



Institut de Recherche  
en Informatique de Toulouse



**INRA**  
SCIENCE & IMPACT

## Assessing Arguments with Schemes and Fallacies

---

Pierre Bisquert, **Florence Dupin de Saint-Cyr**, Philippe Besnard

LPNMR, June 3-7, 2019  
Saint Josephs University, Philadelphia

## Why formalize fallacious reasoning?

- Persuade people (or oneself) = crucial in daily life and important topic in Philosophy, Psychology, Linguistics, Logic, AI, MAS, Legal reasoning..
- “correct” persuasion?

### Examples of fallacies

- “Finish your hot-dog, think about the poor, starving children who were not fortunate enough to have any food at all” (*Appeal to emotion*)
- “Temperatures have been rising over the past few centuries, whilst at the same time the numbers of pirates have been decreasing, thus pirates cool the temperature” (*False cause*)
- “How can you tell me I should stop smoking when you still smoke yourself ?” (*Tu quoque*)

## Motivation (2)

- Fallacies : broadly studied: Aristotle, (*Hamblin 1970*), (*Kelley 2013*).

Websites:

<https://yourlogicalfallacyis.com/>

<https://www.logicallyfallacious.com/>



- (*Woods 1994*):
  - “a fallacy is any error in reasoning”
  - “patterns of errors that occur with some frequency”,
  - “a certain surface plausibility ... premises **seem** relevant to the conclusion ... argument **seems** to have strength”

- ▶ Understanding fallacious reasoning: 2 benefits
  1. Helping people to identify and avoid them;
  2. Progressing in understanding **good reasoning**
- ▶ **Generic logical system to automatically assess argument.**
  - is it sound?
  - does it use a fallacious scheme? which one?
- ▶ idea= define a framework mimicking fallacious (and rational) reasoning
  - handle logical knowledge + user defined inference schemes

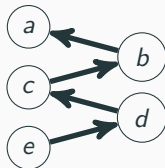
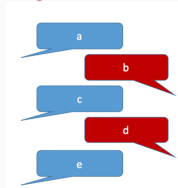
1. Assessing Arguments vs Argumentation
2. Language for reasoning with Logical Knowledge + user defined Inference Schemes
3. To Classical Logic and Beyond
4. (Fallacious + Rational) Schemes and Reasoning

# **Assessing Arguments vs Argumentation**

---

# Assessing argument vs Argumentation

- Two meanings of the word “Argument”:
  - **Argumentation**: interaction about difference of opinion



- **Assessing Argument**:
  - ▶ argument = structure with a claim and its reasons
  - ▶ checking if reasons are good evidence for the claim
- NB: **Argumentation** and **Assessing Arguments** coincide when proof simulated by dialogue PRO/CON

# Representations of an Argument

- Abstract vertex in a graph (*Dung 1995*)
- (*Wigmore 1931*)'s Diagram
- (*Toulmin 1958*): (Claim, Data, Warrant, Backing, Rebuttal) where Claim supported by Data, Warrant provides licence to infer Claim from Data, Backing for the Warrant, Rebuttal condition encapsulates exceptions
- (*Walton and Thomas F. Gordon 2005*): scheme based on critical questions
- ▶ Logic-based argument = (premises, claim)
  - assess argument = check if it is **sound** = valid with true premises



**Language for reasoning with Logical Knowledge +  
user defined Inference Schemes**

---

- $L_0$ : logical knowledge
  - term  $T$ : member of  $\mathcal{X}_0$  (variables) or  $\mathcal{C}_0$  (constants)
  - atom in  $At_0$ :  $p(\vec{T})$  s.t.  $p \in P_0$  (predicates) and  $\vec{T}$  vector of terms
  - variable  $X$  in  $\mathcal{X}_1$ : represent any atom of  $At_0$
  - formula  $\varphi, \psi$  in  $L_0$  ::  $p(\vec{T}); X; \varphi \rightarrow \psi; \neg\varphi$
- $L$ : metalanguage about  $L_0$ 
  - $L$  ::  $licit(\Psi, \varphi); proven(\varphi); sound(\Psi, \varphi)$   
where  $\Psi$  set of formulae of  $L_0$

# Semantics of the Language L (1)

- $K \subseteq L_0$ : knowledge base
- $S \subseteq I \times 2^{L_0} \times L_0$ : recorded schemes of the form  $(id, \Psi, \varphi)$

## Logical Knowledge Base and Recorded Schemes

$K = \{rain, rain \rightarrow take(umbrella)\}$

$S = \{(modPo, \{X, X \rightarrow Y\}, Y)\}$

- **Licit Argument** = obtained by substitution upon a recorded scheme:  
 $K, S \models_L licit(\Psi, \varphi)$  iff  $\exists(id, A, \alpha) \in S$  and  $\exists\sigma : \mathcal{X}_1 \rightarrow At_0$  s.t.

$$(\Psi, \varphi) = (\sigma(A), \sigma(\alpha))$$

## Licit Argument

$K, S \models_L licit(\{rain, rain \rightarrow take(umbrella)\}, take(umbrella))$

( $\sigma = \{rain/X, take(umbrella)/Y\}$  applied to *modPo* scheme)

## Semantics of the Language L (2)

- **Proven Formula** = reached by a sequence of inference schemes from  $K$ :

$K, S \models_L proven(\varphi)$  if<sub>def</sub>  $\varphi \in K$  or  $\exists \Psi \in 2^{L_0}$  s.t.  $K, S \models_L sound(\Psi, \varphi)$

- **Sound Argument** = licit and proven premises:

$K, S \models_L sound(\Psi, \varphi)$  if<sub>def</sub>  $K, S \models_L licit(\Psi, \varphi)$  and  $\forall \psi \in \Psi, K, S \models_L proven(\psi)$

### Sound Argument and Proven Formula

$K_1$

expert(docWho, weather)  
topic(sunny, weather)  
said(docWho, sunny)

$S_1$

( expArg,  $\left\{ \begin{array}{l} expert(Agent, Topic) \\ said(Agent, Claim) \\ topic(Claim, Topic) \end{array} \right\}, Claim)$   
( modPo,  $\{X, X \rightarrow Y\}, Y)$

$K_1, S_1 \models_L proven(expert(docWho, weather))$

$K_1, S_1 \models_L licit(\{ expert(docWho, weather), said(docWho, sunny), topic(sunny, weather) \}, sunny)$

( $\sigma = \{ docWho/Agent, weather/Topic, sunny/Claim \}$  on expertArg scheme)

$K_1, S_1 \models_L sound(\{ expert(docWho, weather), said(docWho, sunny), topic(sunny, weather) \}, sunny)$

$K_1, S_1 \models_L proven(sunny)$

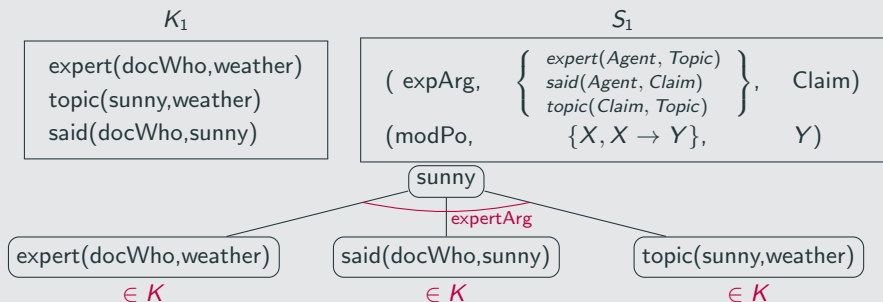
# To Classical Logic and Beyond

---

# Proof Tree

- $\exists$  finite proof tree of root  $\varphi$  wrt  $K$  and  $S$  denoted:  $K \vdash_S \varphi$

## Proof tree for *sunny* ( $K_1 \vdash_{S_1} \text{sunny}$ )



## Proposition

$K, S \models_L \text{proven}(\varphi)$  iff  $K \vdash_S \varphi$

# Soundness of this Framework

$S$  is *cl-valid* iff<sub>def</sub>  $\forall (i, \Psi, \varphi) \in S, \Psi \models \varphi$

$S$  is *cl-valid*,  $S_1$  is not

$S$

$(\text{modPo}, \{X, X \rightarrow Y\}, Y)$

$S_1$

$(\text{expArg}, \left\{ \begin{array}{l} \text{expert}(\text{Agent}, \text{Topic}) \\ \text{said}(\text{Agent}, \text{Claim}) \\ \text{topic}(\text{Claim}, \text{Topic}) \end{array} \right\}, \text{Claim})$   
 $(\text{modPo}, \{X, X \rightarrow Y\}, Y)$

## Proposition (cl-validity)

If  $S$  is *cl-valid* then  $\forall \varphi \in L_0, K \subseteq L_0,$

$K, S \models_L \text{proven}(\varphi) \Rightarrow K \models \varphi$

# Completeness of this Framework

$S$  is *cl-complete* iff<sub>def</sub>  $\forall \varphi \in L_0, \models \varphi$  implies  $\vdash_S \varphi$

$S$  is not *cl-complete*,  $S_H$  is *cl-complete*

$S$

$S_H$

(modPo, { $X, X \rightarrow Y$ },  $Y$ )

(modPo, { $X, X \rightarrow Y$ },  $Y$ )

(hilbertK,  $\emptyset$ ,  $X \rightarrow (Y \rightarrow X)$ )

(hilbertS,  $\emptyset$ ,  $(X \rightarrow (Y \rightarrow Z)) \rightarrow ((X \rightarrow Y) \rightarrow (X \rightarrow Z))$ )

(hilbertNot,  $\emptyset$ ,  $(\neg Y \rightarrow \neg X) \rightarrow ((\neg Y \rightarrow X) \rightarrow Y)$ )

## Proposition (cl-completeness)

If  $S$  is *cl-complete* then  $\forall \varphi \in L_0, K \subseteq L_0,$

$$K \models \varphi \Rightarrow K, S \models_L \text{proven}(\varphi)$$



# Escaping ex-falso quodlibet (1)

## No diploma for Doctor Who

$K_2$

$expert(docWho, weather)$   
 $topic(sunny, weather)$   
 $said(docWho, sunny) \quad K_1$

$nodiapl(docWho, weather)$   
 $nodiapl(Agent, Topic) \rightarrow$   
 $\neg expert(Agent, Topic)$

$S_1$

$(expArg, \left\{ \begin{array}{l} expert(Agent, Topic) \\ said(Agent, Claim) \\ topic(Claim, Topic) \end{array} \right\}, Claim)$   
 $(modPo, \{X, X \rightarrow Y\}, Y)$

- $K_2, S_1 \models_L proven(\neg expert(docWho, weather))$
- still  $K_2, S_1 \models_L proven(expert(docWho, weather)) \wedge proven(sunny)$

## Escaping ex-falso quodlibet (2)

- $L ::$  *licit*( $\Psi, \varphi$ ); *proven*( $\varphi$ ); *sound*( $\Psi, \varphi$ ); *unctrv*( $\varphi$ ); *robust*( $\Psi, \varphi$ )
- **Uncontroversial Formula** = like “proven” but negation not proven:

$K, S \models_L \text{unctrv}(\varphi)$  if<sub>def</sub>  $K, S \not\models_L \text{proven}(\neg\varphi)$   
and  $(\varphi \in K$  or  $\exists \Psi \in 2^{L_0}$  s.t.  $K, S \models_L \text{robust}(\Psi, \varphi)$ )

- **Robust Argument** = licit and **uncontroversial** premises:

$K, S \models_L \text{robust}(\Psi, \varphi)$  if<sub>def</sub>  $K, S \models_L \text{licit}(\Psi, \varphi)$  and  $\forall \psi \in \Psi, K, S \models_L \text{unctrv}(\psi)$

### No diploma for doctor Who

$K_1 +$   
 $K_2$   $\text{nodipl}(\text{docWho}, \text{weather})$   
 $\text{nodipl}(\text{Agent}, \text{Topic}) \rightarrow$   
 $\neg \text{expert}(\text{Agent}, \text{Topic})$

$S_1$   $(\text{expArg}, \dots, \dots)$   
 $(\text{modPo}, \{X, X \rightarrow$   
 $Y\}, Y)$

$K_2, S_1 \models_L \text{unctrv}(\text{topic}(\text{sunny}, \text{weather}))$  but  $K_2, S_1 \not\models_L \text{unctrv}(\text{sunny})$

# **(Fallacious + Rational) Schemes and Reasoning**

---

# Kelley's categories of Fallacies (1)

(Kelley 2013): “the variety of bad reasoning are too numerous to catalog here”  
⇒ restricted to four categories:

## 1. Subjectivist fallacies:

- a) Subjectivism: “I want/believe in  $p$ ” **hence**  $p$  holds.
- b) Appeal to majority: majority believes  $p$  **hence**  $p$  holds.
- c) Appeal to emotion: explicit/implicit emotion instead of evidence.
- d) Appeal to force: threat instead of evidence.

## 2. Fallacies involving credibility:

- a) Appeal to Authority: agent  $A$  says  $p$  **hence**  $p$  holds.
- b) *Ad Hominem*:  $A$  says  $p$ ,  $A$  has negative trait **hence**  $p$  is false.

## 3. Fallacies of Context: significant logical gap

- a) False Alternative: Either  $p$  or  $q$ ,  $\neg q$  **hence**  $p$
- b) *Post Hoc*:  $X$  occurred before  $Y$  **hence**  $X$  caused  $Y$ .

## Kelley's categories of Fallacies (2)

- c) Hasty Generalization: not enough variety to be representative.
- d) Hasty Application: apply generalization to exceptional case
- e) Slippery Slope: action  $X$  lead to  $Y$  lead to  $Z$ ,  $Z$  bad **hence** avoid  $X$
- f) Composition (Division): true of a part (the whole) **hence** true of the whole (a part)

### 4. Fallacies of Logical Structure

- a) Begging the Question:  $p$  **hence**  $p$  (formulated differently)
- b) Equivocation: a word used in premise and conclusion with different meaning.
- c) Appeal to Ignorance:  $\neg p$  not proven **hence**  $p$  holds
- d) Diversion: changing the issue in the middle of an argument.
  - *Straw man*: distorts an opponents position and then refutes it.
  - *Non sequitur*: premises completely unrelated to conclusion.

# Proposal of fallacious schemes encoding

Fallacy	Scheme
Subjectivism	$(f1a, \{likeable(X)\}, X)$
Majority	$(f1b, \{majoritarian(X)\}, X)$
Authority*	$(f2a, \{expert(A, T), topic(X, T), said(A, X)\}, X)$
<i>Ad Hominem</i>	$(f2b, \{said(A, X), \neg likeable(A)\}, \neg X)$
False Alternative*	$(f3a, \{X \rightarrow \neg Y, \neg X \rightarrow Y, \neg X\}, Y)$
<i>Post Hoc</i>	$(f3b, \{before(X, Y)\}, cause(X, Y))$
Hasty Generalization	$(f3c, \{hasProp(X, P), Y \rightarrow X\}, hasProp(Y, P))$
Accident	$(f3d, \{hasProp(X, P), X \rightarrow Y\}, hasProp(Y, P))$
Slippery Slope	$(f3e, \{cause(X, Y), cause(Y, Z) \neg likeable(Z)\}, \neg do(X))$
Composition	$(f3f, \{hasProp(X, P), part(X, Y)\}, hasProp(Y, P))$
Begging the Question*	$(f4a, \{X\}, X)$

\* not fallacious!

# Rationality and consistency

- $S$ : divided into (rational and sophistic) schemes  $S = S_R \cup S_S$
- $a$  is **fallacious** wrt.  $K, S$  iff<sub>def</sub>  $K, S_R \not\models_L \text{robust}(a)$
- Hence 3 cases where  $a = (\Psi, \varphi)$  is fallacious :
  1. premise not proven ( $\exists \psi \in \Psi, K, S_R \not\models_L \text{proven}(\psi)$ )
  2. controversial premise ( $\exists \psi \in \Psi, K, S_R \models_L \text{proven}(\psi) \wedge \text{proven}(\neg\psi)$ )
  3. sophistic or unrecorded scheme ( $K, S_R \not\models_L \text{licit}(a)$ )
    - characterizes the *Non Sequitur* fallacy,
    - cover cases like *Appeal to Emotion* and *Appeal to Force*

## Proposition

$K, S$  s.t.  $K \cup \{(\sigma(\Psi) \rightarrow \sigma(\varphi) \mid \sigma \in \mathcal{X}_1 \rightarrow \text{At}_0, (id, \Psi, \varphi) \in S_R)\} \not\vdash \perp$ ,

$a$  is fallacious wrt.  $K, S$  iff  $K, S_R \not\models_L \text{sound}(a)$

- When  $S_R$  is **not potentially inconsistent with  $K$** , fallacious simply means **unsound** argument (wrt **rational schemes**)

# Conclusion

---



## Conclusion: summing up

**Aim**= help people to build efficiently *sound* arguments and fight fallacies (especially in case of low cognitive availability – Bisquert et al. 2017).

- **Framework to assess arguments wrt**
  - its logical intrinsic content,
  - a logical Knowledge Base
  - a set of recorded argument Schemes
- **Clarify various forms of validity wrt target**
  - Logical Formula: “valid/unvalid” (called “proven”)
  - Instantiated Argument: “licit/illicit” and “sound/unsound”
  - Generic Argument Scheme: “rational/sophism”
- **Framework flexible enough to represent user-defined argument schemes**
- **Logical encoding of 11 classical fallacious inference schemes.**
- **Prolog Tool** to identify the schemes used + check soundness.





- *Rejection Calculi* (Lukasiewicz 1957) developed for classical logic, intuitionistic logic, modal logics (Skura 2011; Oetsch and Tompits 2011; Goranko 1994)
- Detection of *petitio principii* fallacy by a dialog game (Mackenzie 1979)
- Argumentation:
  - ABA (Modgil and Prakken 2013), ASPIC+ (Prakken 2010) and Carneades (Thomas F Gordon, Prakken, and Walton 2007) structured arguments + interactions
  - Assessment only based on interactions: counter-intuitive results (Amgoud and Besnard 2013)
  - No counter-argument stated: does not mean sound argument.

- Schemes for more types of rational reasoning (and their possible flaws)
- Extend language to more complex arguments (meta-arguments)
- Protocol governing authorized moves (“dialog support system”)
  - speech act **Argue**: commits the speaker to be able to provide a sound proof of her claim
  - automatic verification of compliance (given  $K$  and  $S$ )
  - alert user about biased reasoning, incite user to give better grounds
- Encode critical questions (*Toulmin 1958; Walton and Thomas F. Gordon 2005; Besnard et al. 2014*)








## References

---

## References (1)

-  Amgoud, Leila and Philippe Besnard (2013). “Logical limits of abstract argumentation frameworks”. In: *Journal of Applied Non-Classical Logics* 23.3, pp. 229–267.
-  Besnard, Philippe et al. (2014). “Introducing structured argumentation”. In: *Argument & Computation* 5.1, pp. 1–4.
-  Bisquert, Pierre et al. (2017). “Formalizing Cognitive Acceptance of Arguments: Durum Wheat Selection Interdisciplinary Study”. anglais. In: *Minds & Machines* 27.1, pp. 233–252.
-  Dung, Phan Minh (1995). “On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games”. In: *Artificial Intelligence J.* 77, pp. 321–357.

## References (2)

-  Goranko, Valentin (1994). "Refutation systems in modal logic". In: *Studia Logica* 53.2, pp. 299–324.
-  Gordon, Thomas F, Henry Prakken, and Douglas Walton (2007). "The Carneades model of argument and burden of proof". In: *Artificial Intelligence J.* 171.10-15, pp. 875–896.
-  Hamblin, Charles L. (1970). *Fallacies*. London: Methuen.
-  Kelley, David (2013). *The art of reasoning: An introduction to logic and critical thinking*. New York: W.W. Norton & Company.
-  Lukasiewicz, Jan (1957). *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic, 2nd edition*. Oxford: Clarendon Press.
-  Mackenzie, Jim D (1979). "Question-begging in non-cumulative systems". In: *J. of Philosophical Logic* 8.1, pp. 117–133.
-  Modgil, Sanjay and Henry Prakken (2013). "A general account of argumentation with preferences". In: *Artificial Intelligence J.* 195, pp. 361–397.

## References (3)



Oetsch, Johannes and Hans Tompits (2011). “Gentzen-type Refutation Systems for Three-valued Logics with an Application to Disproving Strong Equivalence”. In: *11th Int. Conf. on Logic Programming and Nonmonotonic Reasoning*. Vol. 6645. LNCS, pp. 254–259.



Prakken, Henry (2010). “An abstract framework for argumentation with structured arguments”. In: *Argument & Computation* 1.2, pp. 93–124.



Skura, Tomasz F. (2011). “Refutation Systems in Propositional Logic”. In: *Handbook of Philosophical Logic*. Ed. by Dov Gabbay and Franz Guentner. 2nd. Vol. 16. Springer, pp. 115–157.



Toulmin, Stephen (1958). *The Uses of Argument*. Cambridge University Press.



Walton, Douglas and Thomas F. Gordon (2005). “Critical questions in computational models of legal argument”. In: *Argumentation in Artificial Intelligence and Law Workshop*. Wolf Legal Publishers, pp. 103–111.

## References (4)



Wigmore, John Henry (1931). *The Principles of Judicial Proof*. 2nd. Little, Brown.



Woods, John (1994). "Is the Theoretical Unity of the Fallacies Possible?" In: *Informal Logic* XVI, pp. 77–85.



# Prolog Implementation

- predicate `arg/2`: the user declares all arguments usable in proofs
- encoding of predicates `proven`, `licit`, `sound`, `unctrv`, `robust`
- + unbound variable: schemes used in the proof.

## Umbrella

```
|?- proven([take(umbrella)], Schemes).  
Schemes = [[modPo, fact(rain),  
            fact(implies(rain, take(umbrella))  
            )]]
```

## Hilbert System

```
|?- proven([implies(f,f)], S).  
S = [[modPo, [modPo, [hilbertK], [hilbertS]],  
      [hilbertK]]]
```