self preservation and maintain final goal**)

* *Superintelligence: Paths, Dangers, Strategies.* Nick Bostrom (head of *Future of Humanity Institute*, Oxford University)

Applications: Scaffolding human-AI reasoning

How to align AI and human values ?

- *Future of Humanity Institute (Oxford)*, *Centre for the Study of Existential Risk, Open AI, MIT ….* all working on value loading/alignment problem

- State of the art = *cooperative inverse reinforcement learning* – AI learns reward function of human through observation, querying, and more generally, ***dialogical interaction*** amongst humans and AI ('enculturation')

Value Deliberation Dialogues

- **However**:

  1) Assumption that best source of values is human behaviour ?

  2) Humans often struggle when deciding on challenging moral issues

1 and 2 exacerbated when moral challenges without precedent (saliently exposing Humean *is-ought gap*) e.g. when involving use of new technologies !

Many ***Kinds*** of Minds are Better than One: Human–AI Deliberation Dialogues for Moral Reasoning

Part of the solution is to do what humans do when faced with challenging moral problems

- "Value Deliberation" dialogues involving humans and AI **that are better purposed to decide ethical issues** given vastly superior epistemic and causal reasoning capacities of AI, informed by human considerations of values/preferences/utilities

Requires further research into logic based dialectical formalisms for moral/ethical reasoning

- Moral/ethical prescriptions are not given and unchanging but are themselves the outcome of epistemic and causal reasoning and *evaluative* (values/preferences/utilities) reasoning about consequences of actions

Other Relevant References

- H. Mercier and D. Sperber. *Why do humans reason? arguments for an argumentative theory*. Behavioral and Brain Sciences, 34(2):57–747, 2011.
- H. Mercier and D. Sperber. *The Enigma of Reason: A New Theory of Human Understanding*, 2017
- Catarina D. Novaes. *The enduring enigma of reason*". Mind & Language 2018.
- *The Reasoner. Column* - What's hot in Argumentation. S. Modgil (issues July 2017 – April 2018) (http://www.thereasoner.org)

The ASPIC+ Framework for Argumentative Characterisations of Non-monotonic Reasoning

The ASPIC+ Framework (Key references)

- ASPIC+ : a **general** framework for structured argumentation that specifies guidelines that guarantee rational outcomes ....

M. Caminada and L. Amgoud. On the evaluation of argumentation formalisms. *Artificial Intelligence*, 171(5-6):286–310, 2007

H. Prakken. An abstract framework for argumentation with structured arguments. *Argument and Computation*, 1(2):93–124, 2010.

S. Modgil, H. Prakken. A General Account of Argumentation and Preferences. In: *Artificial Intelligence*. 195(0), 361 - 397, 2013.

S. Modgil and H. Prakken Corrigendum to: ``A General Account of Argumentation with Preferences'' [Artif. Intell. 195 (2013) 361-397]. In *Artificial Intelligence*, 2018.

S. Modgil and H. Prakken. Abstract rule-based argumentation. In P. Baroni, et al, editors, *Handbook of Formal Argumentation, Vol. 1*, pages 286–361. London: College Publications, 2018.

Specifying elements of an ASPIC+ theory that enable construction of Arguments

- Arguments consist of premises from which one infers/derives conclusion (claim)

- When specifying an ASPIC+ theory:

1) choose some language $\mathcal{L}$ in which to represent formulae
(e.g., you might choose $\mathcal{L}$ to be a classical propositional or first order language)

2) Specify two (possibly empty) sets of premises:

- 'ordinary' (defeasible) premises (K_p) that represent fallible facts/data that may be challenged/questioned

- 'axiom' premises (K_n) that represent infallible facts/data that therefore cannot be challenged

Specifying elements of an ASPIC+ theory that enable construction of Arguments

3) Specify *Inference Rules* which are **NOT** object level rules that one might specify in a knowledge/belief base of premises (e.g. a material implication '*penguin* $\supset$ *bird*')

- Inference rules are *metalevel* rules that license an inference (conclusion) from object level formulae in the antecedent of the rule
- Two types of inference rules in ASPIC+ (α_i, β are *wff* in chosen language $\mathcal{L}$)

1) A set of strict inference rules R_s of the form $\alpha_1,...,\alpha_n \rightarrow \beta$ that typically encode inference in a chosen deductive logic *Ds* : $R_s = \{ \alpha_1,...,\alpha_n \rightarrow \beta \mid \alpha_1,...,\alpha_n \vdash_{Ds} \beta \}$

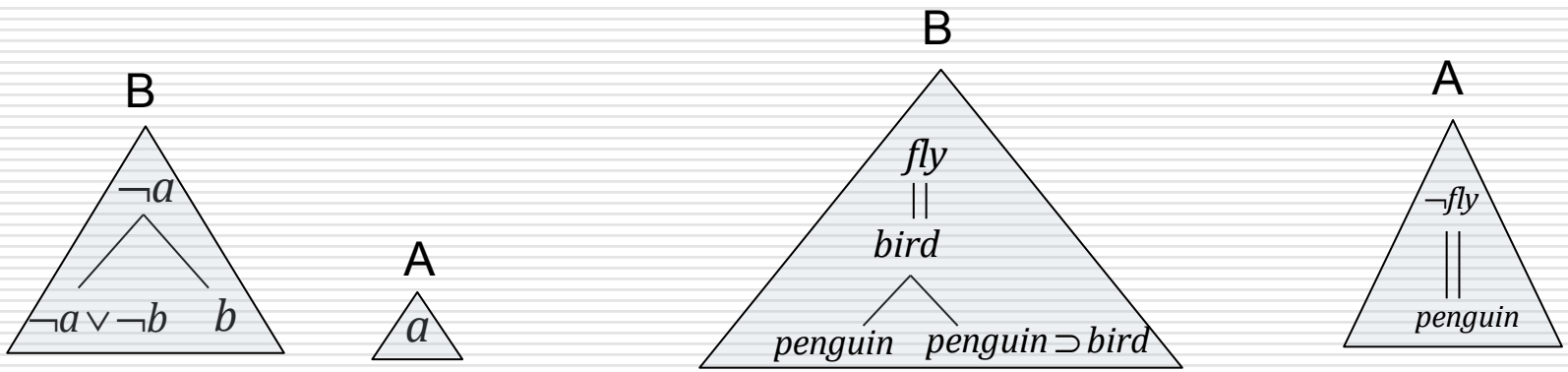
E.g $p, p \supset (q \supset r), \neg r \rightarrow \neg q \in R_s$ given $p, p \supset (q \supset r), \neg r \vdash_{CL} \neg q$

2) A set of defeasible inference rules R_d of the form $\alpha_1,...,\alpha_n \Rightarrow \beta$ that typically encode default inferences β normally/usually/typically (as opposed to necessarily) follows from $\alpha_1,...,\alpha_n$

E.g *bird* $\Rightarrow$ *fly* *penguin* $\Rightarrow \neg$*fly*

ASPIC+ Arguments

ASPIC+ arguments are upside down trees chaining leaves (premises) to conclusion through application of 0 or more inference rules



Specifying elements of an ASPIC+ theory that enable identification of *Attacks*

4) Specify a function that declares for *wff* in $\mathcal{L}$ whether they (symmetrically or asymmetrically) conflict with each other

Eg:

$\forall \alpha, \neg\alpha \in \mathcal{L}$: α and $\neg\alpha$ are ***contradictories*** (symmetric conflict)

$\forall \alpha, NOT\alpha \in \mathcal{L}$: α is a **contrary** of $NOT\alpha$ (asymmetric conflict)

married and *bachelor* are contradictories

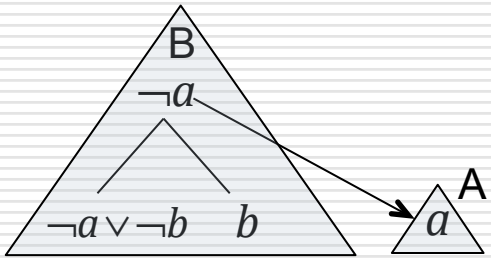
5) Specify a function *n* that names defeasible inference rules

Eg:

$n(penguin \Rightarrow \neg\ fly) = appPF$

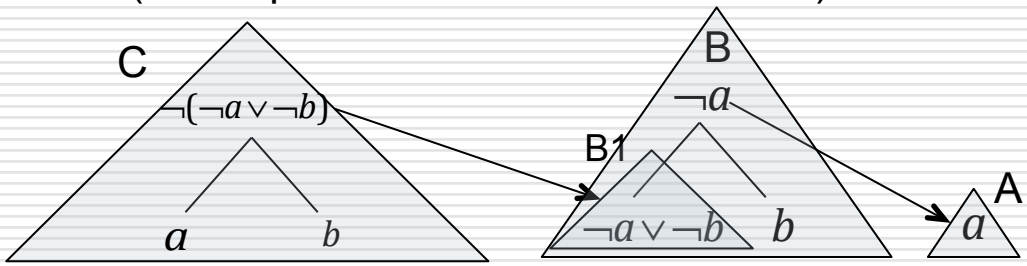
ASPIC+ Attacks

Undermine attacks from a conclusion that is a contradictory/contrary of an **ordinary premise** (axiom premises cannot be attacked) in attacked argument



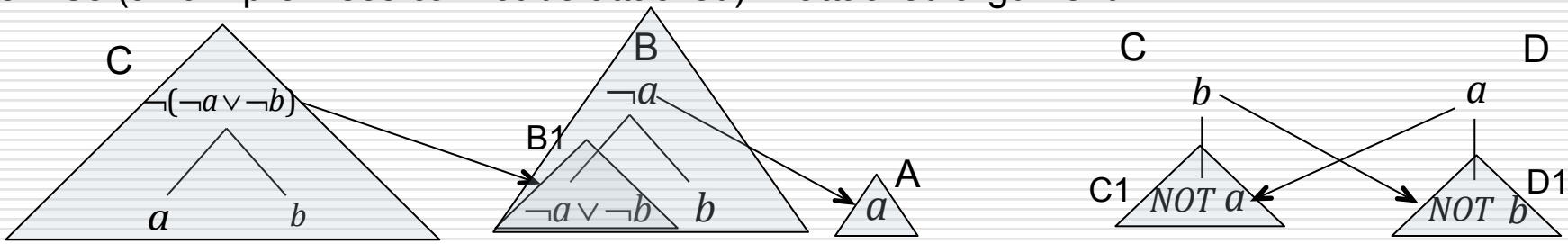
ASPIC+ Attacks

Undermine attacks from a conclusion that is a contradictory/contrary of an **ordinary premise** (axiom premises cannot be attacked) in attacked argument



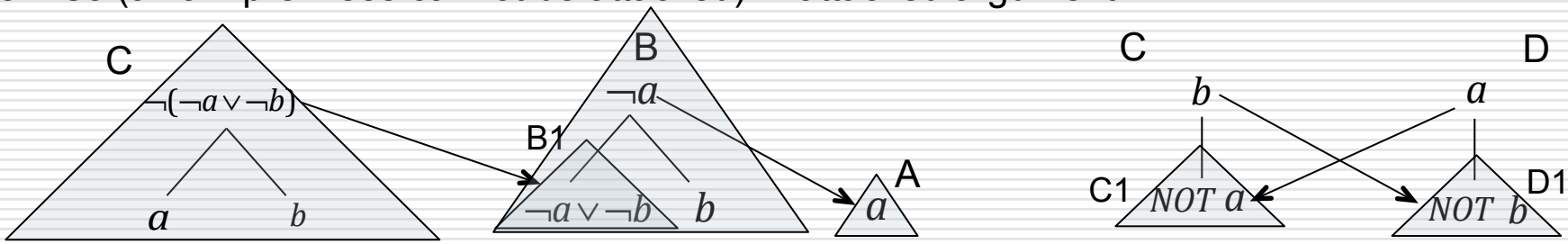
ASPIC+ Attacks

Undermine attacks from a conclusion that is a contradictory/contrary of an **ordinary premise** (axiom premises cannot be attacked) in attacked argument

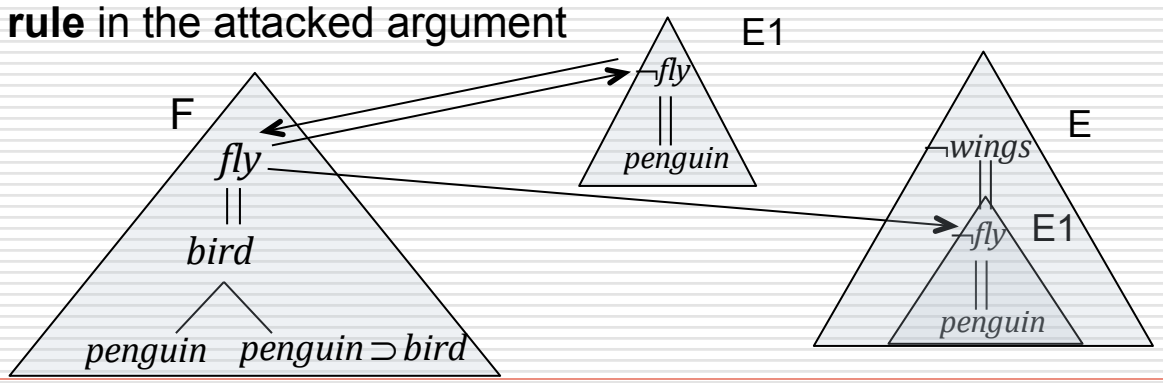


ASPIC+ Attacks

Undermine attacks from a conclusion that is a contradictory/contrary of an **ordinary premise** (axiom premises cannot be attacked) in attacked argument



Rebut attacks from a conclusion that is a contradictory/contrary of the consequent of a **defeasible rule** in the attacked argument



ASPIC+ Attacks

Undercut attacks from a conclusion that negates the name of a defeasible rule – the attacking argument provides rationale invalidating defeasible/presumptive inference from antecedent of attacked rule to consequent

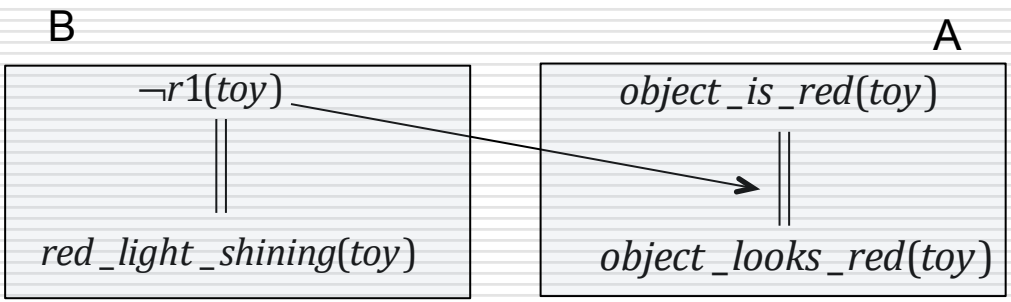
$$\forall x.red\_light\_shining(X) \Rightarrow \neg r1(X)$$

$$\forall x.object\_looks\_red(X) \Rightarrow object\_is\_red(X)$$

$$n(\forall x.object\_looks\_red(X) \Rightarrow object\_is\_red(X)) = r1(X)$$

$$red\_light\_shining(toy)$$

$$object\_looks\_red(toy)$$



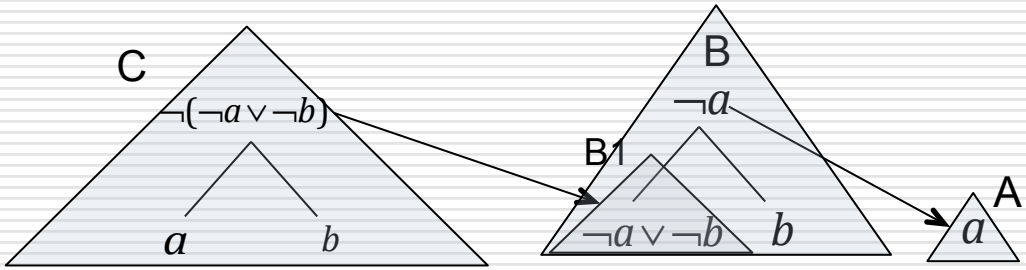
ASPIC+ Defeats

Given a strict partial ordering $\prec$ over arguments one then determines which undermine attacks and rebuts on **contradictories** succeed as defeats

$B \prec A$ denotes that *A* is strictly preferred to (stronger than) B where this preference relation can be defined in your way of choosing (e.g., based on the relative certainty/entrenchment of arguments' constituent ordinary premises and/or defeasible rules, or the relative reliability of the sources of these argument)

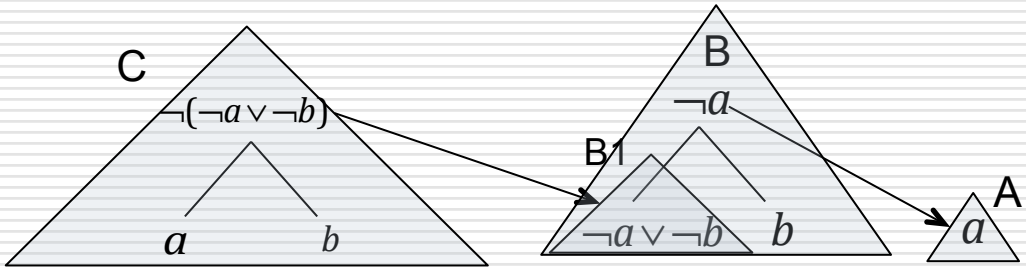
Undercut attacks and attacks on **contraries** always succeed as defeats (i.e., independently of preferences)

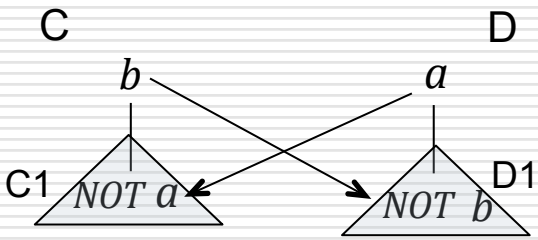
ASPIC+ Defeats



- C *attacks* B on B1 and *defeats* B on B1 only if C $\not\prec$ B1
- B *attacks* A on A and *defeats* A on A only if B $\not\prec$ A

ASPIC+ Defeats



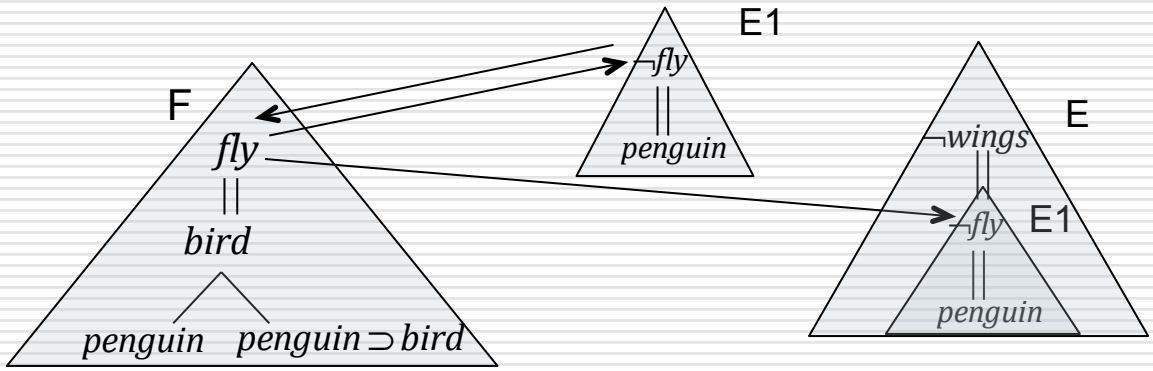


- C *attacks* B on B1 and *defeats* B on B1 only if C $\not\prec$ B1
- B *attacks* A on A and *defeats* A on A only if B $\not\prec$ A

C attacks D on D1 and D attacks C on C1

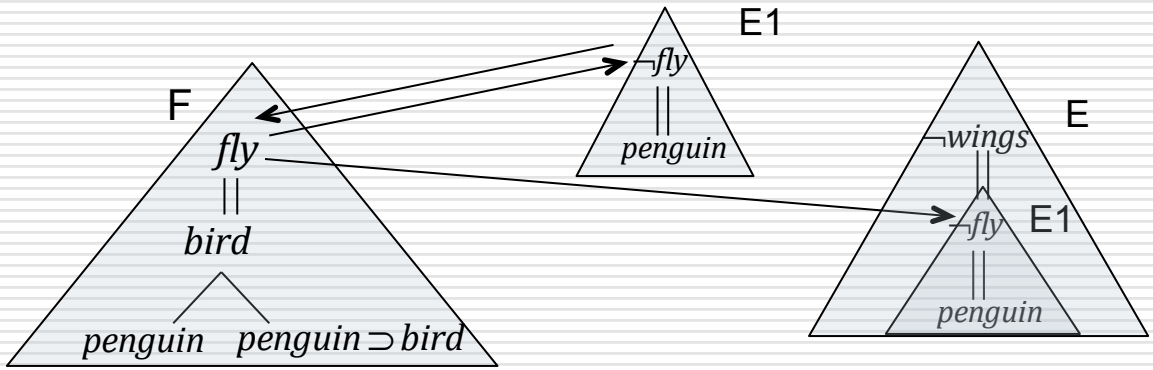
Both attacks are on contraries and so succeed as defeats independently of preferences

ASPIC+ Defeats

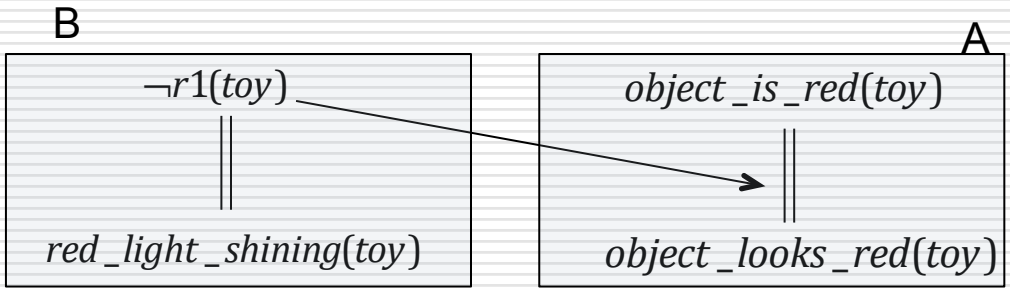


- F *attacks* E1 on E1 and *defeats* E1 on E1 only if F $\not\prec$ E1
- F *attacks* E on E1 and *defeats* E on E1 only if F $\not\prec$ E1

ASPIC+ Defeats



- F *attacks* E1 on E1 and *defeats* E1 on E1 only if F $\not\prec$ E1
- F *attacks* E on E1 and *defeats* E on E1 only if F $\not\prec$ E1



B attacks A on A and defeats A independently of preferences

To summarise ….

- Define for some arbitrary language $\mathcal{L}$:

1) KB of infallible (axiom) and/or fallible (defeasible) premises that are *wff* in $\mathcal{L}$
2) Strict and/or defeasible rules *inference* rules respectively encoding inference in some deductive logic and domain specific defeasible/default inferences
3) Function declaring when one formula symmetrically or asymmetrically conflicts with another
4) Naming function assigning names (*wff* in $\mathcal{L}$) to defeasible inference rules

- Construct arguments $\mathcal{A}rgs$ from KB of premises and inference rules and define binary attack relation $\mathcal{A}tt \subseteq \mathcal{A}rgs$ x$\mathcal{A}rgs$
- Given $\prec \subseteq \mathcal{A}rgs$ x$\mathcal{A}rgs$ define binary defeat relation $\mathcal{D}ef \subseteq \mathcal{A}rgs$ x$\mathcal{A}rgs$
- Define extensions and so justified arguments of Dung AF ($\mathcal{A}rgs$,$\mathcal{D}ef$)

Argumentation-based characterisations of non-monotonic inference relations in ASPIC+

$(\mathcal{Args},\mathcal{Def})_\Delta \mid\!\sim \alpha$ (the claim of an argument in grounded extension)

iff

$\Delta \mid\!\sim_{LP} \alpha$ (under well founded semantics of ***logic programming***)

$(\mathcal{Args},\mathcal{Def})_\Delta \mid\!\sim \alpha$ (claim of argument in stable ext.)

iff

$\Delta \mid\!\sim_{PDL} \alpha$ (conclusion of formula in default extension)

PDL – ***Prioritised Default Logic***

(W,D,<) where < is a priority ordering over defaults in D

(W = axiom premises, D = defeasible inference rules + $R_{S(CL)}$)

Other Instances of ASPIC+

- One can define novel non-monotonic logics in ASPIC+ e.g.,

H. Dong, B. Liao, R. Markovich, and L. van der Torre. From classical to non-monotonic deontic logic using ASPIC+. In: *International Workshop on Logic, Rationality and Interaction (LORI2019)*, pages 71–85. Springer, 2019.

- Other approaches to argumentation can be formalised as instances of ASPIC+ e.g. *Assumption Based Argumentation*, *Classical Logic Argumentation*, *Carneades*, *Schemes and Critical Questions*. ....

For more details, see in particular:

S. Modgil and H. Prakken. Abstract rule-based argumentation. In P. Baroni, et al, editors, *Handbook of Formal Argumentation, Vol. 1*, pages 286–361., 2018.

Example: An Argumentative characterisation of *Preferred Subtheories*

- Brewka's Preferred Subtheories is a well known non-monotonic formalism

Totally ordered set of classical wff $\Delta = (\Delta_1,\ldots,\Delta_n)$ s.t. $\alpha \in \Delta_i, \beta \in \Delta_j . i < j \rightarrow \alpha > \beta$

Preferred Subtheories of $(\Delta,>)$ are $>$ ordered maximal consistent subsets of Δ

$$\frac{\dfrac{a,\neg a}{c, c \supset e, \neg e}}{f}$$

$\{a\}$ → $\{a,c,c\supset e\}$, $\{a,c,\neg e\}$, $\{a,c\supset e,\neg e\}$

$\{\neg a\}$ → $\{\neg a,c,c\supset e\}$, $\{\neg a,c,\neg e\}$, $\{\neg a,c\supset e,\neg e\}$

$E1=\{a,c,c\supset e,f\}$
$E2=\{a,c,\neg e,f\}$
$E3=\{a,c\supset e,\neg e,f\}$

$E1=\{\neg a,c,c\supset e,f\}$
$E2=\{\neg a,c,\neg e,f\}$
$E3=\{\neg a,c\supset e,\neg e,f\}$

$\Delta \mid\!\sim_{PS\text{-}cr} \alpha$ if $\exists E_i \alpha \in Cn(E_i)$

$\Delta \mid\!\sim_{PS\text{-}scp} \alpha$ if $\forall E_i \alpha \in Cn(E_i)$

ASPIC+ argumentative characterisation of Preferred Subtheories

Given (Δ,>) : $R_d = \emptyset$, $R_s = \{\alpha_1,\ldots,\alpha_n \rightarrow \beta \mid \alpha_1,\ldots,\alpha_n \vdash_{CL} \beta\}$ $K_n = \emptyset$, $K_p = \Delta$

$\mathcal{A}rgs = \{\Gamma \rightarrow \beta \mid \Gamma$ is a *consistent subset minimal* subset of Δ that classically entails $\beta\}$

Writing (Γ,β) instead of $\Gamma \rightarrow \beta$:

X = (Γ,β) *undermine attacks* Y = (Π,α) if $-\beta \in \Pi$

X attacks Y on subargument Y' = $(\{-\beta\},-\beta)$

Define argument ordering $\prec$ on basis of underlying ordering on formulae in Δ
X $\prec$ Y if $\exists \delta \in prem(X), \forall \varphi \in prem(Y) : \delta < \varphi$

X *defeats* Y if X *attacks* Y on Y' and X $\not\prec$ Y'

ASPIC+ argumentative characterisation of Preferred Subtheories

$(\mathcal{A}rgs, \mathcal{D}ef)_{\Delta} \mid\!\sim \alpha$ (α is the conclusion of an argument in a stable extension)

iff

$\Delta \mid\!\sim_{PS\text{-}cr} \alpha$

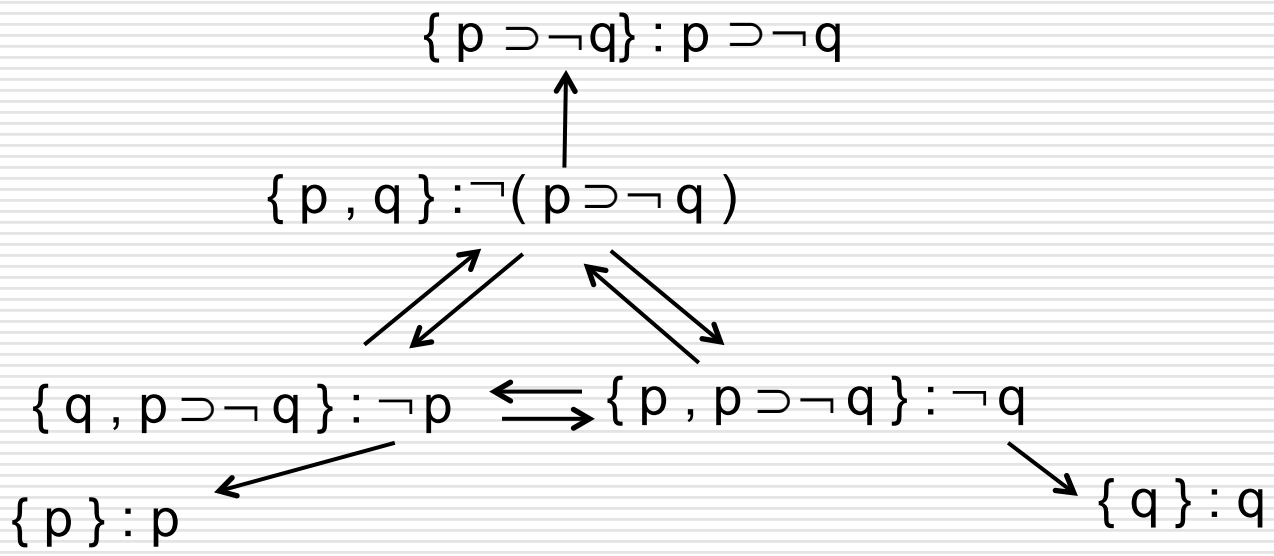
$(\mathcal{A}rgs, \mathcal{D}ef)_{\Delta} \mid\!\sim \alpha$ (every stable extension contains an argument concluding α)

iff

$\Delta \mid\!\sim_{PS\text{-}scp} \alpha$

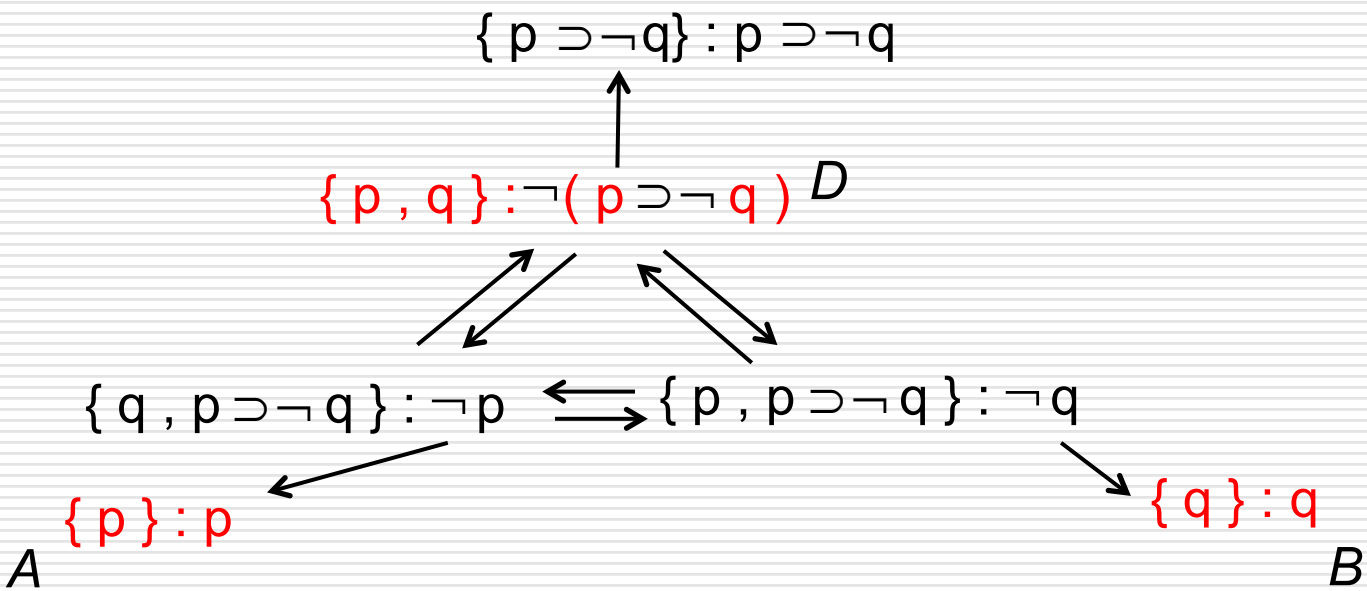
An Example

- Part of AF defined by $\Delta = ((p , q , p \supset \neg q), > = \emptyset)$



An Example

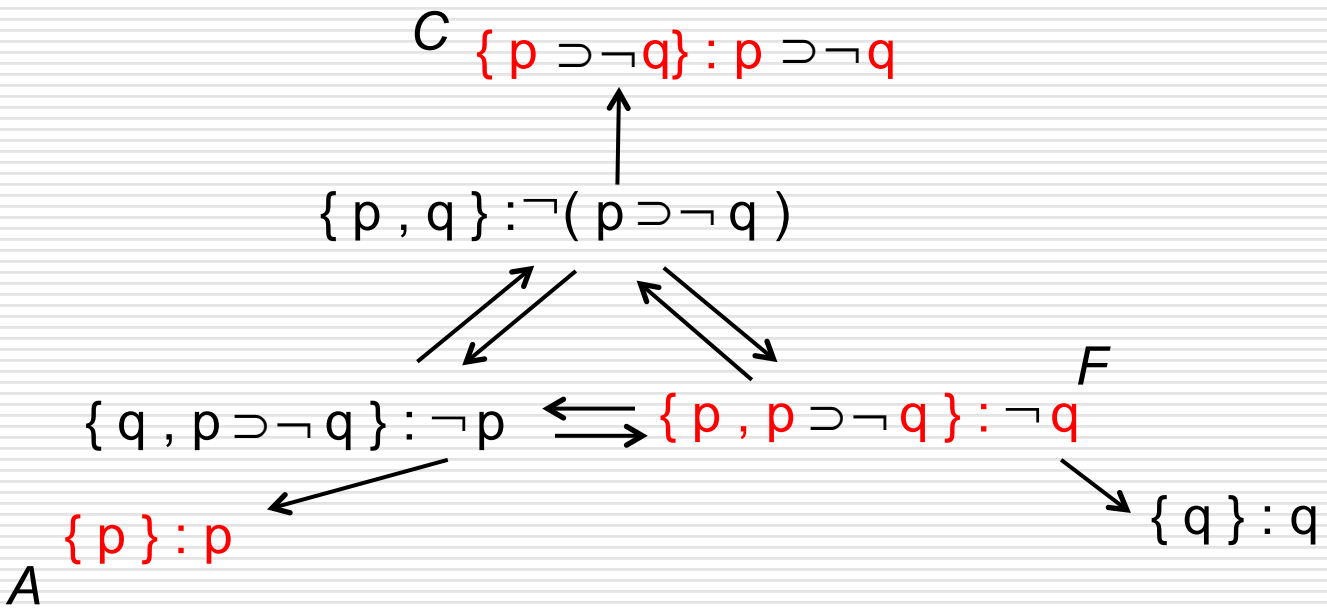
- Framework defined by $\Delta = ((p , q , p \supset \neg q), > = \varnothing)$



$E1 = \{p,q\}$ $\quad$ $S1 = \{A,B,D,E = (\{p,q\},p \wedge q)\ldots\ldots\ldots\}$

An Example

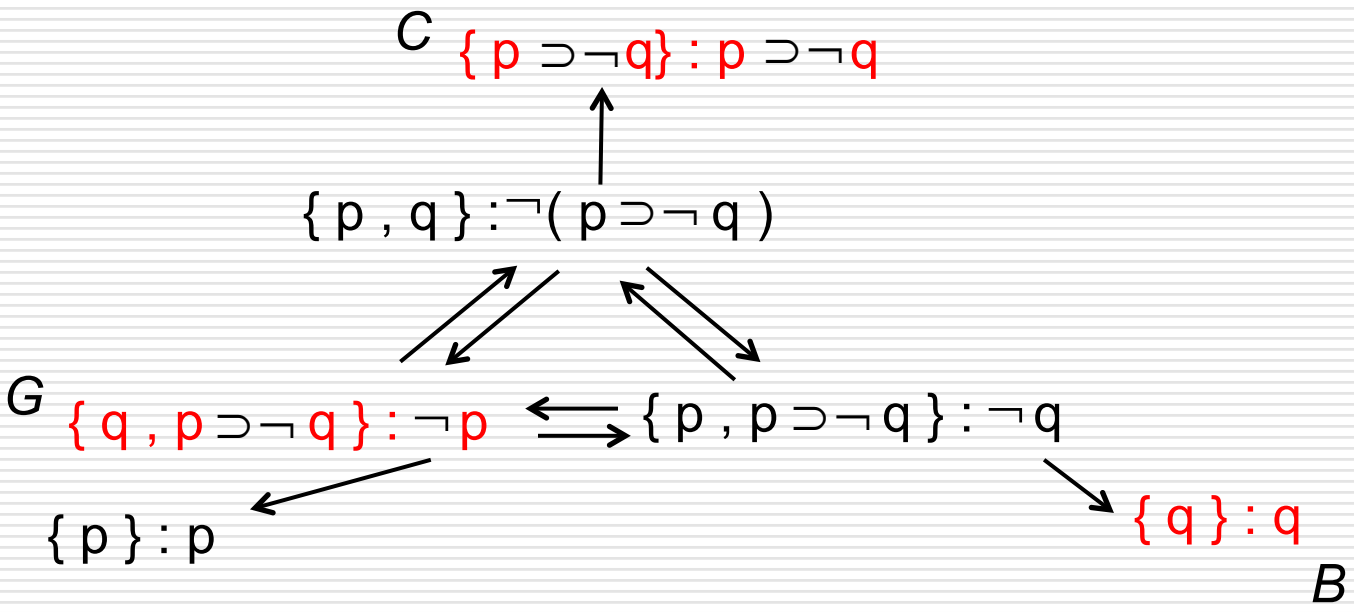
- Framework defined by $\Delta = ((p , q , p \supset \neg q), > = \varnothing)$



$E2 = \{p, p \supset \neg q\}$ $\quad$ $S2 = \{A, C, F, ...\}$

An Example

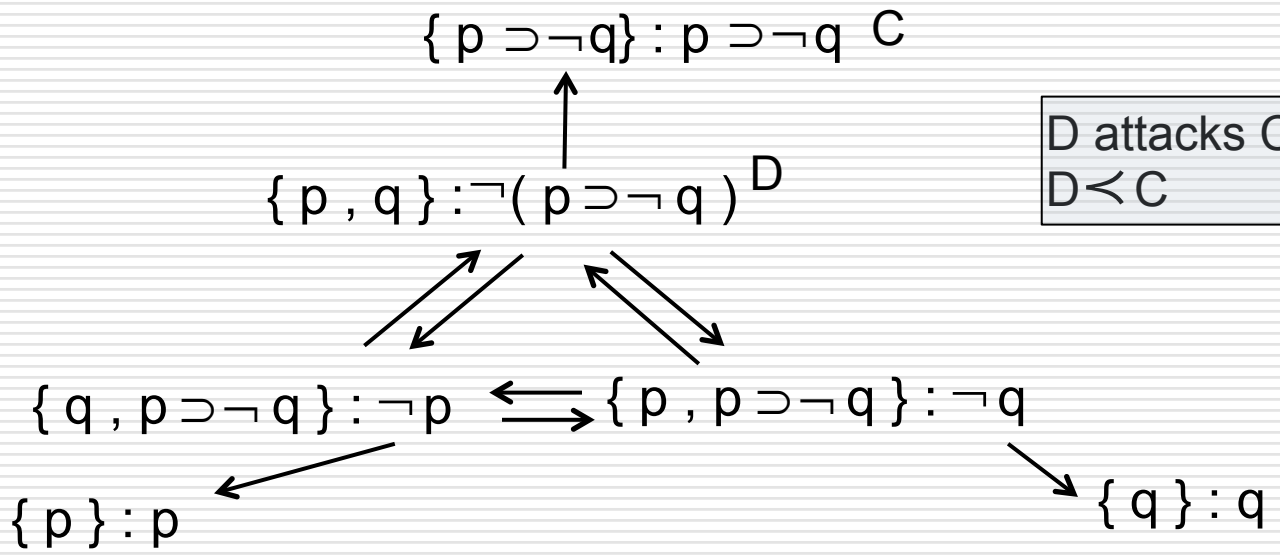
- Framework defined by $\Delta = ((p , q , p \supset \neg q), > = \varnothing)$



$E3 = \{q, p \supset \neg q\}$ $S3 = \{B, C, G, ...\}$

Example with Preferences

- Framework defined by $\Delta = ((p , q , p \supset \neg q), p\ p \supset \neg q > q)$



D attacks C on C
$D \prec C$

Example with Preferences

- Framework defined by $\Delta = ((p , q , p \supset \neg q), p, p \supset \neg q > q)$

$$\{ p \supset \neg q \} : p \supset \neg q$$

$$\{ p , q \} : \neg (p \supset \neg q)$$

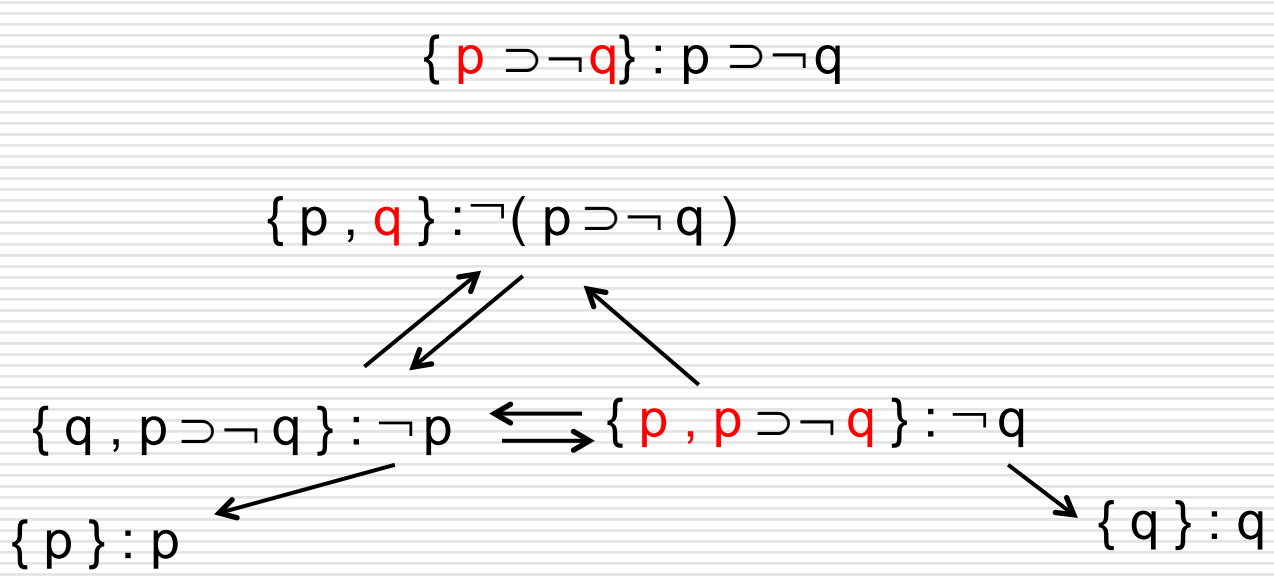
$$\{ q , p \supset \neg q \} : \neg p \rightleftarrows \{ p , p \supset \neg q \} : \neg q$$

$$\{ p \} : p$$

$$\{ q \} : q$$

Example with Preferences

- Framework defined by $\Delta = ((p , q , p \supset \neg q), p, p \supset \neg q > q)$



Example with Preferences

- Framework defined by $\Delta = ((p , q , p \supset \neg q), p, p \supset \neg q > q)$

$$\{ p \supset \neg q \} : p \supset \neg q$$

$$\{ p , q \} : \neg (p \supset \neg q)$$

$$\{ q , p \supset \neg q \} : \neg p \rightleftarrows \{ p , p \supset \neg q \} : \neg q$$

$$\{ p \} : p$$

$$\{ q \} : q$$

Example with Preferences

- Framework defined by $\Delta = ((p , q , p \supset \neg q), p, p \supset \neg q > q)$

$$\{ p \supset \neg q \} : p \supset \neg q$$

$$\{ p , q \} : \neg (p \supset \neg q)$$

$$\{ q , p \supset \neg q \} : \neg p \longleftarrow \{ p , p \supset \neg q \} : \neg q$$

$$\{ p \} : p \qquad\qquad \{ q \} : q$$

Example with Preferences

- Framework defined by $\Delta = ((p , q , p \supset \neg q), p, p \supset \neg q > q)$

$$\{ p \supset \neg q \} : p \supset \neg q$$

$$\{ p , q \} : \neg (p \supset \neg q)$$

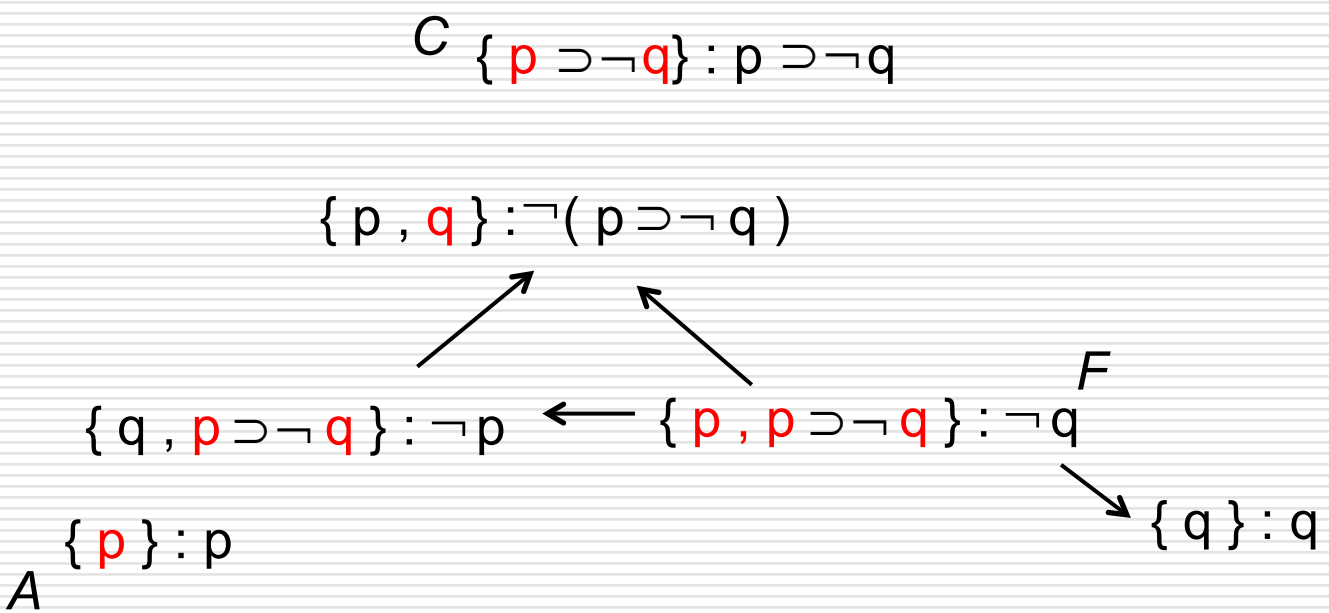
$$\{ q , p \supset \neg q \} : \neg p \longleftarrow \{ p , p \supset \neg q \} : \neg q$$

$$\{ p \} : p$$

$$\{ q \} : q$$

Example with Preferences

- Framework defined by $\Delta = ((p , q , p \supset \neg q), p, p \supset \neg q > q)$



$E2 = \{p, p \supset \neg q\}$ $\quad$ $S2 = \{A, C, F, ...\}$

Another Example: Using ASPIC+ to Formalise Argumentation with Schemes and Critical Questions

- ASPIC+ formalises argumentation using logical representations of *schemes and critical questions (*widely studied in philosophy by Doug Walton * and many others)

- Over 100 schemes which are generic templates capturing stereotypical patterns of arguments and can be used to constructing arguments when formalised as defeasible inference rules in ASPIC+

- *Critical questions* identify presumptions in arguments that can be challenged, and so act as pointers to counter-arguments that are themselves instances of schemes

- D. Walton, *Argument Schemes for Presumptive Reasoning,* Lawrence Erlbaum Associates, 1996.

Schemes and Critical Questions

Action Scheme

In situation S
A achieves goal G
Which promotes value V
And so action A is recommended

16 CQ

E.g., Is S true ?
Does A achieve G ?
Is there an alternative A' that achieves G ?

Schemes and Critical Questions

- *Schemes formalised as defeasible inference rules and critical questions* used as pointers to counter-arguments that are themselves instances of schemes

S = Kurdish ethnic cleaning
A = enforce protected zone
G = stop cleansing
V = right to life

Schemes and Critical Questions

- *Schemes formalised as defeasible inference rules and critical questions* used as pointers to counter-arguments that are themselves instances of schemes

S = Kurdish ethnic cleaning
A = enforce protected zone
G = stop cleansing
V = right to life

Does A achieve G ?

Expert opinion scheme

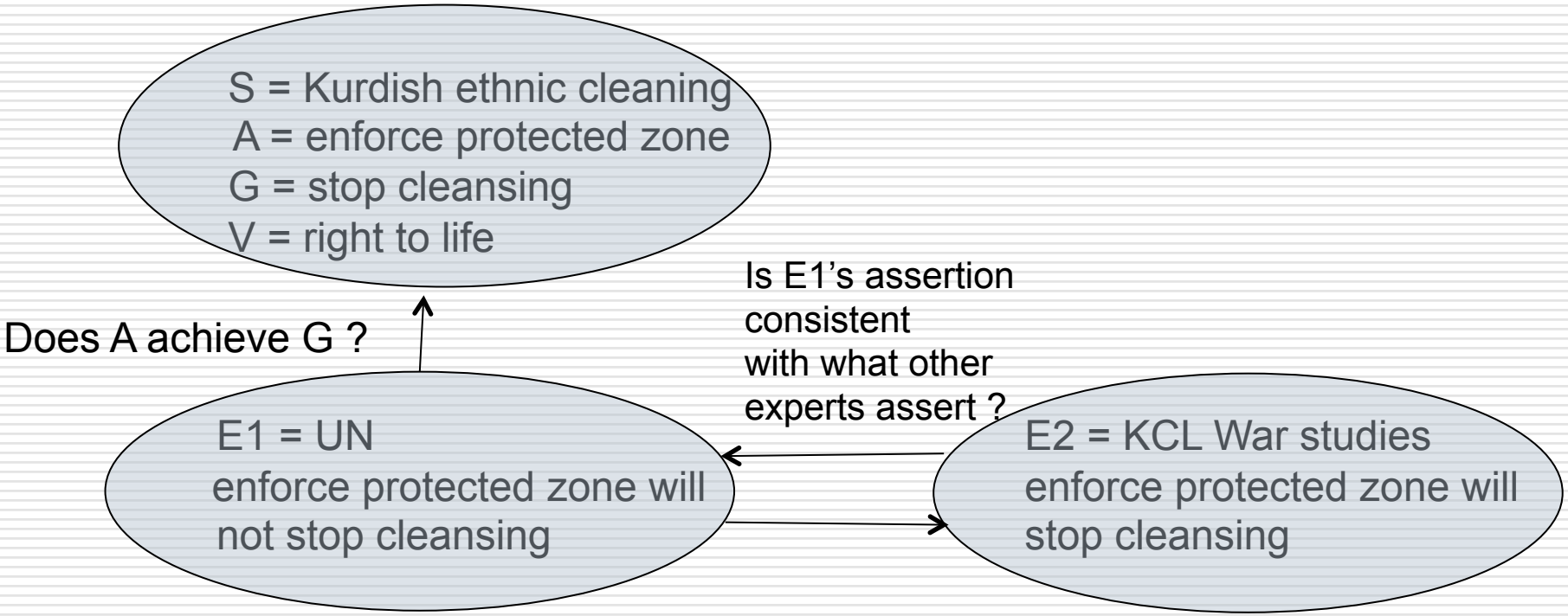
E is an expert in domain D
E says S is true/false
S is a statement in domain D
Therefore S is true/false

Is E reliable ?
Does domain D contain statement S ?
Is E's assertion consistent with what other experts assert ?

.

Argumentation using schemes and critical questions in ASPIC+ *

- *Schemes formalised as defeasible inference rules in ASPIC+ and critical questions* used as pointers to counter-arguments that are themselves instances of schemes



- See Section 3.5 S. Modgil and H. Prakken. Abstract rule-based argumentation. In *Handbook of Formal Argumentation, Vol. 1*, pages 286–361., 2018 for first order examples

ASPIC+ and Rationality

ASPIC+ and the Rationality Postulates

ASPIC+ specifies guidelines on how to specify:

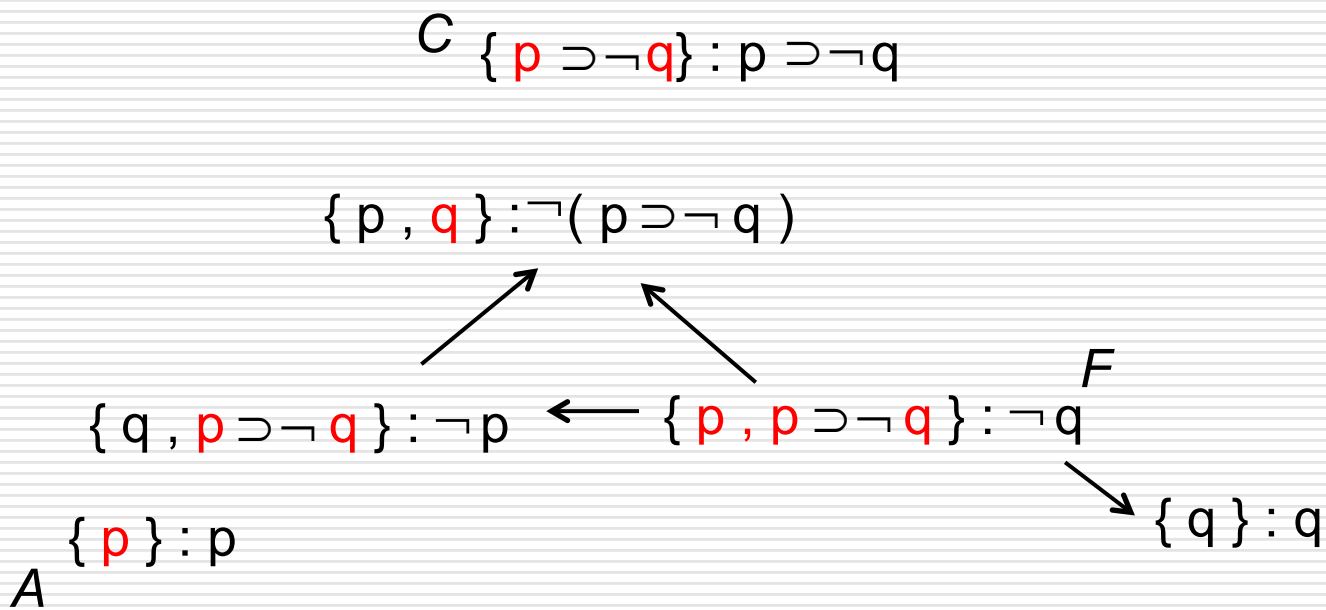
1) axiom premises;
2) strict inference rules;
3) definition of attacks, and
4) the ways in which preferences are defined,

so as to guarantee satisfaction of consistency and closure rationality postulates

- **Consistency** *Claims and premises of arguments in* ***complete*** *extension of AF are mutually consistent*
- **Sub-argument closure** *If X is in a* ***complete*** *extension E of AF then all sub-arguments of X are also in* ***E***
- **Closure under deductive\strict rules** *If conclusions of arguments in a* ***complete*** *extension E deductively (strictly) entail β then E contains an argument concluding β*

Closure under Strict Rules

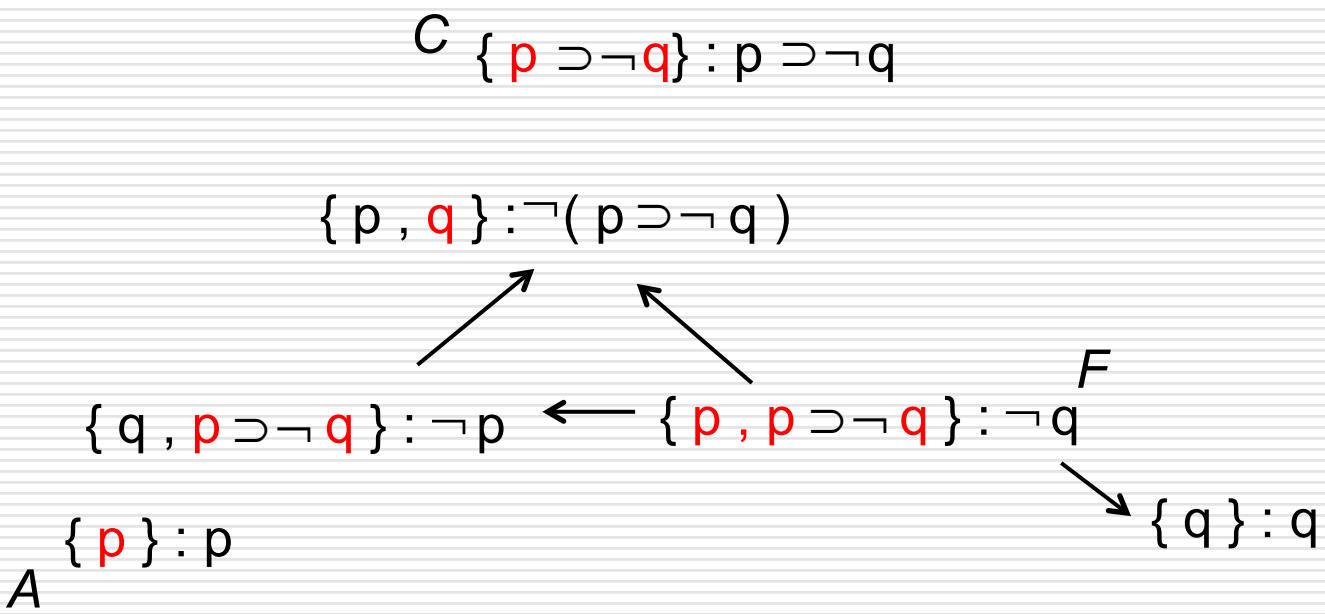
- Part of AF defined by $\Delta = ((p, q, p \supset \neg q), p, p \supset \neg q > q)$



Complete/preferred/grounded\stable extension E includes {A, F, C } and is closed under $Rs = \vdash_{CL}$

$$\forall \alpha \in Cn_{CL}(p \supset \neg q, p, \neg q) \quad \exists X \in E \quad \text{s.t.} \quad Conc(X) = \alpha$$

Sub-argument Closure



E includes {A, F, C } and all sub-arguments of A, F and C, i.e.,

$$\{(\Delta,\alpha) \mid \Delta \subseteq \{p, p \supset \neg q, \neg q\}, \Delta \vdash_{CL} \alpha\}$$

Consistency

C $\{ p \supset \neg q \} : p \supset \neg q$

$\{ p , q \} : \neg (p \supset \neg q)$

F

$\{ q , p \supset \neg q \} : \neg p$ $\longleftarrow$ $\{ p , p \supset \neg q \} : \neg q$

$\{ q \} : q$

$\{ p \} : p$

A

E = {A, F, C } satisfies consistency

$$\boxed{\nexists (\Delta, \alpha), (\Gamma, \neg\alpha) \in E}$$

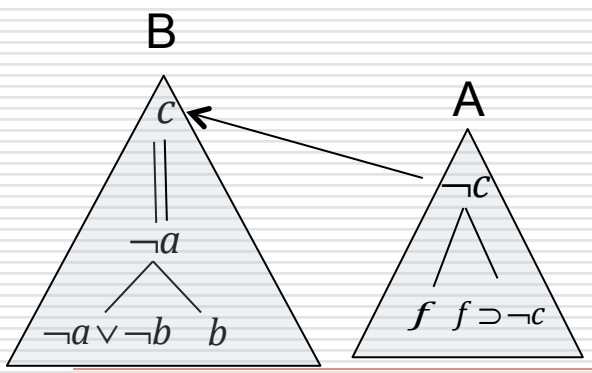
Key Guidelines for satisfying Rationality Postulates

1) Obvious requirement that axiom premises (K_n) are consistent

$$\nexists \Gamma \subseteq K_n \text{ s.t. } \Gamma \rightarrow \alpha \quad \Gamma \rightarrow -\alpha$$

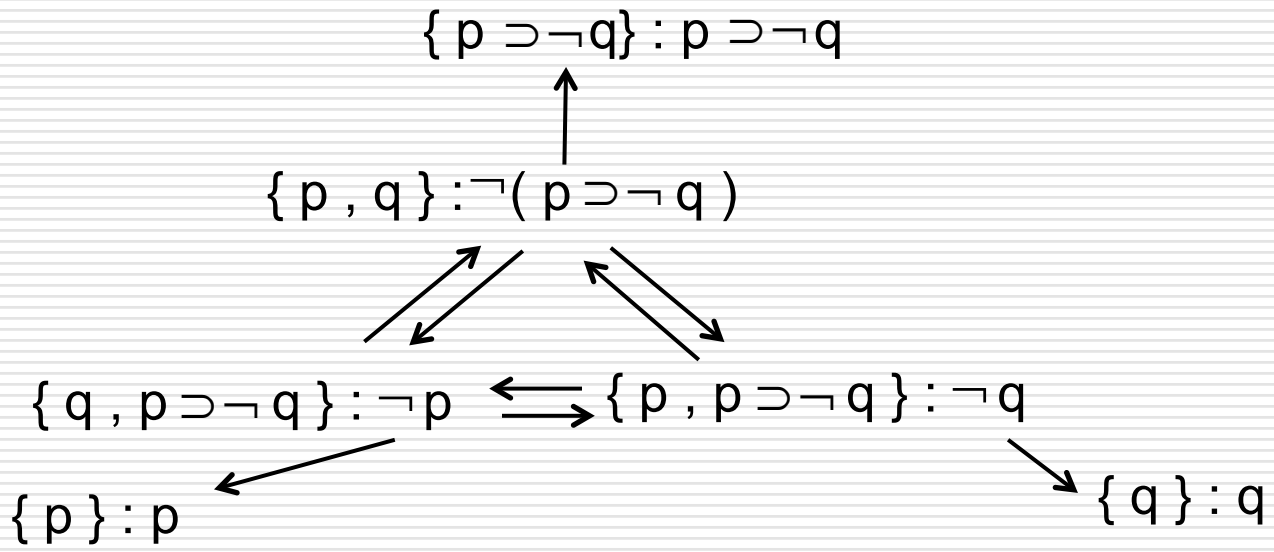
2) Restricted Rebut

Rebut attacks cannot target the conclusions of strict rules; only the conclusions of defeasible rules can be targeted

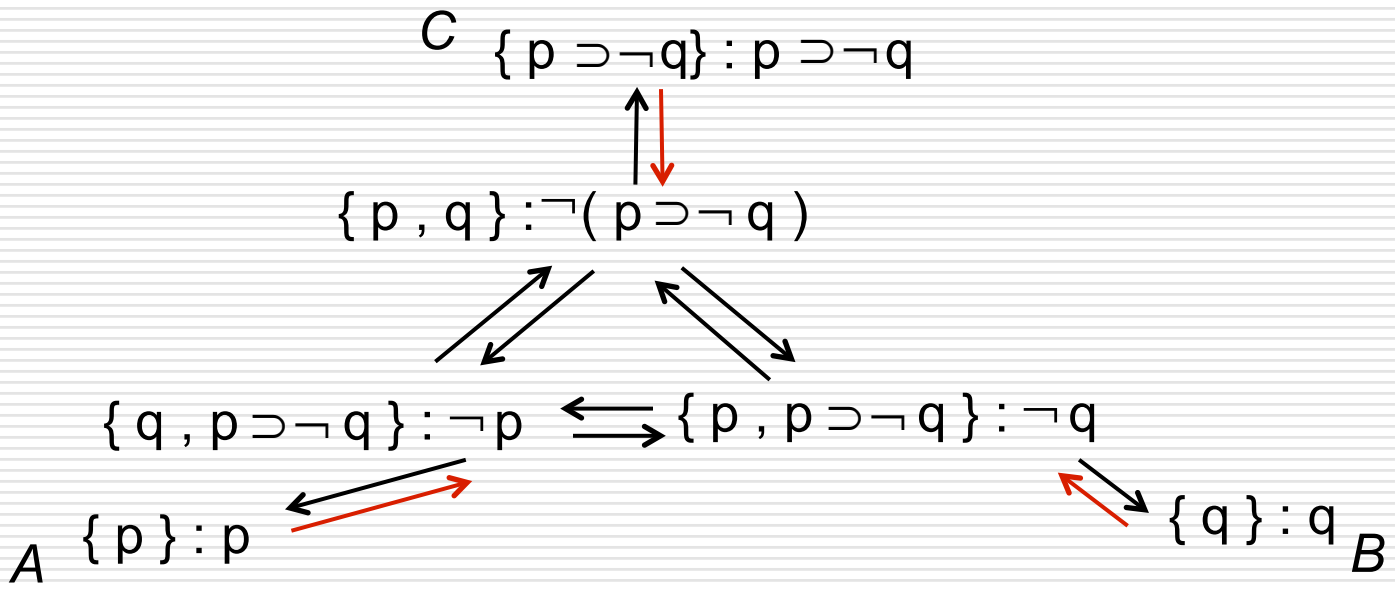


Suppose we allowed rebuts on the conclusions of strict rules

- $\Delta = ((p , q , p \supset \neg q), > = \emptyset)$



Suppose we allowed rebuts on the conclusions of strict rules



{*A,B,C* ...} is admissible and so a subset of a complete extension that does not satisfy consistency (under closure of strict rules = *indirect consistency*)

Key Guidelines for satisfying Rationality Postulates

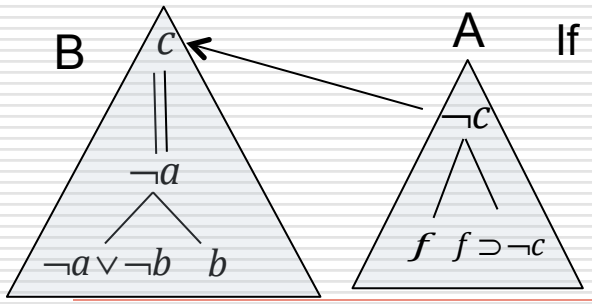
1) Obvious requirement that axiom premises (K_p) are consistent

$$\nexists \Gamma \subseteq K_p \text{ s.t. } \Gamma \rightarrow \alpha \quad \Gamma \rightarrow -\alpha$$

2) Restricted Rebut

Rebut attacks cannot target the conclusions of strict rules; only the conclusions of defeasible rules can be targeted

3) Strict Rules are closed under 'contraposition/transposition'

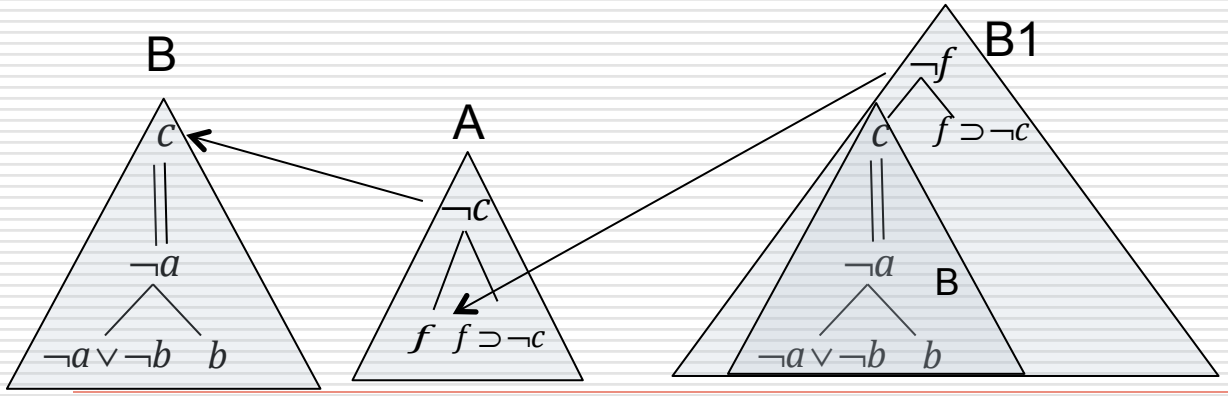


If $f, f \supset \neg c \rightarrow \neg c \in R_S$ then

$c, f \supset \neg c \rightarrow \neg f \in R_S$ and $f, c \rightarrow \neg(f \supset \neg c) \in R_S$

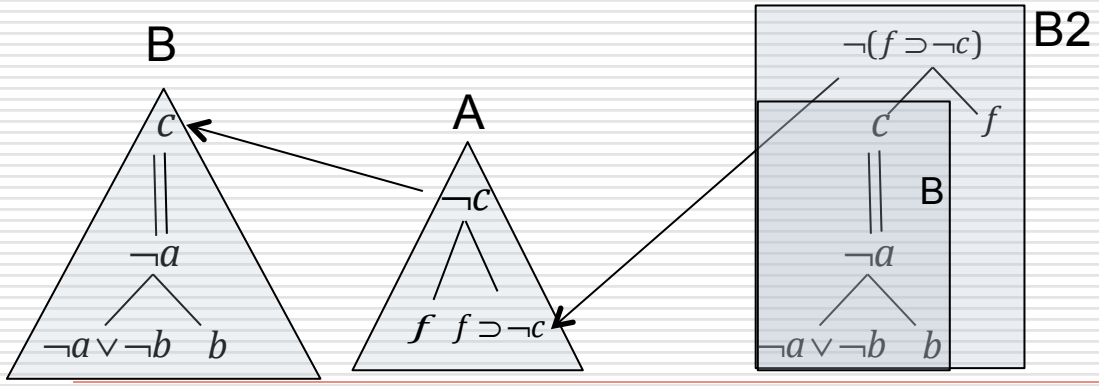
Key Guidelines for satisfying Rationality Postulates

Extending B with $c, f \supset \neg c \rightarrow \neg f \in R_S$ and so we also have B1



Key Guidelines for satisfying Rationality Postulates

Extending B with $f, c \rightarrow \neg(f \supset \neg c) \in R_S$ and so we also have B2



Key Guidelines for satisfying Rationality Postulates

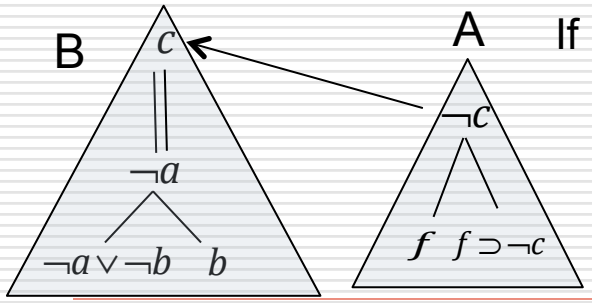
1) Obvious requirement that axiom premises (K_p) are consistent

$$\nexists \Gamma \subseteq K_p \text{ s.t. } \Gamma \rightarrow \alpha \quad \Gamma \rightarrow -\alpha$$

2) Restricted Rebut

Rebut attacks cannot target the conclusions of strict rules; only the conclusions of defeasible rules can be targeted

3) Strict Rules are closed under 'contraposition/transposition'



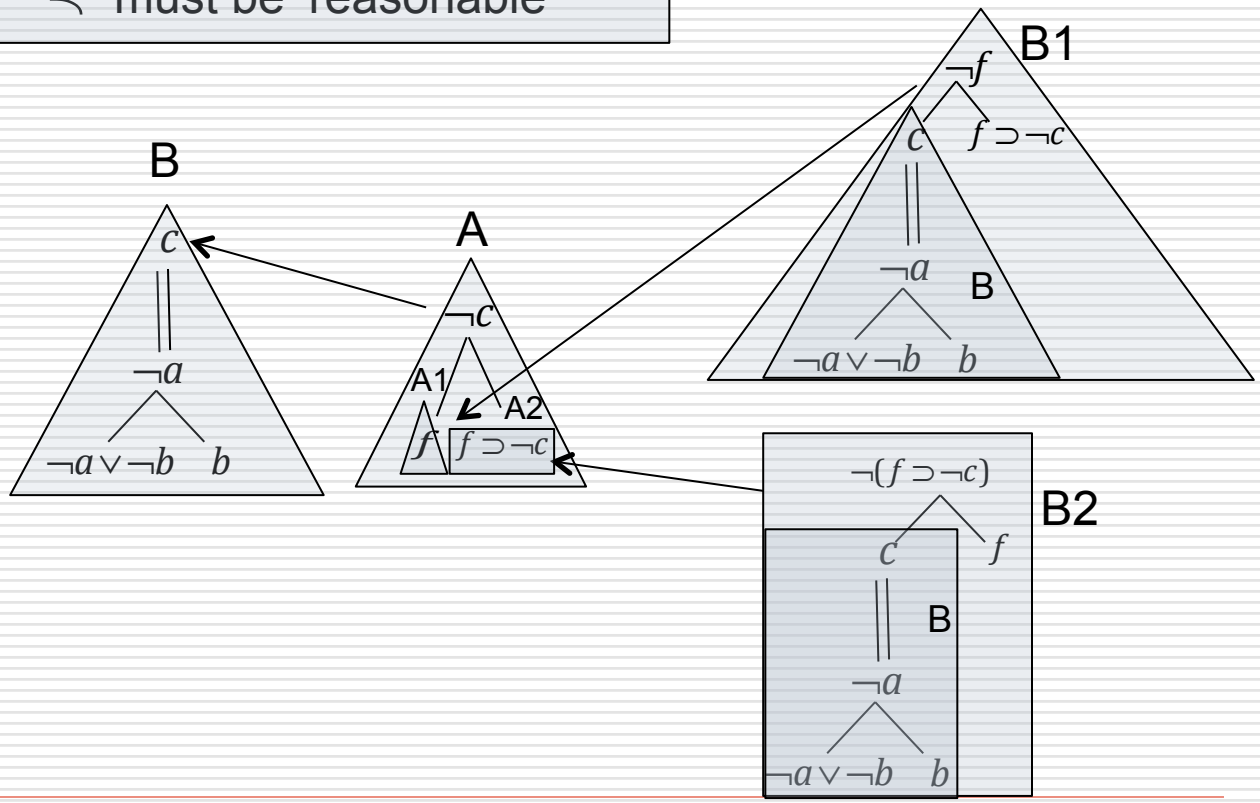
If $f, f \supset \neg c \rightarrow \neg c \in R_S$ then $c, f \supset \neg c \rightarrow \neg f \in R_S$ and $f, c \rightarrow \neg(f \supset \neg c) \in R_S$

Above satisfied given that $R_S = \{\Gamma \rightarrow \alpha \mid \Gamma \vdash_{CL} \alpha\}$

Key Guidelines for satisfying Rationality Postulates

4) Preference relation/ordering over arguments must satisfy certain properties

$\prec$ must be 'reasonable'

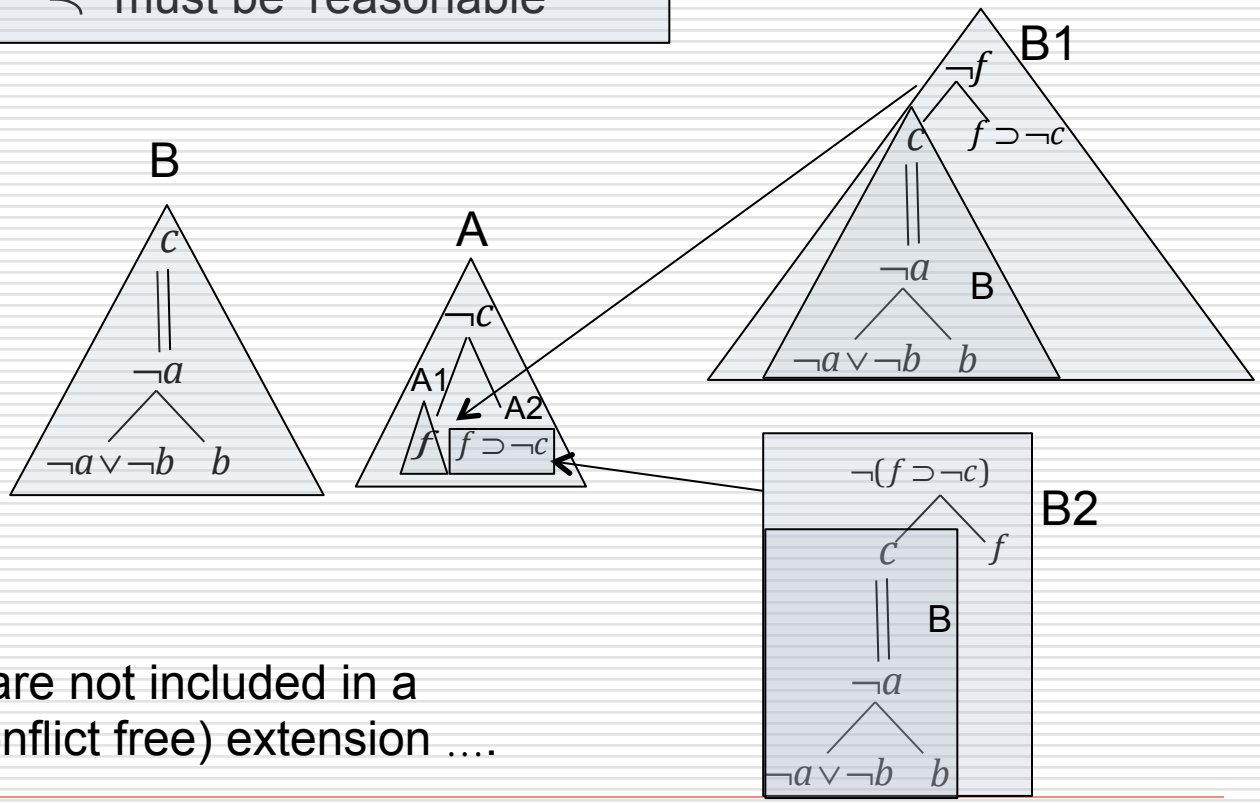


Suppose A $\prec$ B

Key Guidelines for satisfying Rationality Postulates

4) Preference relation/ordering over arguments must satisfy certain properties

$\prec$ must be 'reasonable'



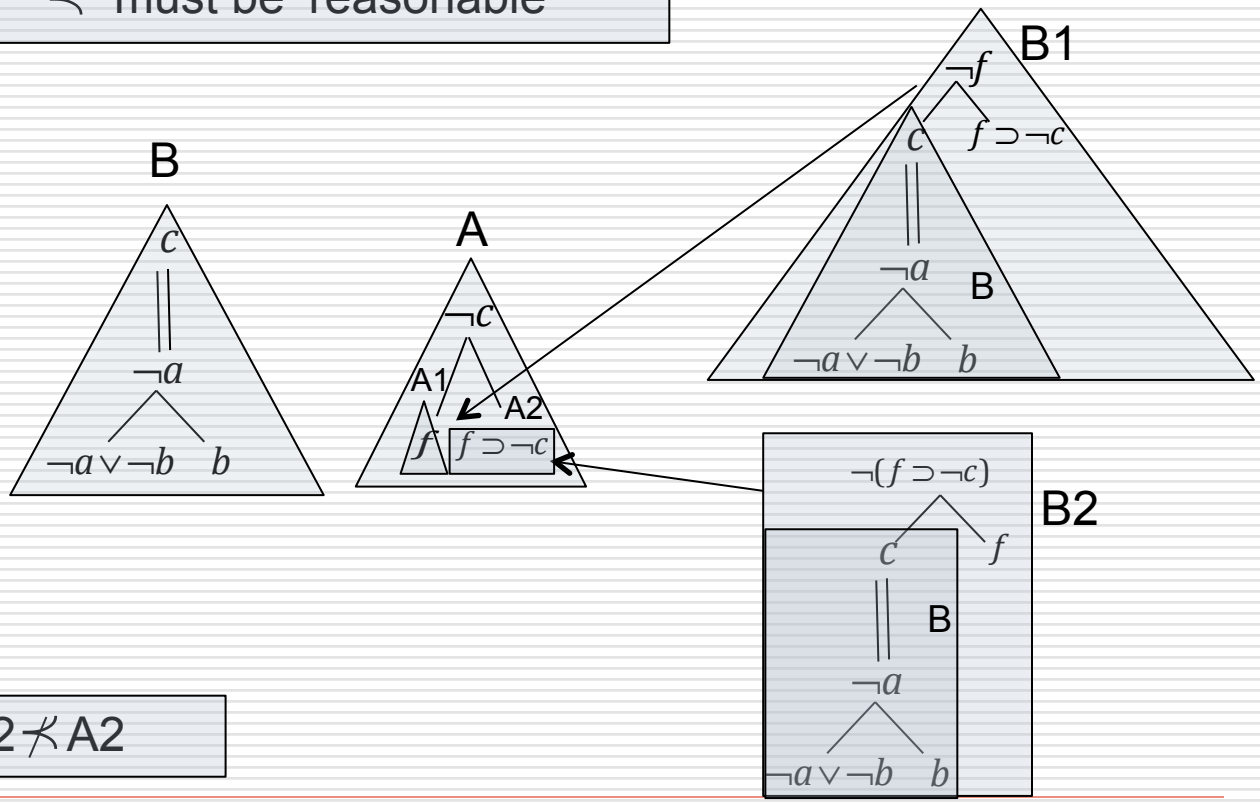
Suppose A $\prec$ B

To ensure that A and B are not included in a complete (and hence conflict free) extension ....

Key Guidelines for satisfying Rationality Postulates

4) Preference relation/ordering over arguments must satisfy certain properties

$\prec$ must be 'reasonable'

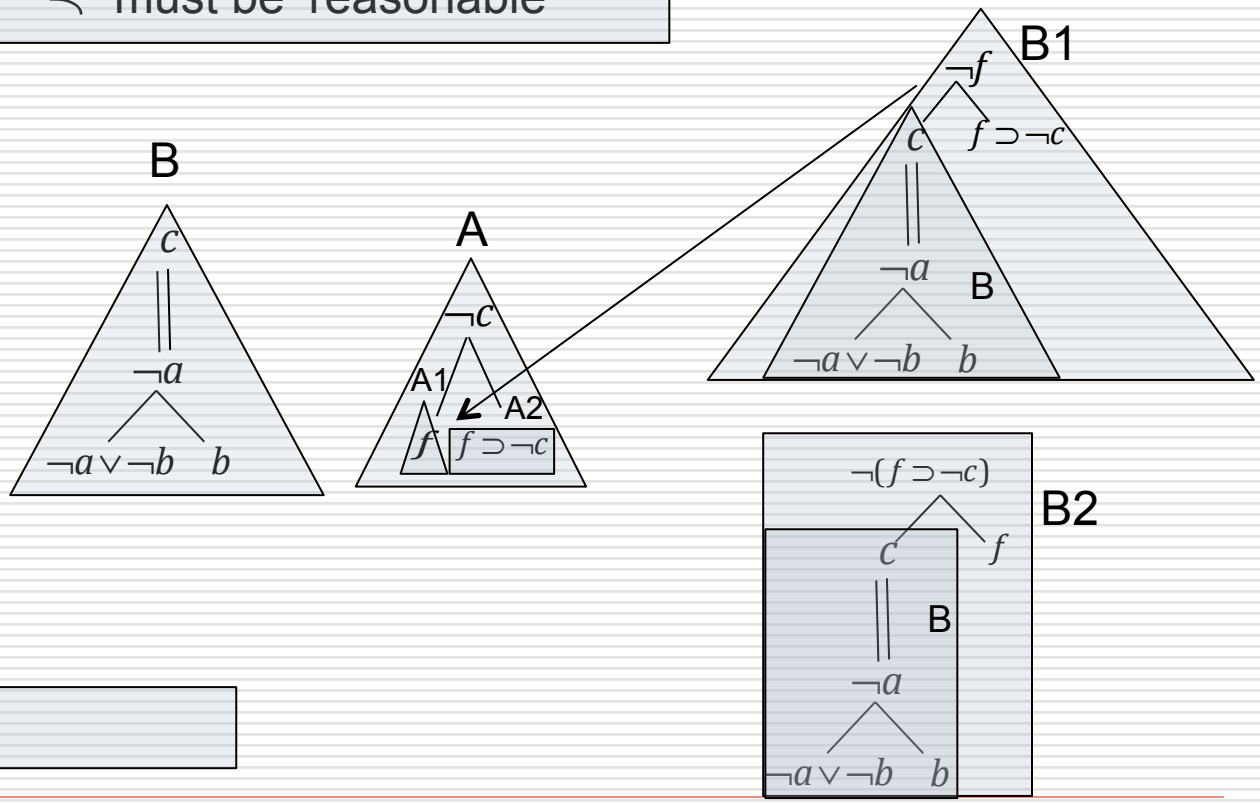


Either B1 $\not\prec$ A1 or B2 $\not\prec$ A2

Key Guidelines for satisfying Rationality Postulates

4) Preference relation/ordering over arguments must satisfy certain properties

$\prec$ must be 'reasonable'

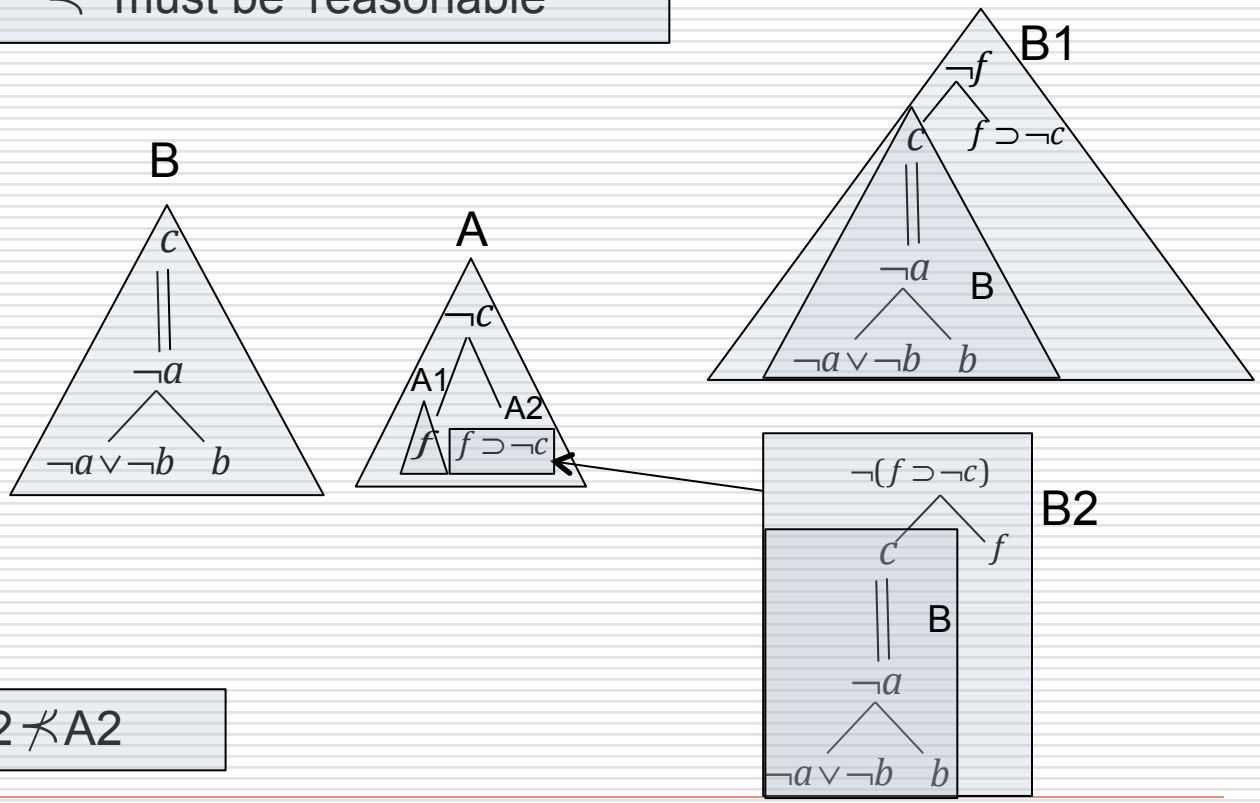


B1 $\not\prec$ A1

Key Guidelines for satisfying Rationality Postulates

4) Preference relation/ordering over arguments must satisfy certain properties

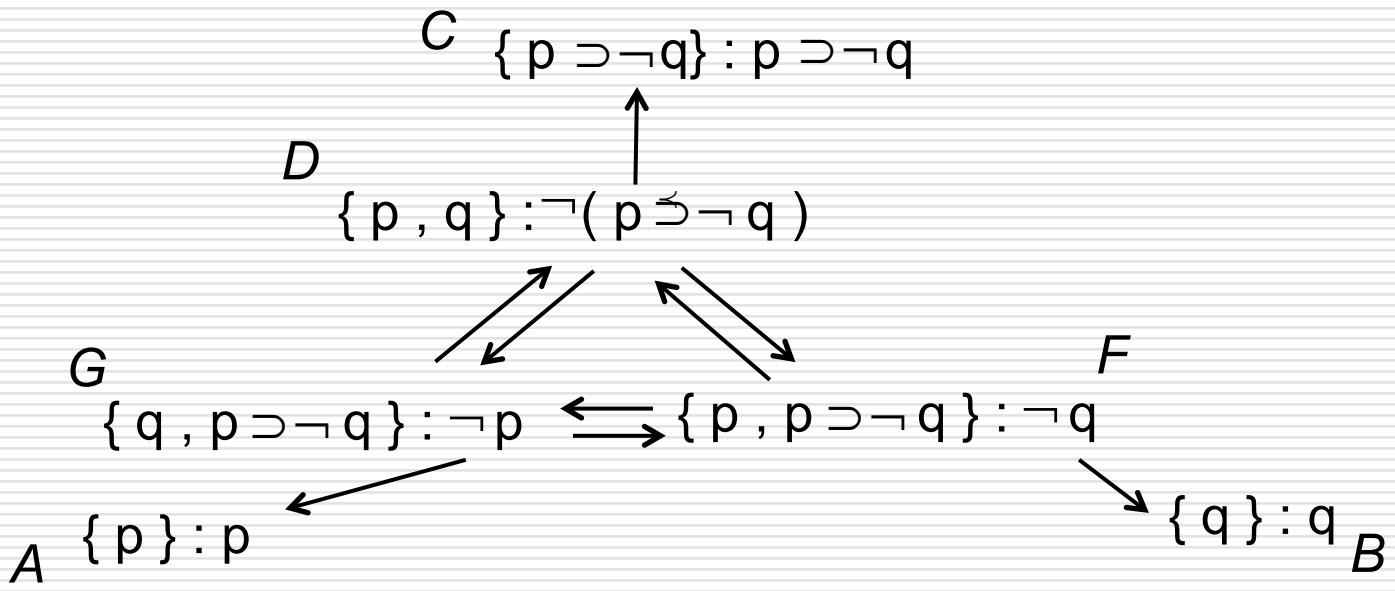
$\prec$ must be 'reasonable'



B2 $\not\prec$ A2

Suppose Preference Ordering not Reasonable

- Suppose $\prec$ is not reasonable: $D \prec C$ *and* $G \prec A$ *and* $F \prec B$



Suppose Preference Ordering not Reasonable

- Suppose $\prec$ is not reasonable: $D \prec C$ *and* $G \prec A$ *and* $F \prec B$

C $\{ p \supset \neg q \} : p \supset \neg q$

D $\{ p , q \} : \neg (p \supset \neg q)$

G $\{ q , p \supset \neg q \} : \neg p$

F $\{ p , p \supset \neg q \} : \neg q$

A $\{ p \} : p$

$\{ q \} : q$ B

Single complete/preferred/grounded extension containing all (inconsistent) arguments

Dialectical ASPIC+ : Full Rationality under Resource Bounds

Limitations of ASPIC+ : Consistency

- ***(Indirect) Consistency***: premises, intermediate conclusions and claims of arguments in an extension are mutually consistent

- Shown under two assumptions:

  1) **Logical Omniscience**: $(\mathcal{A}rgs, \mathcal{D}efeats)$ includes ***all*** arguments that can be constructed using strict rules that encode ***all*** deductive inferences

Limitations of ASPIC+ : Consistency satisfied assuming logical omniscience

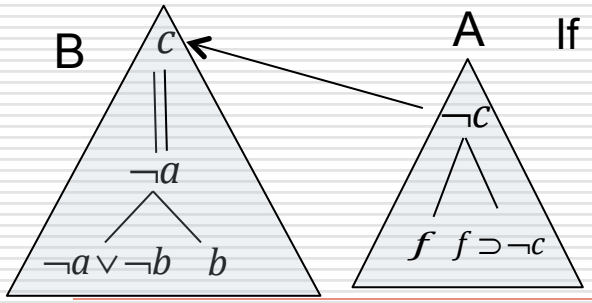
1) Obvious requirement that axiom premises (K_p) are consistent

$$\nexists \Gamma \subseteq K_p \text{ s.t. } \Gamma \rightarrow \alpha \quad \Gamma \rightarrow -\alpha$$

2) Restricted Rebut

Rebut attacks cannot target the conclusions of strict rules; only the conclusions of defeasible rules can be targeted

3) Strict Rules are closed under 'contraposition/transposition'



If $f, f \supset \neg c \rightarrow \neg c \in R_S$ then $c, f \supset \neg c \rightarrow \neg f \in R_S$ and $f, c \rightarrow \neg(f \supset \neg c) \in R_S$

Above satisfied given that $R_S = \{\Gamma \rightarrow \alpha \mid \Gamma \vdash_{CL} \alpha\}$

ASPIC+ argumentative characterisation of Preferred Subtheories

Given (Δ,>) :

$R_d = \emptyset$

$R_s = \{\alpha_1,...,\alpha_n \rightarrow \beta \mid \alpha_1,...,\alpha_n \vdash_{CL} \beta\}$

$K_n = \emptyset$

$K_p = \Delta$

$\mathcal{A}rgs = \{\Gamma \rightarrow \beta \mid \Gamma$ is a *consistent subset minimal* subset of Δ that classically entails $\beta\}$

Hence $\mathcal{A}rgs$ is in general infinite

Clearly this is not practical for resource bounded real-world agents

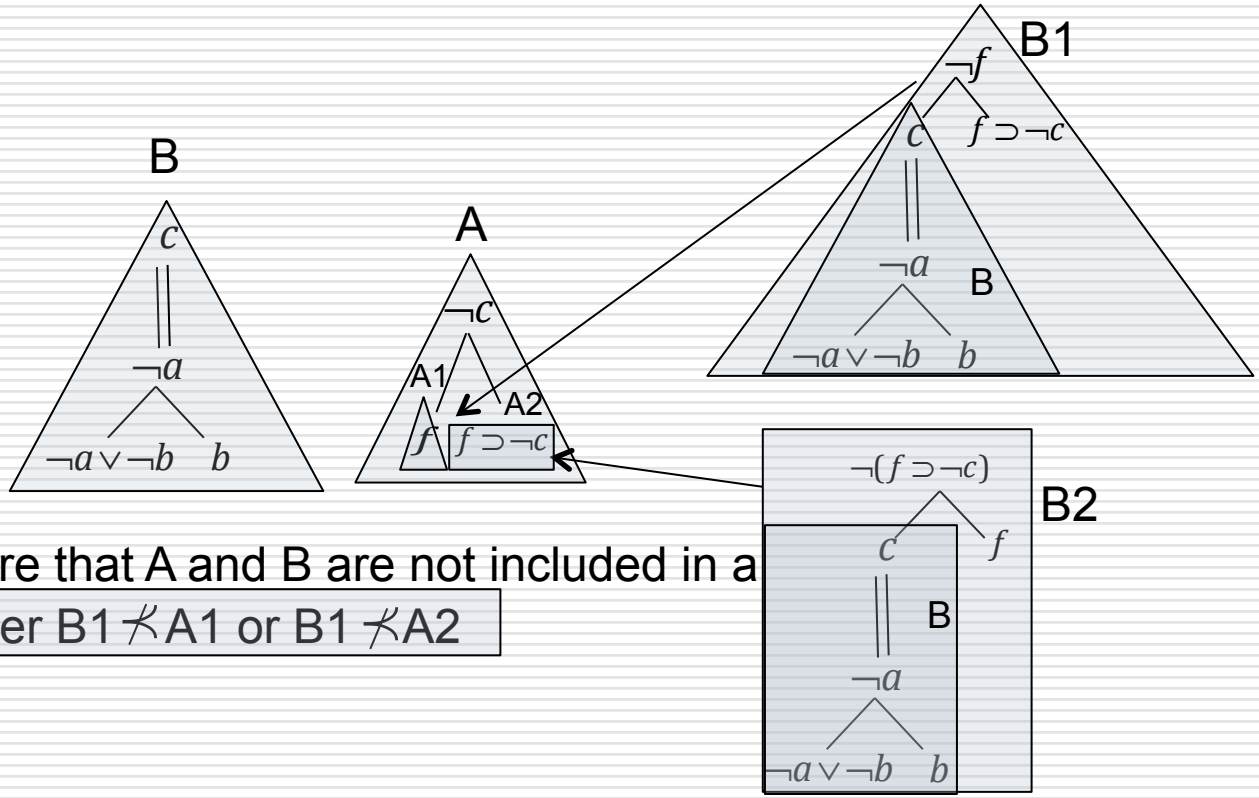
Limitations of ASPIC+ : Consistency

- ***Consistency***: premises, intermediate conclusions and claims of arguments in an extension are mutually consistent

- Shown under two assumptions: Logical Omniscience and

2) **'Reasonable' Preference Relations** over arguments must satisfy certain properties

Limitations of ASPIC+ : Consistency



Suppose A $\prec$ B. To ensure that A and B are not included in a complete extension either B1 $\not\prec$ A1 or B1 $\not\prec$ A2

But surely real-world agents should immediately be able to recognise that A and B cannot be jointly acceptable (i.e., be included in a complete extension)

Limitations of ASPIC+ : Non-contamination

- ***Non-contamination*** postulates proposed some time after consistency and closure postulates *

  Argumentation defined inferences from KB and inference rules R are not invalidated when adding premises and rules *syntactically disjoint* from KB and R

- Satisfied *only by ASPIC+ formalisations of classical logic argumentation (e.g. Preferred Subtheories)* **and only if arguments' premises checked for *consistency and subset minimality***

$\mathcal{A}rgs = \{\Gamma \rightarrow \beta \mid \Gamma$ is a *consistent subset minimal* subset of Δ that classically entails $\beta\}$

- M. Caminada, W. Carnielli, and P. Dunne. Semi-stable semantics. *Journal of Logic and Computation*, 22(5):1207–1254, 2012.

Limitations of ASPIC+ : Non-contamination

- Suppose consistency check not implemented

KB1 = {s} and so ({s},s) is in single grounded extension. Now add syntactically disjoint p and ¬ p to KB

⇨

KB2 = {s,p,¬ p} and now ({p,¬ p},¬ s) defeats ({s},s) which is now no longer in grounded extension !

(Grounded extension is non-empty only if there is at least one undefeated argument)

- Dropping subset minimality check on premises may also result in contamination (as will be illustrated later)

Limitations of ASPIC+ : Non-contamination

- Logical Omniscience and subset minimality/consistency checks on arguments' premises are not feasible for real world resource bounded agents

- Until recently no solution to contamination problem for full ASPIC+ framework (e.g. when arguments incorporate defeasible inference rules)

Although solutions for fragments of ASPIC+ are provided (see work of J. Heyninck, C. Straßer and A. Borg [1,2])

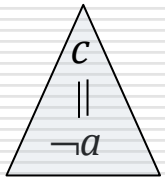
1 J. Heyninck and C. Straßer. *A fully rational argumentation system for pre- ordered defeasible rules*. In *AAMAS'19*

2. A, Borg and C. Straßer. *Relevance in structured argumentation*. In *IJCAI-18*,

Limitations of ASPIC+ : Non-contamination

- Eg

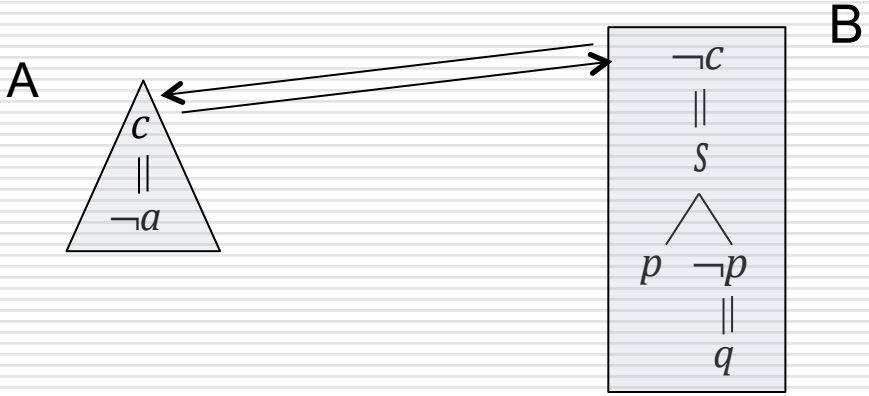
Suppose ASPIC+ theory *T* consisting of single premise $\neg a$ the defeasible inference rules $\neg a \Rightarrow c$ and $s \Rightarrow \neg c$ and let strict rules encode classical logic $R_S = \{\Gamma \rightarrow \alpha \mid \Gamma \vdash_{CL} \alpha\}$

A = 

is in the single grounded extension

- Suppose we add the **syntactically disjoint** p, q and $q \Rightarrow \neg p$ to the theory *T*

Limitations of ASPIC+ : Non-contamination



We then obtain an additional argument B that symmetrically attacks (defeats A)

Now A is no longer in the grounded extension !

Claim (inference *c*) has been invalidated upon adding syntactically disjoint information !

What is required …

- A framework for dialectical formalisations of non-monotonic reasoning for use by resource bounded agents reasoning individually and via dialogue, that:

  1) Drops computationally expensive consistency and subset minimality checks on arguments

  2) Drops assumption of logical omniscience (not all arguments defined by a theory are assumed to be constructed and included in Dung AF)

  3) Drops assumptions on preference relations required to maintain consistency

  4) While still being **fully** rational (satisfying *consistency* **and** *non-contamination*)

A *Dialectical* formulation of ASPIC+ that is fully rational under resource bounds

Joint Work with Marcello D'Agostino, Dept. of Philosophy Milan

M. D'Agostino and S.Modgil *Classical Logic, Argument and Dialectic*. In *Artificial Intelligence* (AIJ). 262, 15 - 51, 2018.

M. D'Agostino and S.Modgil. *A Study of Argumentative Characterisations of Preferred Subtheories* In: 27th Int. Joint Conference on Artificial Intelligence (IJCAI-ECAI-18), 1788-1794, 2018

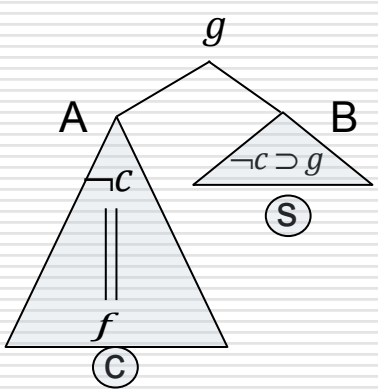
M. D'Agostino and S.Modgil. *A Fully Rational Account of Structured Argumentation Under Resource Bounds*. In: 29th International Joint Conference on Artificial Intelligence (IJCAI'20). 2020.

A Dialectical Ontology for Arguments

- The solution is to define an ontology and evaluation of arguments that accounts for their *dialectical* use
- In practice, arguments are of the following form :

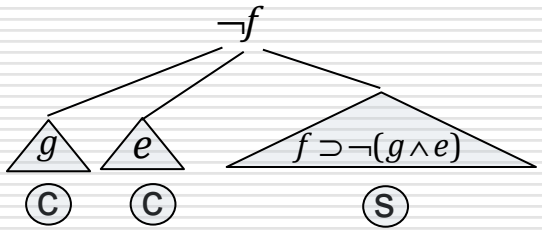
Given that I am committed to the claims Δ and supposing for the sake of argument your commitment to the claims Γ, it then necessarily (deductively) follows that α

So an argument is now a *triple* (*Δ, Γ, α*) – **no subset minimality or consistency checks**
Δ are the *commitments* and *Γ* the *suppositions*



({*A*}, {*B*}, *g*)

$(\{ g \ , \ e \}, \{ f \supset \neg(g \wedge e) \ \}, \neg f \ )$



Dialectical Defeat and Defense

- An acceptable set E defends its arguments against all defeats

$$E = \{ X_1 = (\Phi_1, \Sigma_1, \beta_1) \quad X_2 = (\Phi_2, \Sigma_2, \beta_2) \quad X_3 = (\Phi_3, \Sigma_3, \beta_3) \}$$

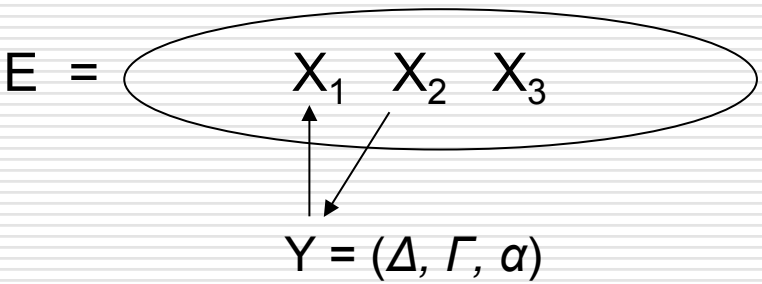
Y

- Y = (Δ, Γ, α) *dialectically defeats* $X_1 = (\Phi_1, \Sigma_1, \beta_1)$ if Y attacks some argument X' in the commitments Φ_1 of X_1 and Y $\not<$ X'

  and suppositions Γ of Y are a subset of the commitments of $X_{1\text{-}3} = \Phi_1 \cup \Phi_2 \cup \Phi_3$

Intuitively, given that I commit to Δ and supposing for the sake of argument your commitments Γ in E, then Y is a counter-argument to X_1

Dialectical Defeat and Defense

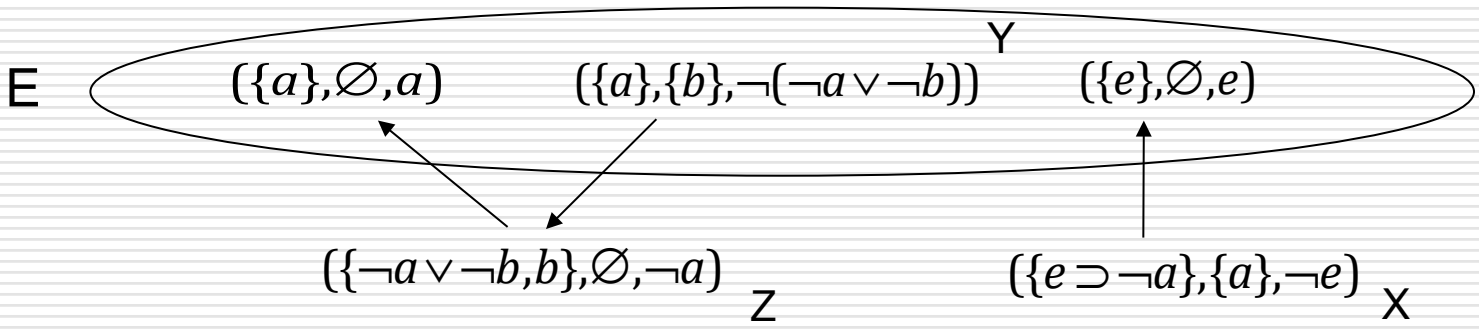


- $X_2 = (\Phi_2, \Sigma_2, \beta_2)$ counter-argues Y (and so defends X_1) if X_2 attacks some argument Y' in the commitments *Δ of Y,* and $X_2 \not\leftarrow$ Y'

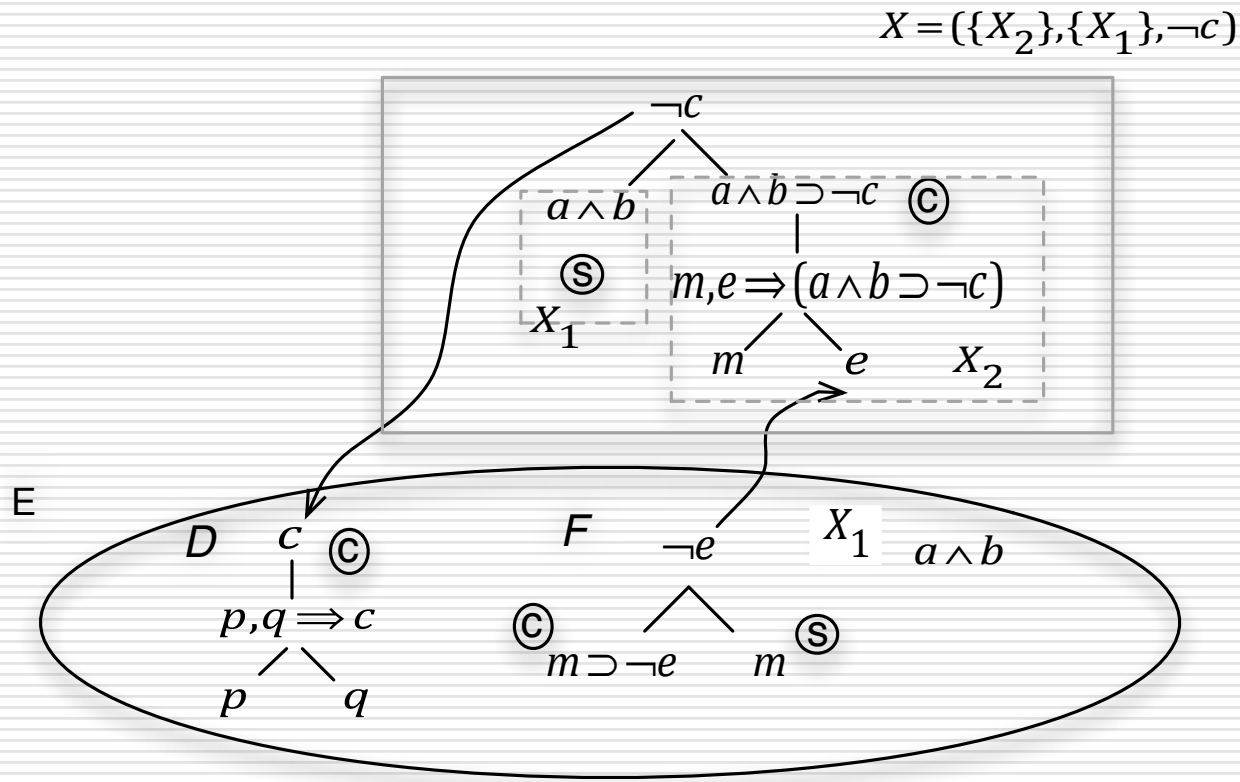
  and the suppositions Σ_2 of X_2 are a subset of the commitments Δ of Y

Intuitively, given my premises Φ_2 and supposing for the sake of argument Σ_2 that you've committed to (in Y), then X_2 is a counter-argument to Y

Classical Logic Example



Example with defeasible inference rules



Dialectical Demonstrations of Inconsistency

- Preferences over dialectical arguments are used in the usual way to determine the success of attacks as defeats, except that

attacks from falsum arguments $(\varnothing,\Delta,\perp)$ ***always succeed as defeats (independently of preferences)***

- Arguments of the form $(\varnothing,\Delta,\perp)$ cannot be defeated since they have empty commitments – they are said to be *unassailable*

Consistency under standard ASPIC+ formalisation of Classical Logic Argumentation

Logical omniscience and *conditions on preference relations* assumed as sufficient conditions to guarantee consistency

E.g., to ensure that A, E, F cannot coexist in an extension ...

$A = \{p \rightarrow \neg q\}: p \rightarrow \neg q$ $\quad E = \{p\}: p \quad F = \{q\}: q$

$B = \{p,q\}: \neg(p \rightarrow \neg q) \quad C = \{q, p \rightarrow \neg q\}: \neg p \quad D = \{p, p \rightarrow \neg q\}: \neg q$

need to assume $B, C, D \in Args$ **and** that either $B \not\prec A$ or $C \not\prec E$ or $D \not\prec F$

Satisfying Consistency in Dialectical Formalisation of ASPIC+ (Classical Logic Example)

If resources suffice to 1) recognise inconsistency amongst premises in A,E and F via construction of at least **one** argument with a conflicting claim e.g
$D = \{p, p \rightarrow \neg q\}:\neg q$

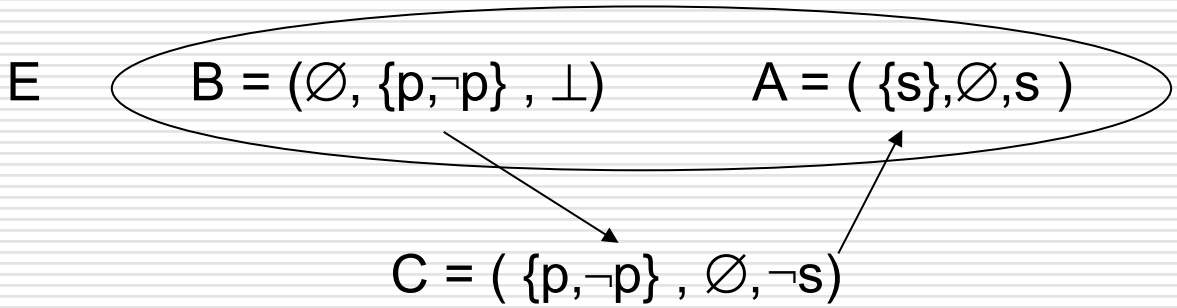
and 2) combine premises of these arguments to obtain unassailable X , then X defeats each argument whose premises contribute to the inconsistency

Consistency postulates satisfied ***independently of preferences*** and without assuming construction of B, C and D (= arguments closed under contraposition/transposition)

$A = (\{p \rightarrow \neg q\}, \{\}, p \rightarrow \neg q)$ $E = (\{p\},\{\},p)$ $F = (\{q\},\{\},q)$

$X = (\{\}, \{p, p \rightarrow \neg q, \neg q\} : \bot)$

Satisfying Non-contamination in Dialectical Formalisation of ASPIC+ (Classical Logic Example)



Despite dropping consistency checks on arguments' premises, *explosivity* does not result in contamination:

A = ({s},∅,s) is in the grounded extension since B defeats C (independently of preferences) and so defends A, and B itself cannot be defeated

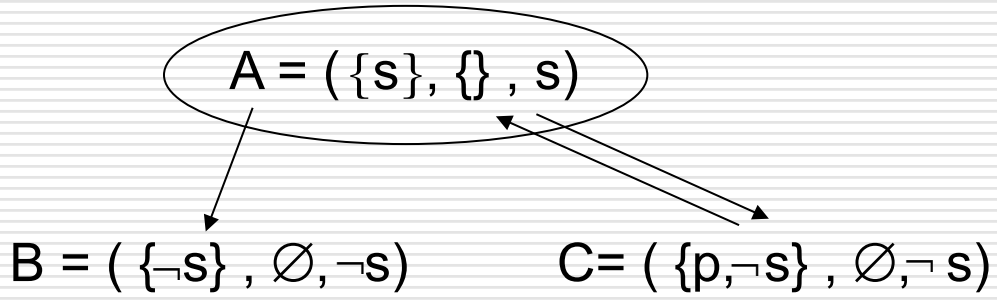
Satisfying Non-contamination in ASPIC+
The problem of relevance

$A = (\{s\}, \{\}, s)$

$B = (\{\neg s\}, \emptyset, \neg s)$

$B \prec A$ and so B does not defeat A and A is in the grounded extension

Satisfying Non-contamination in ASPIC+ : The problem of relevance

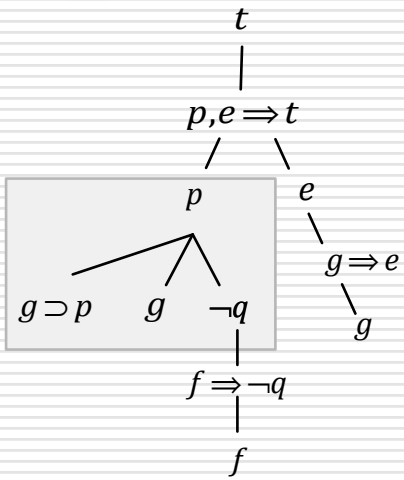


$B < A$ but $C \not< A$ and so C defeats A and A is not in the grounded extension

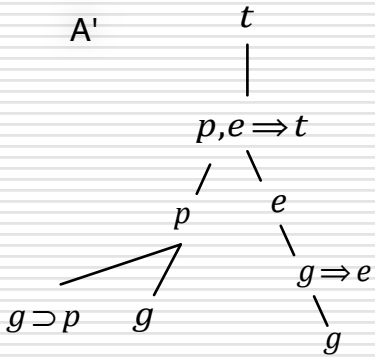
- Checking subset minimality is not a computationally feasible means of enforcing relevance
- We require a notion of relevance that can be enforced *proof theoretically* which means a proof system for classical logic that does not allow one to infer $\neg s$ from p and $\neg s$

Proof theoretic exclusion of arguments that are contaminated dues to non-explosive redundant components

A



A'

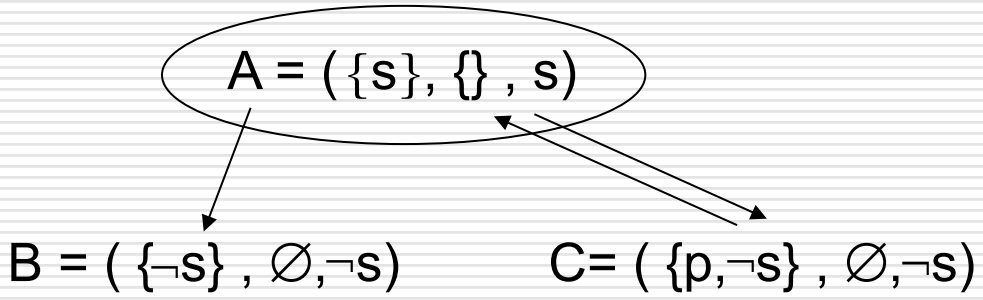


Redundancy due to non-relevant deductive inference can be excluded proof theoretically e.g., use of *Intelim* classical natural deduction system in

M. D'Agostino, D. Gabbay and S.Modgil *Normality, non-contamination and logical depth in classical natural deduction. In: Studia Logica,* pp 1–67 *Feb, 2019.*

will not license redundant inference of *p* from: $g, g \supset p, \neg q$.
Hence only non redundant argument (A') can be constructed

Satisfying Non-contamination in ASPIC+ (Classical Logic Example)



$B < A$ but $C \not< A$ and so C defeats A and A is not in the grounded extension

- Either C is excluded proof theoretically (e.g. through use of *Intelim* natural deduction system)

or

- Preferences must be such that arguments are not strengthened when adding syntactically disjoint premises/rules. Hence $B < A$ implies $C < A$ and so .....

Satisfying Non-contamination in ASPIC+ (Classical Logic Example)

$A = (\{s\}, \{\} , s)$

$B = (\{\neg s\} , \emptyset, \neg s)$

$C = (\{p, \neg s\} , \emptyset, \neg s)$

C **does not now defeat A** and A **is** in the grounded extension

A Fully Rational ASPIC+ For Resource Bounded Agents

Let ($\mathcal{A}rgs$,$\mathcal{D}efeats$) be defined by ASPIC+ theory Δ = (*KB,R*), where $\mathcal{A}rgs$ is ***any*** subset of the dialectical arguments defined by Δ such that

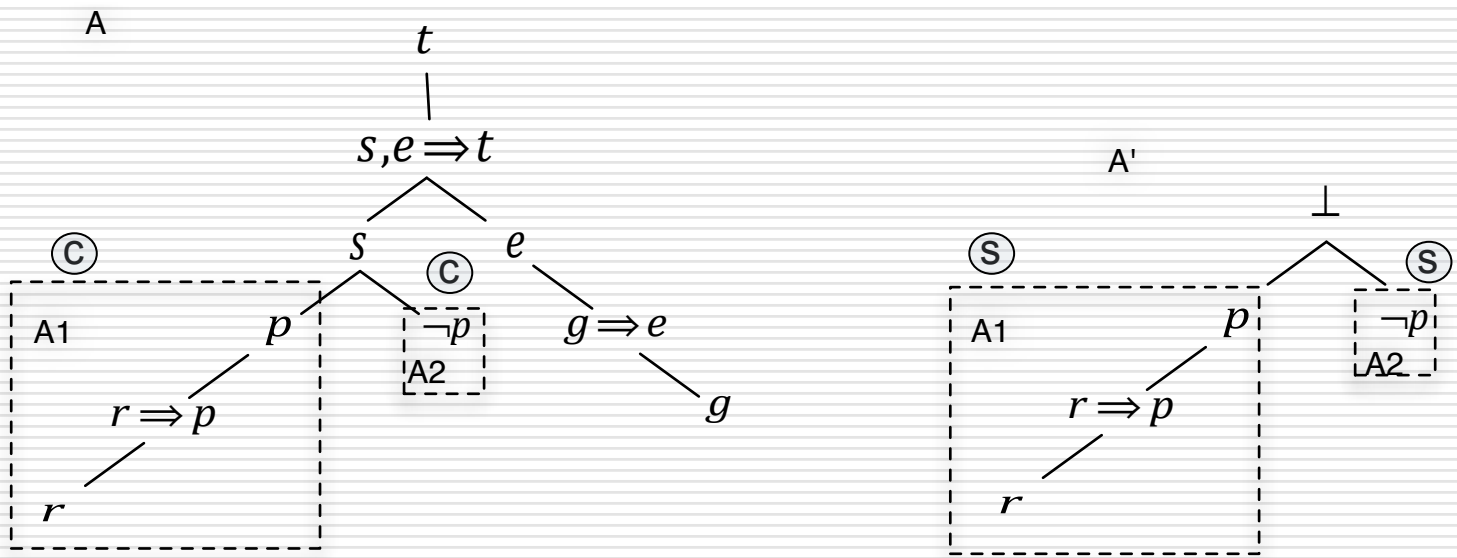
1) If $\boldsymbol{\alpha}$ is a premise in KB then $(\{\boldsymbol{\alpha}\},\emptyset,\boldsymbol{\alpha}) \in \mathcal{A}rgs$

2) If $(\Delta,\emptyset,\boldsymbol{\alpha})$ and$(\Gamma,\emptyset,\text{-}\boldsymbol{\alpha}) \in \mathcal{A}rgs$ then $(\Delta \cup \Gamma,\emptyset,\perp)$ and so $(\emptyset,\Delta \cup \Gamma,\perp) \in \mathcal{A}rgs$

3) If $(\Delta \cup \Gamma,\emptyset,\boldsymbol{\alpha}) \in \mathcal{A}rgs$ and Δ syntactically disjoint from $\Gamma \cup \{\boldsymbol{\alpha}\}$ then

i) if redundant arguments are proof theoretically excluded (i.e., $\Gamma = \emptyset$) then $(\Delta,\emptyset,\perp) \in \mathcal{A}rgs$

Excluding arguments that are contaminated due to explosivity



$A \in \mathcal{A}rgs$ implies $A' \in \mathcal{A}rgs$

A Fully Rational ASPIC+ For Resource Bounded Agents

Let ($\mathcal{Args}$,$\mathcal{Defeats}$) be defined by ASPIC+ theory Δ = (*KB,R*), where $\mathcal{Args}$ is ***any*** subset of the dialectical arguments defined by Δ such that

1) If **α** is a premise in KB then ({**α**},{},**α**) ∈ $\mathcal{Args}$

2) If (Δ,{},**α**) and(Δ',{},-**α**) ∈ $\mathcal{Args}$ then (Δ∪Δ',{},⊥) and so ({},Δ∪Δ', ⊥) ∈ $\mathcal{Args}$

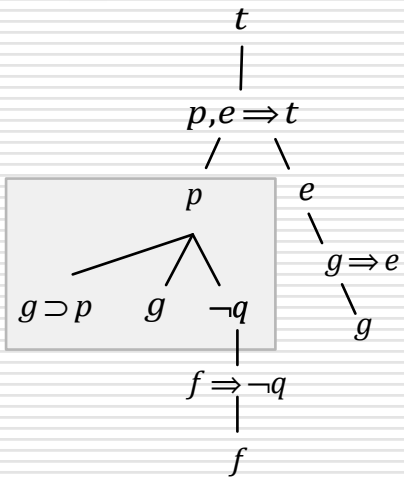
3) If (Δ∪Γ,∅,**α**) ∈ *Args* and Δ syntactically disjoint from Γ∪{**α**} then

i) if redundant arguments are proof theoretically excluded (i.e., Γ = {}) then (Δ,{},⊥) ∈ $\mathcal{Args}$

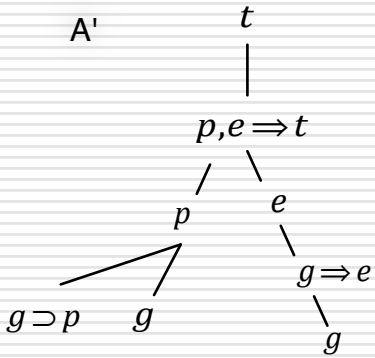
ii) else (Δ,{},⊥) ∈ $\mathcal{Args}$ or (Γ,{},**α**) ∈ $\mathcal{Args}$ and (Δ∪Γ,∅,**α**) and (Γ,{},**α**) are of the same strength

If Proof Theory does not exclude redundant arguments

A



A'



$A \in \mathcal{A}rgs$ implies $A' \in \mathcal{A}rgs$ and A and A' are equally strong

A Fully Rational ASPIC+ For Resource Bounded Agents

Let ($\mathcal{Args}$,$\mathcal{Defeats}$) be defined by ASPIC+ theory Δ = (*KB,R*), where $\mathcal{Args}$ is ***any*** subset of the dialectical arguments defined by Δ such that

1) If $\boldsymbol{\alpha}$ is a premise in KB then $(\{\boldsymbol{\alpha}\},\{\},\boldsymbol{\alpha}) \in \mathcal{Args}$

2) If $(\Delta,\{\},\boldsymbol{\alpha})$ and$(\Delta',\{\},-\boldsymbol{\alpha}) \in \mathcal{Args}$ then $(\Delta \cup \Delta',\{\},\perp)$ and so $(\{\},\Delta \cup \Delta',\ \perp) \in \mathcal{Args}$

3) If $(\Delta \cup \Gamma,\varnothing,\boldsymbol{\alpha}) \in$ *Args* and Δ syntactically disjoint from $\Gamma \cup \{\boldsymbol{\alpha}\}$ then

i) if redundant arguments are proof theoretically excluded (i.e., $\Gamma = \{\}$) then $(\Delta,\{\},\perp) \in \mathcal{Args}$

ii) else $(\Delta,\{\},\perp) \in \mathcal{Args}$ or $(\Gamma,\{\},\boldsymbol{\alpha}) \in \mathcal{Args}$ and $(\Delta \cup \Gamma,\varnothing,\boldsymbol{\alpha})$ and $(\Gamma,\{\},\boldsymbol{\alpha})$ are of the same strength

Then all rationality postulates are satisfied

Conclusions

- The ASPIC+ structured argumentation framework enables argumentative characterisations of non-monotonic reasoning for use by single agents and multi-agent distributed non-monotonic reasoning
- These dialogical formalisations of reasoning are particularly important for scaffolding human-human and human-AI reasoning
- ASPIC+ satisfies consistency under conditions that are not practical for real-world agents with bounded computational/cognitive resources, and does not satisfy non-contamination
- Dialectical ASPIC+ is **fully** rational under resource bounds

Extras on Dialectical ASPIC+

- Grounded semantics applied to AFs instantiated by *Dialectical* ASPIC+ theories are less sceptical than Grounded semantics applied to AFs instantiated by ASPIC+ theories

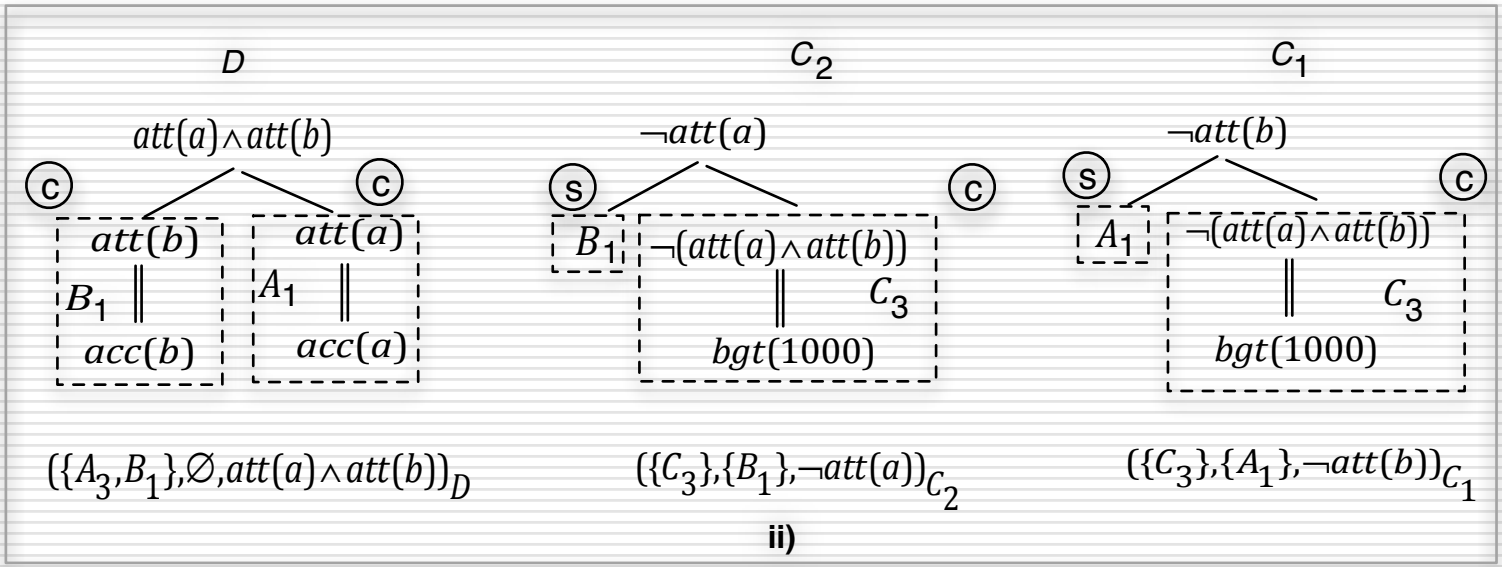
  (see Section 4 https://www.ijcai.org/Proceedings/2018/0247.pdf)

- Dialectical ASPIC+ solves foreign commitment problem (see following slides)

- Suppose proof system licenses construction of redundant arguments – can we then avoid requirement that preference relation does not change strength of arguments upon adding syntactically disjoint rules/ premises ? Yes, if we extend the range of dialectical moves (see following slides)

Other Features of Dialectical ASPIC+ : Solving the Foreign Commitment Problem [1]

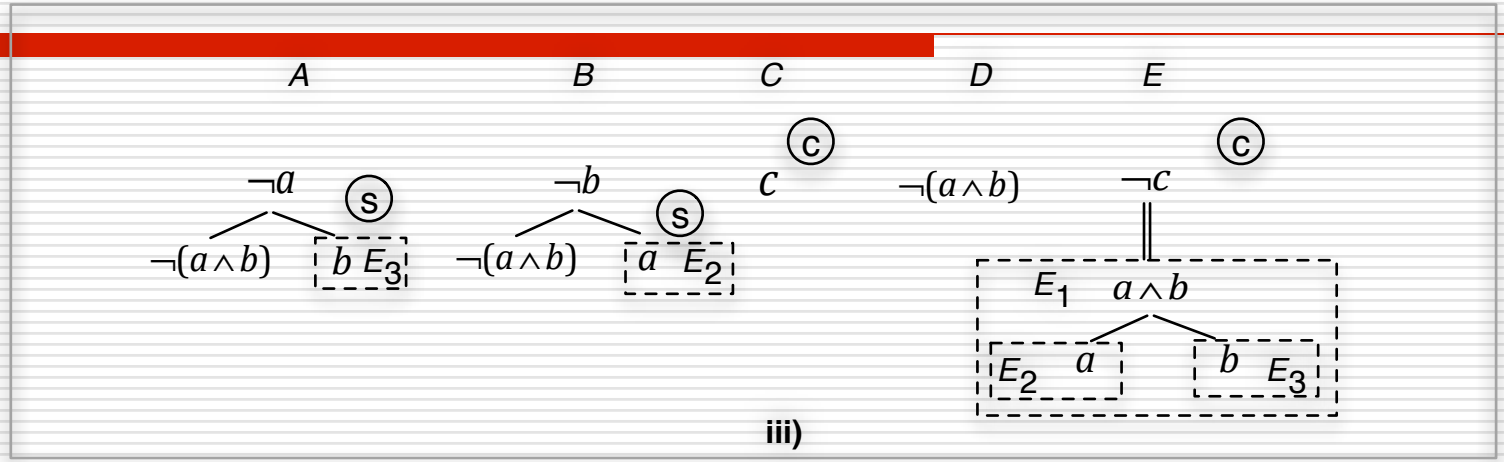
att(x) = attend conference x , *acc(x)* = paper accepted at x, *bgt(1000)* = budget of £1000



C_2 and C_1 can attack *D* claiming at attendance at both conferences, without having to commit to (and instead suppose only for the sake of argument) attendance at *b*, respectively *a*

1. M. Caminada, S.Modgil, and N. Oren. Preferences and unrestricted rebut. In COMMA 2014, pages 209–220. 2014.

Other Features of Dialectical ASPIC+ : Less sceptical than ASPIC+ under grounded semantics



E = "if attend *a* and *b* then cannot attend *c*" .Subsequently learn scheduling constraint: *D* = axiom premise – "cannot attend conferences *a* and *b*"

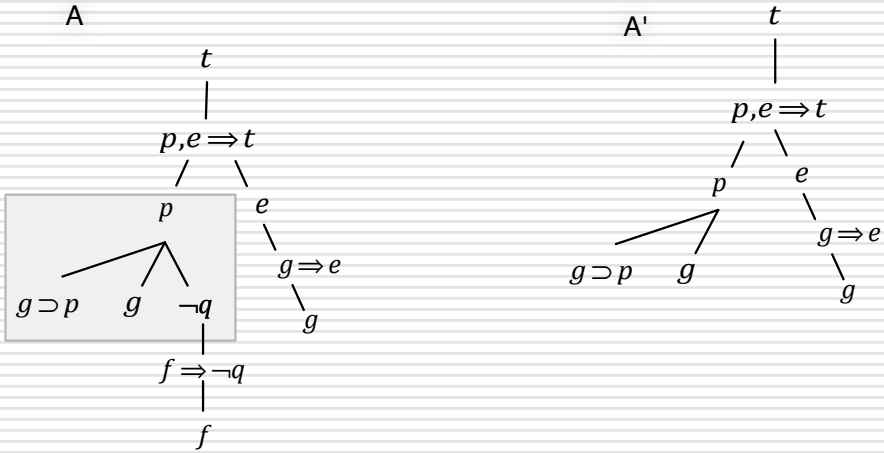
One then wants *C* in grounded extension, but defending *C* against *E* would in standard ASPIC+ require either ($\{\neg(a \wedge b), b\}, \neg a$) or ($\{\neg(a \wedge b), a\}, \neg b$) in grounded extension G_E. However these symmetrically defeating arguments are not in G_E

However dialectical arguments *A* and *B* (which respectively suppose b/a and so cannot be defeated) are in grounded extension. Both defeat *E* and so defend C which is now in the grounded extension

Other Features of Dialectical ASPIC+ : Scope for extending range of dialectical moves

For example (∅, {A'}, t) defeats ({A}, ∅, t) independently of preferences

Assuming no commitments of mine, and supposing a subset of the rules and premises in your argument A, one can construct a non-redundant argument A' claiming t, without making use of irrelevant (syntactically disjoint) premises and rules



Then even if proof theory licenses construction of A. non-contamination now satisfied without needing to assume that A and A' are of the same strength