Argumentation Logic

Antonis Kakas **Francesca Toni**

Paolo Mancarella

Department of Computer Science
University of Cyprus

Department of Computing
Imperial College London

Dipartimento di Informatica
Universita di Pisa

16th July, 2013

Panhellenic Logic Symposium

Athens

AL Introduction

- A re-examination of "classical" logical reasoning - Propositional Logic- as a Logic of Arguments.
  - Closer to original inception of logic?
  - Closer to Common Sense Human Reasoning?
- Methods: Argumentation Theory from AI and Syllogistic Roots of Logic
  - Natural Acceptability Semantics for Argumentation
  - Re-examine Reductio ad Absurdum in Natural Deduction

Motivation from AI

- **Computing and Artificial Intelligence**
  - **Common Sense Human Reasoning?**
    - **Default Reasoning, e.g. Temporal Persistence**
    - **Reasoning about actions and change**
    - **Knowledge Qualification, e.g. Resolving contradictory information, or Legal Reasoning**
  - **Text Comprehension**
    - **"I am attending the 8th Panhellenic Logic Symposium in Athens in July."**
      - **Elaborative Inferences, e.g. "I will be in Athens sometime in July", "I am an academic/logician"…**
      - **Conflict resolution, e.g. 8th PLS?**
- **Case of "Logic from Computer Science".**

The (traditional) logic side of things

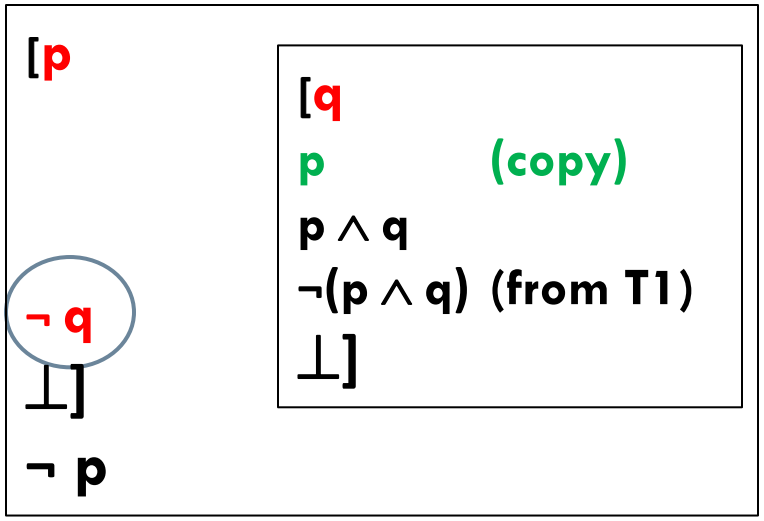
Part 1: "Syllogistic roots" of Logic

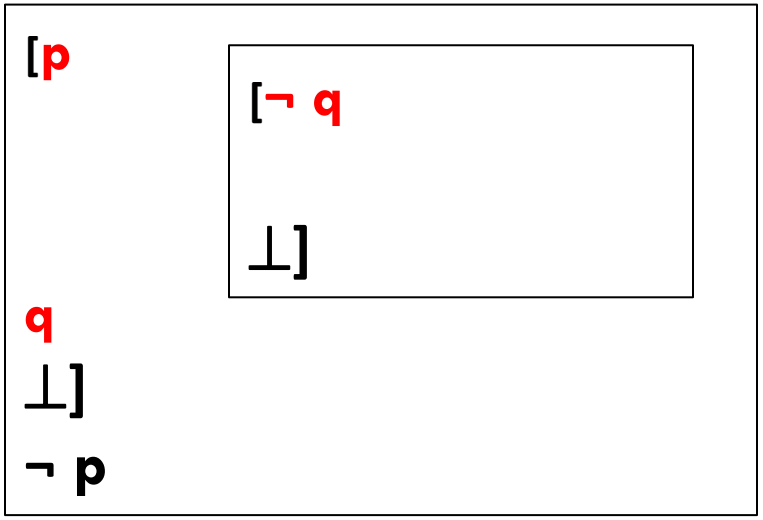
- Consider Propositional Logic (PL) and its Natural Deduction (ND) proof system.
- Separate out the Reductio ad Absurdum (RA) rule (¬I rule) as a different type of proof rule or argument.
  - Is it an argument at all? Is RA an axiomatic part of Logic?
- Call (c.f. Archimedes) the rest of ND, Direct Logic/Proofs, $\vdash_{MRA}$
  - Direct Logic: basic logic underlying Argumentation Logic
- Note that in any RA derivation, [ϕ … … ⊥], we have a direct derivation of the contradiction.

Reductio ad Absurdum in ND

Example 1

- T1 = {¬ (p ∧ q), ¬ q → ⊥}





Note 1: Direct (sub) proofs under $\vdash_{MRA}$: " $\vdash_{ND}$ minus RA"

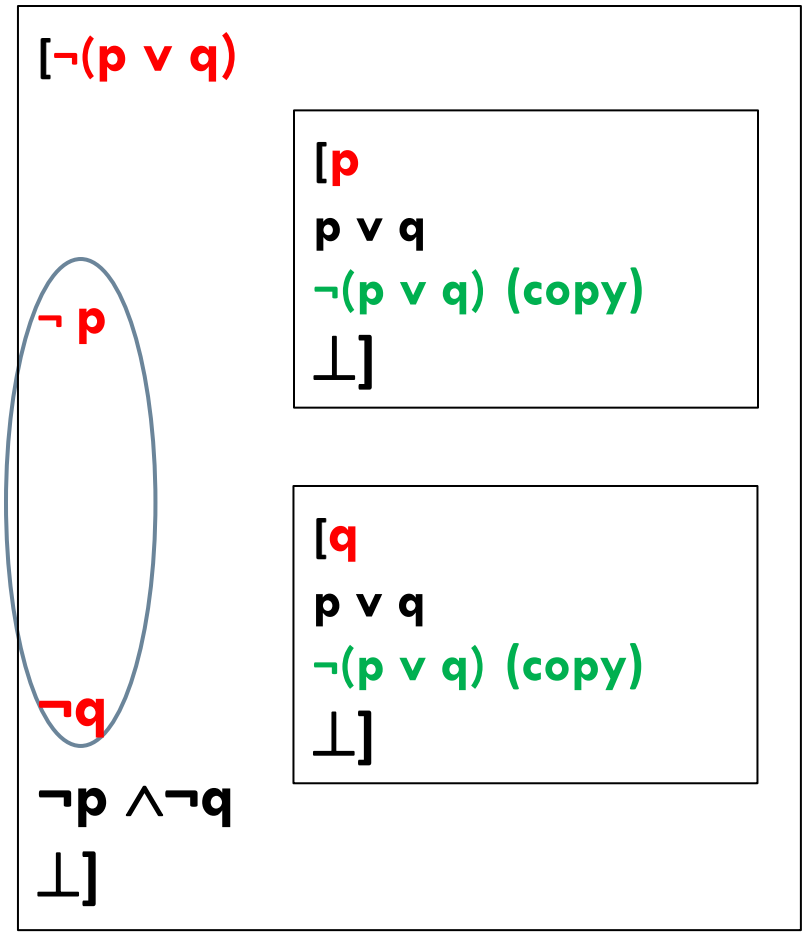
Note 2: Relevance of hypothesis to inconsistency:

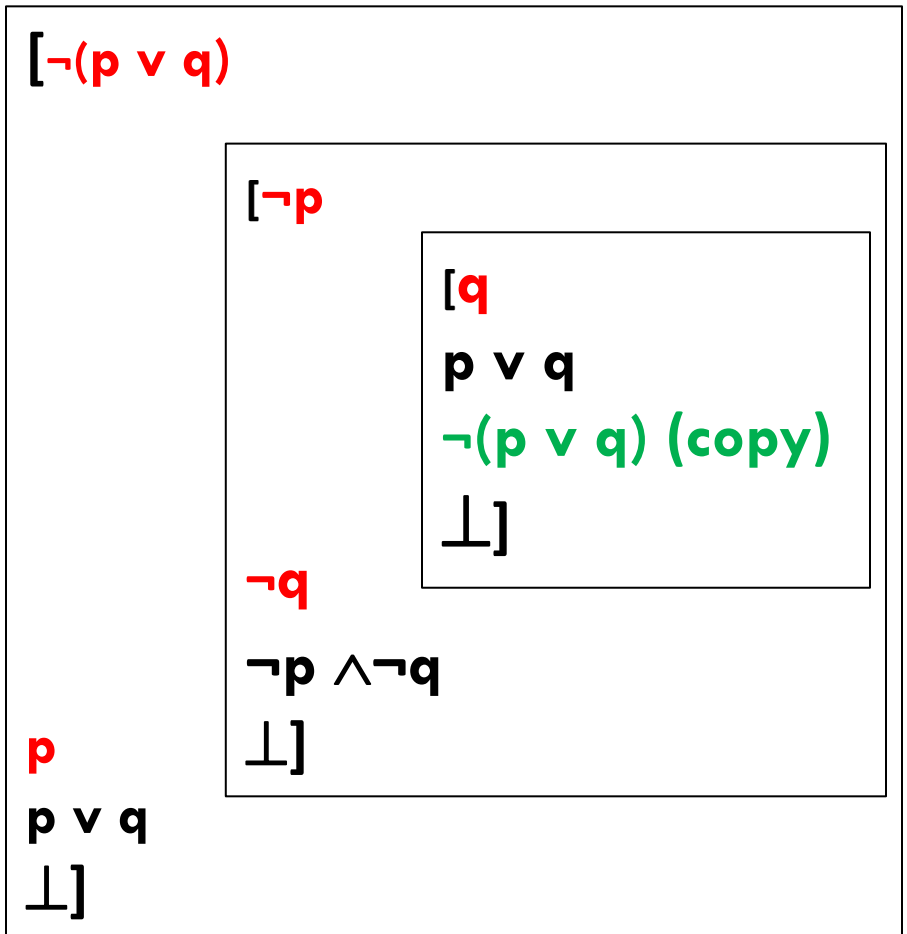
Genuine Absurdity Property

Reductio ad Absurdum in ND

Example 2

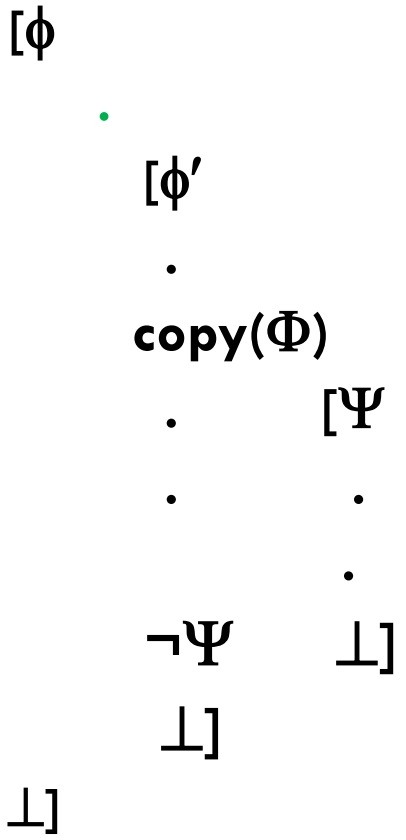
- T2 = {¬ (¬ p ∧ ¬ q)} $\vdash_{ND}$ p ∨ q





Genuine RAND derivations

Genuine RAND derivations:



$\mathbf{T} \cup \Phi \cup \neg\Psi \not\vdash_{MRA} \bot$

ϕ′ is necessary for the direct derivation of ⊥

- **Do Genuine RAND derivations always exist?**

AL equivalent to PL ("restricted" to ¬, ∧)

- **Main Lemma:** For **consistent** theories (in ¬, ∧) if there is a RAND derivation from ϕ then there is a **Genuine** RAND derivation from ϕ.
  - **Proof: Is this result known?**
  - **Hence the Restricted form of RA does not compromise completeness of ND.**
- **Equivalence** through the universality of ¬, ∧.
  - If we interpret V and $\rightarrow$ through their classical equivalence in terms of ¬, ∧ then **AL=PL.**
  - **But this is not necessary (see below part 2).**

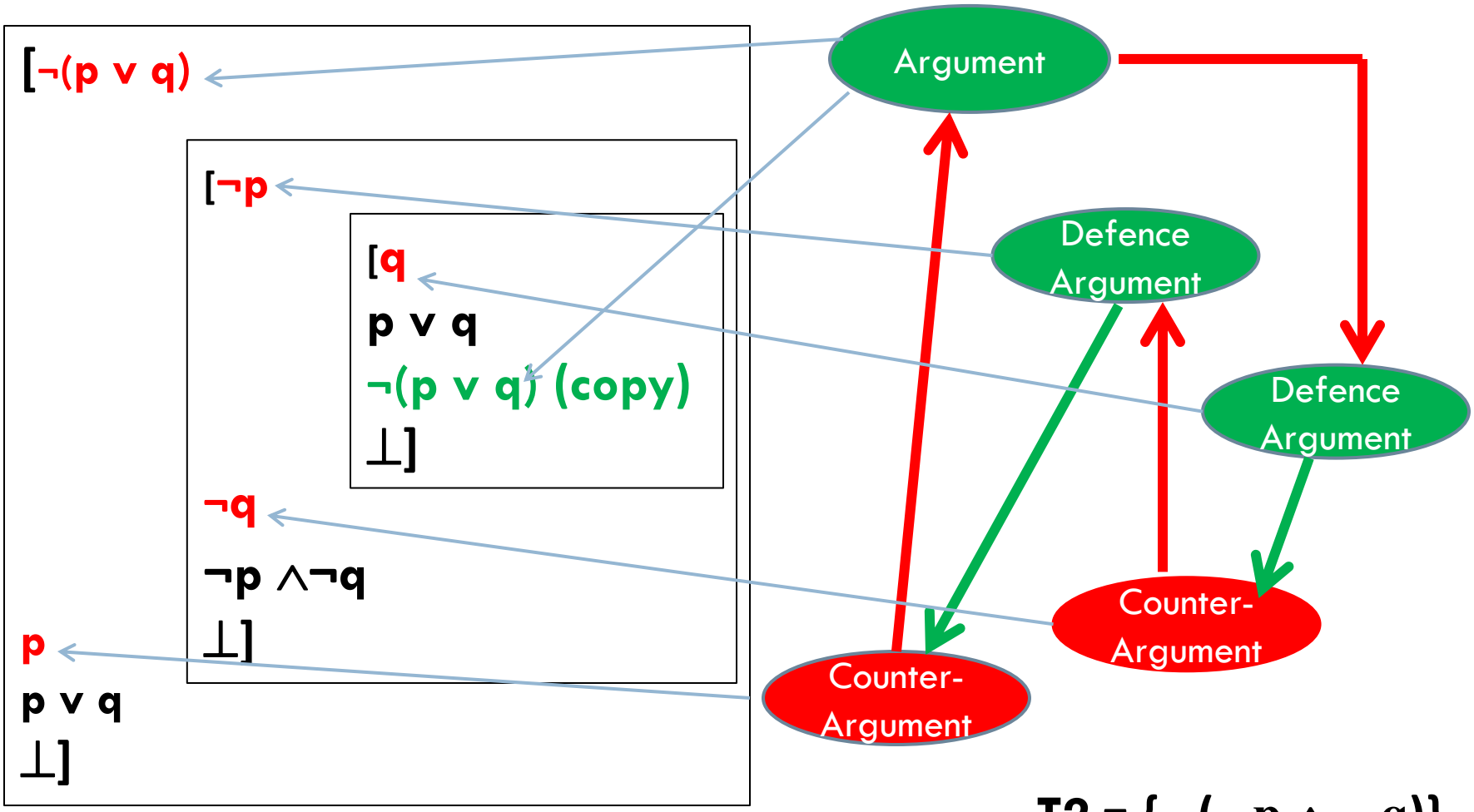
The Argumentation side of things

The other Argumentation side

- Can we see Logic as a Theory of Arguments?
  - How can we do this?
    - Can we have Logical Formulae as Arguments?
    - Entailment through Acceptable Arguments?
  - How can we link this to Classical logic (PL)?
    - Reformulate PL as a Logic of Arguments?

- Can we formulate Natural Deduction with restricted Reductio ad Absurdum as a Logic of Arguments?
  - Using argumentation theory/semantics from AI?
  - Build on the "success" of Argumentation in AI and CSR.

Argumentation Interpretation of Reduction ad Absurdum - Informal



T2 = {¬ (¬ p ∧ ¬ q)}

Central Idea of Argumentation Logic

- **Logical Formulae as Arguments.** Arguments attack each other through a Direct Derivation of inconsistency.

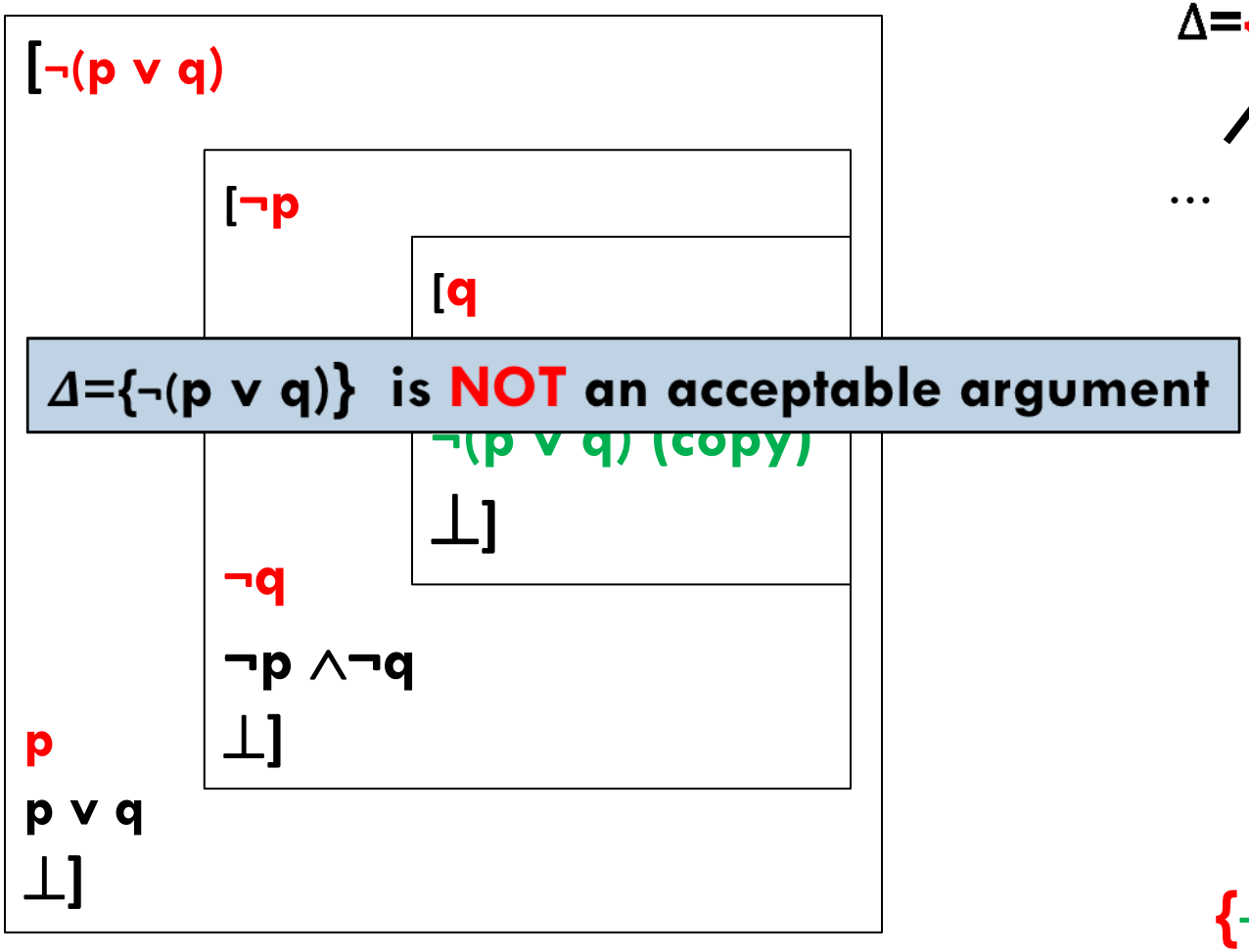
- Argumentation Framework <Args, Att, Def> for Logic:
  - Args: Sets of Propositional Formulae: Δ
    (Direct proofs from Δ and the given theory, T)
  - Att: A attacks Δ : $T \cup \Delta \cup A \vdash_{MRA} \bot$
  - Def: Def $\subseteq$ Att

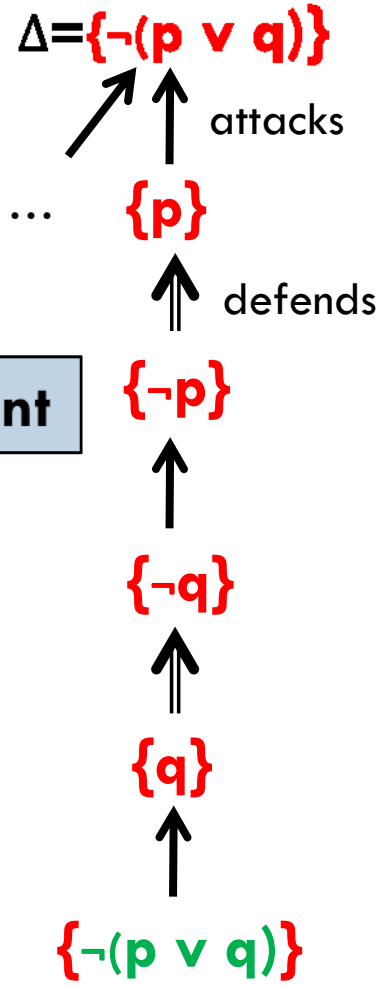
- Recover Reductio ad Absurdum through the semantics of argumentation. BUT WHICH SEMANTICS?

- The problem of Logic ανάγεται to the question: What is a good, or acceptable, argument?

Argumentation Interpretation of Reduction ad Absurdum

T2 = {¬ (¬ p ∧ ¬ q)}





Argumentation in AI - Basics

- **Argumentation Framework:** <Args, Attacks, Defence>

- **Semantics:** Δ is an **admissible** set of arguments iff:
  - **Δ does not attack itself.**
  - **Δ defends against all sets that attack it:**
    - **Δ attacks back A.**

- **Is the Admissible semantics "complete"?**
  - **What if an attack is by itself "no good", e.g. self-attacking?**
  - **Do we still need to explicitly defend/attack it back?**

- **Admissibility semantics => Acceptability semantics**

Acceptability Semantics
Informal Motivation

- **Acceptability:** Follow the "universal" intuition:

> **An argument (or a set of arguments) can be accepted iff all its counter-arguments can be rejected**

- Can we formalize directly this intuition?
  - How are we to understand the "Rejection of Arguments"?
    - As "Can not be Accepted"?
  - An argument can play a role in rejecting its counter-arguments
    - The Acceptance of arguments is a RELATIVE notion.

Acceptability Semantics Definition

- A set Δ is acceptable relative to Δ': $Acc(\Delta,\Delta')$.

  $Acc(\Delta,\Delta')$ iff $\Delta \subseteq \Delta'$, or
  for any A s.t. A attacks Δ:
  there exists D s.t. D defends/attacks back A
  and $acc(D, \Delta' \cup \Delta \cup A)$.

- Acceptability, Acc(-,-), is defined as the least fixed point of a monotonic operator, F_{ACC} ,on the binary relations on sets of arguments.

- Acceptability Semantics: Δ is acceptable iff $Acc(\Delta,\{\})$ holds.

Central use of Acceptability Semantics

- Acceptability of arguments is a relative notion.
- Captures a semantic notion of self-defeating (set of) argument(s):
  - S is self-defeating iff there exists an attacking set, A, against S such that ¬Acc(A, {}) and Acc(A, S) hold.
- Self-defeating S: renders one of its attacks acceptable
  - This is a kind of Reductio ad Absurdum Principle!
- Acceptability deals with (odd) cycles of attacks.
  - Compare with "cyclic reasoning" of Reduction ad Absurdum!

Argumentation Logic

Self-defeat ↔ Reductio ad Absurdum

ϕ Is AL-entailed iff_{def} ACC($\{\phi\},\{\}$) and $\neg$ACC($\{\neg\phi\},\{\}$)

Theorem

$\neg$Acc($\{\phi\},\{\}$) $\Leftrightarrow$ Genuine RAND derivation for ϕ

Corollary (from Lemma)

For consistent T: AL = PL

Argumentation Logic Results (1)

- AL distinguishes two forms of Inconsistency of T
  - Classically inconsistent but directly consistent (under $\vdash_{MRA}$)
    - Violation of rule of «Excluded Middle».
    - For some, ϕ, neither ϕ nor $\neg\phi$ is acceptable: $T = \{ a \rightarrow \bot, \neg a \rightarrow \bot \}$
      - $a \vee \neg a$ not AL-entailed, but $b \vee \neg b$ is AL-entailed
  - Directly inconsistent
    - For some ϕ, T has a direct argument for ϕ and $\neg\phi$: $T = \{ \phi , \neg\phi \}$

AL is a paraconsistent logic.

Example of Directly Consistent: Logical Paradox
"**Not a contradiction but a paradox**"

- **"A barber shaves anyone that does not shave himself"**
  - ¬ ShavesHimself(Person)→ ShavedByBarber(Person)
  - ShavesHimself(Person) → ¬ ShavedByBarber(Person)
- **Self-reference: When Person = barber**
  - ShavedByBarber(barber) → ShavesHimself(barber)
  - ¬ ShavedByBarber(barber) → ¬ ShavesHimself(barber)

Logic Paradox Example in AL

- Classically Inconsistent due to the law of **excluded middle**
  - **SB(P) or ¬ SB(P) , for any person P, even for P=barber.**

- In AL the law of excluded middle for SB(b) does not hold
  - ¬ ACC({SB(b)},{}) **SB(b)** is **non-acceptable**
  - ¬ ACC({¬SB(b)},{}) **¬ SB(b)** is **non-acceptable**

  - The law **(SB(b) v ¬ SB(b))** is **non-acceptable.**

- **AL gives up the law of excluded middle (when needed)!**

Argumentation Logic Results (2)

- For classically consistent theories AL = PL (for the restricted language of ¬ and ∧)

- But we can define AL directly on the whole language of PL with V and →
  - Interpretation of implication in AL differs from PL, e.g.
    - Both a→b and ¬(a→b) are acceptable w.r.t. to T={¬a}

- Can also take different Direct Logic underlying AL.

AL – What does it mean?

- Computing (on the Web) today is "demanding" Common Sense Human Reasoning

- Human oriented Computing: Agency + Human Interaction
  - "I am attending the 8th Panhellenic Logic Symposium in Athens in July."

- "Conjecture:" For Common Sense Reasoning we need to challenge Classical Logic.

Loogle

- QUERY: "I am attending the 8th Panhellenic Logic Symposium in Athens in July. Please suggest places to stay."
  - Data/Information integration over the database/Knowledge base of the Web

- ANSWER: "The Golden Age hotel : this is close to the Music Hall where a concert will take place in its gardens."
  - Personalized, Justified (and persuasive) recommendations

Conclusions

- A reformulation of PL in terms of argumentation under acceptability semantics => Argumentation Logic (AL)
  - AL is a conservative extension of PL into a type of Relevance Para-consistent Logic
  - Only genuine use of Reductio ad Absurdum
  - Implication in AL differs from classical material implication
- Implication is a hybrid of default rule and contrapositive reasoning
  - UNIFY classical and defeasible reasoning under argumentation???
- This questioning of CL by AL is rooted in (a part of) "AI Computing" that needs the automation of Common Sense Human Reasoning
  - Not driven from the needs of strict Mathematical Reasoning but from open Human Reasoning, e.g. Natural Language or "linguistic" reasoning.

AL – What does it mean?

- **Philosophy (of Science):**
  - **Logic Describes vs Logic Captures**
  - **Logic: Language vs Realism**