

Abductive Reasoning with Sequent-Based Argumentation

(Extended Abstract)

Ofer Arieli¹, AnneMarie Borg², Matthis Hesse³ and Christian Straßer³

¹School of Computer Science, Tel-Aviv Academic College, Israel

²Department of Information and Computing Sciences, Utrecht University, The Netherlands

³Institute for Philosophy II, Ruhr University Bochum, Germany

Abstract

We show that logic-based argumentation, and in particular sequent-based frameworks, is a robust argumentative setting for abductive reasoning and explainable artificial intelligence.

1. Introduction

Abduction is the process of deriving a set of explanations of a given observation relative to a set of assumptions. The systematic study of abductive reasoning goes back to Peirce (see [1]). Abduction is closely related to ‘inference to the best explanation (IBE)’ [2]. However, it is often distinguished from the latter in that abductive inference may provide explanations that are not known as *the* best explanation available, but that are merely worthy of conjecturing or entertaining (see, e.g., [3, 4, 5]).

In this work, we model abduction (not in the strict sense of IBE) by computational argumentation, and show that sequent-based argumentation frameworks [6, 7] are a solid argumentative base for abductive reasoning. According to our approach, abductive explanations are handled by ingredients of the framework, and so different considerations and principles concerning those explanations are expressed *within the framework*. The advantages of this are discussed in the last section of the paper.

2. Sequent-Based Argumentation

We denote by \mathcal{L} a propositional language. Atomic formulas in \mathcal{L} are denoted by p, q, r , formulas are denoted by $\phi, \psi, \delta, \gamma, \epsilon$, sets of formulas are denoted by $\mathcal{X}, \mathcal{S}, \mathcal{E}$, and finite sets of formulas are denoted by $\Gamma, \Delta, \Pi, \Theta$, all of which can be primed or indexed. The set of atomic formulas appearing in the formulas of \mathcal{S} is denoted $\text{Atoms}(\mathcal{S})$. The set of the (well-formed) formulas of \mathcal{L} is denoted $\text{WFF}(\mathcal{L})$, and its power set is denoted $\wp(\text{WFF}(\mathcal{L}))$.

• **The base logic** is an arbitrary propositional logic, namely a pair $\mathbf{L} = \langle \mathcal{L}, \vdash \rangle$ consisting of a language \mathcal{L} and a consequence relation \vdash on $\wp(\text{WFF}(\mathcal{L})) \times \text{WFF}(\mathcal{L})$. The relation \vdash is assumed to be *reflexive* ($\mathcal{S} \vdash \phi$ if $\phi \in \mathcal{S}$),

monotonic (if $\mathcal{S}' \vdash \phi$ and $\mathcal{S}' \subseteq \mathcal{S}$, then $\mathcal{S} \vdash \phi$), and *transitive* (if $\mathcal{S} \vdash \phi$ and $\mathcal{S}', \phi \vdash \psi$ then $\mathcal{S}, \mathcal{S}' \vdash \psi$).

• **The language** \mathcal{L} contains at least a \vdash -negation operator \neg , satisfying $p \not\vdash \neg p$ and $\neg p \not\vdash p$ (for atomic p), and a \vdash -conjunction operator \wedge , for which $\mathcal{S} \vdash \psi \wedge \phi$ iff $\mathcal{S} \vdash \psi$ and $\mathcal{S} \vdash \phi$. We denote by $\bigwedge \Gamma$ the conjunction of the formulas in Γ . We sometimes assume the availability of a deductive implication \rightarrow , satisfying $\mathcal{S}, \psi \vdash \phi$ iff $\mathcal{S} \vdash \psi \rightarrow \phi$.

A set \mathcal{S} of formulas is \vdash -consistent, if there are no formulas $\phi_1, \dots, \phi_n \in \mathcal{S}$ for which $\vdash \neg(\phi_1 \wedge \dots \wedge \phi_n)$.

• **Arguments** based on a logic $\mathbf{L} = \langle \mathcal{L}, \vdash \rangle$ are single-conclusioned *L-sequents* [8], i.e., expressions of the form $\Gamma \Rightarrow \psi$, where \Rightarrow is a symbol that does not appear in \mathcal{L} , and such that $\Gamma \vdash \psi$. Γ is called the argument’s support (also denoted $\text{Supp}(\Gamma \Rightarrow \psi)$) and ψ is its conclusion (denoted $\text{Conc}(\Gamma \Rightarrow \psi)$). An \mathcal{S} -based argument is an \mathbf{L} -argument $\Gamma \Rightarrow \psi$, where $\Gamma \subseteq \mathcal{S}$. We denote by $\text{Arg}_{\mathbf{L}}(\mathcal{S})$ the set of the \mathbf{L} -arguments that are based on \mathcal{S} .

We distinguish between two types of premises: a \vdash -consistent set \mathcal{X} of strict premises, and a set \mathcal{S} of defeasible premises. We write $\text{Arg}_{\mathbf{L}}^{\mathcal{X}}(\mathcal{S})$ for $\text{Arg}_{\mathbf{L}}(\mathcal{X} \cup \mathcal{S})$.

• **Attack rules** are sequent-based inference rules for representing attacks between sequents. Such rules consist of an attacking argument (the first condition of the rule), an attacked argument (the last condition of the rule), conditions for the attack (the other conditions of the rule) and a conclusion (the eliminated attacked sequent). The elimination of $\Gamma \Rightarrow \phi$ is denoted by $\Gamma \not\Rightarrow \phi$.

Given a set \mathcal{X} of strict (non-attacked) formulas, we shall concentrate here on the following two attack rules:

Direct Defeat (for $\gamma \notin \mathcal{X}$):

$$\frac{\Gamma_1 \Rightarrow \psi_1 \quad \psi_1 \Rightarrow \neg \gamma \quad \Gamma_2, \gamma \Rightarrow \psi_2}{\Gamma_2, \gamma \not\Rightarrow \psi_2}$$

Consistency Undercut (for $\Gamma_2 \neq \emptyset, \Gamma_2 \cap \mathcal{X} = \emptyset, \Gamma_1 \subseteq \mathcal{X}$):

$$\frac{\Gamma_1 \Rightarrow \neg \bigwedge \Gamma_2 \quad \Gamma_2, \Gamma'_2 \Rightarrow \psi}{\Gamma_2, \Gamma'_2 \not\Rightarrow \psi}$$

NMR 2022: 20th International Workshop on Non-Monotonic Reasoning, August 7-9, 2022, Haifa, Israel

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)



Direct Defeat (DirDef) indicates that if the conclusion (ψ_1) of the attacker entails the negation of a formula (γ) in the support of an argument, the latter is eliminated. When $\Gamma_1 = \emptyset$, consistency undercut (ConUcut) eliminates an argument with an inconsistent support.

• **A (sequent-based) argumentation framework (AF)**, based on the logic L and the attack rules in AR , for a set of defeasible premises \mathcal{S} and a \vdash -consistent set of strict premises \mathcal{X} , is a pair $\mathbb{AF}_{L,AR}^{\mathcal{X}}(\mathcal{S}) = \langle \text{Arg}_L^{\mathcal{X}}(\mathcal{S}), A \rangle$ where $A \subseteq \text{Arg}_L^{\mathcal{X}}(\mathcal{S}) \times \text{Arg}_L^{\mathcal{X}}(\mathcal{S})$ and $(a_1, a_2) \in A$ iff there is a rule $R_{\mathcal{X}} \in AR$, such that a_1 $R_{\mathcal{X}}$ -attacks a_2 . We shall use AR and A interchangeably, denoting both of them by A .

• **Semantics** of sequent-based frameworks are defined as usual by Dung-style extensions [9]: Let $\mathbb{AF} = \mathbb{AF}_{L,A}^{\mathcal{X}}(\mathcal{S}) = \langle \text{Arg}_L^{\mathcal{X}}(\mathcal{S}), A \rangle$ be an AF and let $\mathbb{E} \subseteq \text{Arg}_L^{\mathcal{X}}(\mathcal{S})$. \mathbb{E} attacks a if there is an $a' \in \mathbb{E}$ such that $(a', a) \in A$. \mathbb{E} defends a if \mathbb{E} attacks every attacker of a , and \mathbb{E} is *conflict-free* (cf) if for no $a_1, a_2 \in \mathbb{E}$ it holds that $(a_1, a_2) \in A$. \mathbb{E} is *admissible* if it is conflict-free and defends all of its elements. A *complete* (cmp) extension of \mathbb{AF} is an admissible set that contains all the arguments that it defends. The *grounded* (grd) extension of \mathbb{AF} is the \subseteq -minimal complete extension of $\text{Arg}_L^{\mathcal{X}}(\mathcal{S})$, a *preferred* (prf) extension of \mathbb{AF} is a \subseteq -maximal complete extension of $\text{Arg}_L^{\mathcal{X}}(\mathcal{S})$, and a *stable* (stb) extension of \mathbb{AF} is a conflict-free set in $\text{Arg}_L^{\mathcal{X}}(\mathcal{S})$ that attacks every argument not in it.¹ We denote by $\text{Ext}_{\text{sem}}(\mathbb{AF})$ the set of all the extensions of \mathbb{AF} of type *sem*.

• **Entailments** of $\mathbb{AF} = \mathbb{AF}_{L,A}^{\mathcal{X}}(\mathcal{S}) = \langle \text{Arg}_L^{\mathcal{X}}(\mathcal{S}), A \rangle$ with respect to a semantics *sem* are defined as follows:

- *Skeptical entailment*: $\mathcal{S} \vdash_{L,A,\mathcal{X}}^{\cap, \text{sem}} \phi$ if there is an argument $a \in \bigcap \text{Ext}_{\text{sem}}(\mathbb{AF})$ such that $\text{Conc}(a) = \phi$.
- *Weakly skeptical entailment*: $\mathcal{S} \vdash_{L,A,\mathcal{X}}^{\cap, \text{sem}} \phi$ if for every extension $\mathbb{E} \in \text{Ext}_{\text{sem}}(\mathbb{AF})$ there is an argument $a \in \mathbb{E}$ such that $\text{Conc}(a) = \phi$.
- *Credulous entailment*: $\mathcal{S} \vdash_{L,A,\mathcal{X}}^{\cup, \text{sem}} \phi$ iff there is an argument $a \in \bigcup \text{Ext}_{\text{sem}}(\mathbb{AF})$ such that $\text{Conc}(a) = \phi$.

Example 1. Consider a sequent-based AF, based on classical logic CL and the set \mathcal{S} of defeasible assumptions:

$$\left\{ \begin{array}{l} \text{clear_skies, rainy, clear_skies} \rightarrow \neg \text{rainy,} \\ \text{rainy} \rightarrow \neg \text{sprinklers, rainy} \rightarrow \text{wet_grass,} \\ \text{sprinklers} \rightarrow \text{wet_grass} \end{array} \right\}$$

Suppose further that $\mathcal{X} = \emptyset$ and the attack rules are DirDef and ConUcut. Then, for instance, the arguments

$$a_1: \text{clear_skies, clear_skies} \rightarrow \neg \text{rainy} \Rightarrow \neg \text{rainy}$$

$$a_2: \text{rainy, clear_skies} \rightarrow \neg \text{rainy} \Rightarrow \neg \text{clear_skies}$$

DirDef-attack each other. There are two stable/preferred extensions \mathbb{E}_1 and \mathbb{E}_2 , where $a_1 \in \mathbb{E}_1$ and $a_2 \in \mathbb{E}_2$ (see Fig. 1). Thus, with respect to stable or preferred semantics, `wet_grass` credulously follows from the framework

¹For an in-depth discussion of extension types see [10].

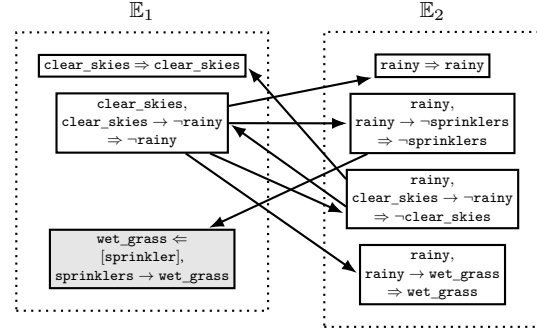


Figure 1: Part of the AF of Example 1 (without the gray node) and of Example 2 (with the gray node).

(since the argument `rainy, rainy` \rightarrow `wet_grass` \Rightarrow `wet_grass` is in \mathbb{E}_2), but it is not skeptically deducible (there is no $a \in \mathbb{E}_1$ such that $\text{Conc}(a) = \text{wet_grass}$).

3. Abductive Reasoning

For supporting abductive explanations in sequent-based argumentation, we introduce *abductive sequents*, which are expressions of the form $\phi \Leftarrow \Gamma, [\epsilon]$, intuitively meaning that ‘(the explanandum) ϕ may be inferred from Γ , assuming that ϵ holds’. While $\Gamma \subseteq \mathcal{S} \cup \mathcal{X}$, ϵ may not be an assumption, but rather a hypothetical explanation of the conclusion.

Abductive sequents are produced by the following rule that models abduction as ‘backwards reasoning’:

$$\bullet \text{ Abduction: } \frac{\epsilon, \Gamma \Rightarrow \phi}{\phi \Leftarrow \Gamma, [\epsilon]}$$

This rule allows us to produce abductive sequents like `wet_grass` \Leftarrow `[sprinklers]`, `sprinklers` \rightarrow `wet_grass` that provides an explanation to `wet_grass`.

Since abductive reasoning is a form of non-monotonic reasoning, we need a way to attack abductive sequents. To this end, we consider rules like those from Section 2:

$$\bullet \text{ Abductive Direct Defeat (for } \gamma \in (\Gamma_2 \cup \{\epsilon\}) \setminus \mathcal{X} \text{):}$$

$$\frac{\Gamma_1 \Rightarrow \phi_1 \quad \phi_1 \Rightarrow \neg \gamma \quad \phi_2 \Leftarrow [\epsilon], \Gamma_2}{\phi_2 \Leftarrow [\epsilon], \Gamma_2}$$

Note that this attack rule assures, in particular, the consistency of explanations with the strict assumptions, thus it renders the following rule admissible:

$$\bullet \text{ Consistency (for } \Gamma_1 \subseteq \mathcal{X} \text{):}$$

$$\frac{\Gamma_1 \Rightarrow \neg \epsilon \quad \phi \Leftarrow [\epsilon], \Gamma_2}{\phi \Leftarrow [\epsilon], \Gamma_2}$$

Abductive explanations should meet certain requirements to ensure their behavior (see, e.g., [11]). Below,

we express some of the common properties in terms of attack rules that may be added to the framework.

- Non Vacuousity:
$$\frac{\vdash \epsilon \rightarrow \phi \quad \phi \Leftarrow [\epsilon]}{\phi \not\Leftarrow [\epsilon]}$$

This rule prevents self-explanations. Thus, in the running example, $\text{wet_grass} \Leftarrow [\text{wet_grass}]$ is excluded.

- Minimality:

$$\frac{\phi \Leftarrow [\epsilon_1], \Gamma \quad \epsilon_2 \Rightarrow \epsilon_1 \quad \epsilon_1 \not\Rightarrow \epsilon_2 \quad \phi \Leftarrow [\epsilon_2], \Gamma}{\phi \not\Leftarrow [\epsilon_2], \Gamma}$$

This rule assures the generality of explanations. Thus, in our example, $\text{sprinklers} \wedge \text{irrelevant_fact}$ should not explain wet_grass , since sprinklers is a more general and so more relevant explanation.

- Defeasible Non-Idleness:
$$\frac{\Gamma_1 \Rightarrow \phi \quad \phi \Leftarrow [\epsilon], \Gamma_2}{\phi \not\Leftarrow [\epsilon], \Gamma_2}$$

- Strict Non-Idleness ($\Gamma_1 \subseteq \mathcal{X}$):
$$\frac{\Gamma_1 \Rightarrow \phi \quad \phi \Leftarrow [\epsilon], \Gamma_2}{\phi \not\Leftarrow [\epsilon], \Gamma_2}$$

The two rules above assure that assumptions shouldn't already explain the explanandum. Defeasible non-idleness rules out explaining wet_grass by sprinklers , since the former is already inferred from the defeasible assumptions (assuming that it is rainy), while strict non-idleness allows this alternative explanation (wet_grass cannot be inferred from the strict assumptions). These two attack rules are particularly interesting when abductive reasoning is used to generate novel hypotheses explaining observations that are not already explained by a given theory resp. the given background assumptions.²

Next, we adapt sequent-based argumentation frameworks to an abductive setting, using abductive sequents, the new inference rule, and additional attack rules.

Given a sequent-based framework $\mathbb{A}\mathbb{F}_{\mathcal{L},\mathcal{A}}^{\mathcal{X}}(\mathcal{S})$, an *abductive sequent-based framework* $\mathbb{A}\mathbb{A}\mathbb{F}_{\mathcal{L},\mathcal{A}^*}^{\mathcal{X}}(\mathcal{S})$ is constructed by adding to the arguments in $\text{Arg}_{\mathcal{L}}^{\mathcal{X}}(\mathcal{S})$ also abductive arguments, produced by Abduction, and where \mathcal{A}^* is obtained by adding to the attack rules in \mathcal{A} also (some of) the rules for maintaining explanations that are described above. Explanations are then defined as follows:

Definition 1. Let $\mathbb{A}\mathbb{A}\mathbb{F}_{\mathcal{L},\mathcal{A}^*}^{\mathcal{X}}(\mathcal{S})$ be an abductive sequent-based argumentation framework as described above. A finite set \mathcal{E} of \mathcal{L} -formulas is called:

- *skeptical sem-explanation* of ϕ , if there is $\Gamma \subseteq \mathcal{S}$ s.t. $\phi \Leftarrow [\bigwedge \mathcal{E}], \Gamma$ is in every sem-extension of $\mathbb{A}\mathbb{A}\mathbb{F}_{\mathcal{L},\mathcal{A}^*}^{\mathcal{X}}(\mathcal{S})$.

²In some accounts of abduction, e.g. [5], it is argued that the abductively inferred ϵ should be of lesser epistemic status than the reasoner's starting point and so "the fundamental conceptual fact about abduction is that abduction is ignorance-preserving reasoning" (p. 40). Our attack rules ensure that the reasoner faces what Gabbay & Woods call an 'ignorance problem' (p. 42, Def. 3.2).

- *weakly-skeptical sem-explanation* of ϕ , if in every sem-extension of $\mathbb{A}\mathbb{A}\mathbb{F}_{\mathcal{L},\mathcal{A}^*}^{\mathcal{X}}(\mathcal{S})$ there is an abductive argument $\phi \Leftarrow [\bigwedge \mathcal{E}], \Gamma$ for some $\Gamma \subseteq \mathcal{S}$.

- *credulous sem-explanation* of ϕ , if there is $\Gamma \subseteq \mathcal{S}$ such that $\phi \Leftarrow [\bigwedge \mathcal{E}], \Gamma$ is in some sem-extension of $\mathbb{A}\mathbb{A}\mathbb{F}_{\mathcal{L},\mathcal{A}^*}^{\mathcal{X}}(\mathcal{S})$.

Example 2. As mentioned, the abductive sequent $\text{wet_grass} \Leftarrow [\text{sprinklers}]$, $\text{sprinklers} \rightarrow \text{wet_grass}$ is producible by Abduction from the sequent-based framework in Example 1, and belongs to a stable/preferred extension of the related abductive sequent-based framework (see again Fig. 1). Therefore, sprinklers is a credulous (but not [weakly] skeptical) stb/prf-explanation of wet_grass .

Example 3. Let $\mathcal{L} = \text{CL}$, $\mathcal{A} = \{\text{DirDef}, \text{ConUcut}\}$ with $\mathcal{S} = \{p, \neg p \wedge q\}$ and $\mathcal{X} = \{q \wedge r \rightarrow s\}$. For $\text{sem} \in \{\text{stb}, \text{prf}\}$, $q \wedge r$ is a weakly-skeptical sem-explanation of s , since the corresponding abductive framework has two sem-extensions, one with $s \Leftarrow [q \wedge r]$, p , $q \wedge r \rightarrow s$ and the other with $s \Leftarrow [q \wedge r]$, $\neg p \wedge q$, $q \wedge r \rightarrow s$. This holds also when the non-vacuousity or the strict non-idleness attack rules are part of the framework. However, $q \wedge r$ is *no longer* a weakly-skeptical sem-explanation of s when minimality attack is added, since the extension that contains $s \Leftarrow [q \wedge r]$, $\neg p \wedge q$, $q \wedge r \rightarrow s$ includes a minimality attacker, $s \Leftarrow [r]$, $\neg p \wedge q$, $q \wedge r \rightarrow s$.

Example 4. Consider now $\mathcal{S} = \{p \wedge q, \neg p \wedge q\}$. This time, with minimality, $q \wedge r$ is not even a credulous sem-explanation of s ($\text{sem} \in \{\text{stb}, \text{prf}\}$), since each of the two sem-extensions contains a minimality attacker ($s \Leftarrow [r]$, $p \wedge q$, $q \wedge r \rightarrow s$ or $s \Leftarrow [r]$, $\neg p \wedge q$, $q \wedge r \rightarrow s$). So, $q \wedge r$, unlike r , does not sem-explain s .

4. Discussion and Conclusion

Abduction has been widely applied in different deductive systems (such as adaptive logics [12]) and AI-based disciplines (e.g., logic programming [13]), including in the context of formal argumentation (see the survey in [14]).

This ongoing work offers several novelties. In terms of knowledge representation we transparently represent abductive inferences by an explicit inference rule that produces abductive arguments. The latter are a new type of hypothetical arguments that are subjected to potential defeats. Specifically designed attack rules address the quality of the offered explanation and thereby model critical questions [15] and meta-argumentative reasoning [16]. This is both natural and philosophically motivated, as argued in [17]. Our framework offers a high degree of modularity, and may be based on a variety of propositional logics. Desiderata on abductive arguments can be disambiguated in various ways by simply changing the attack rules, all in the same base framework.

References

- [1] C. S. Peirce, *The Collected Papers of Charles Sanders Peirce, Vol. I: The Principles of Philosophy*, Harvard University Press, Cambridge, 1931.
- [2] P. Lipton, *Inference to the Best Explanation*, Routledge, 2004. 2nd edition.
- [3] S. Yu, F. Zenker, Peirce knew why abduction isn't IBE – A scheme and critical questions for abductive argument, *Argumentation* 32 (2017) 569–587.
- [4] G. Minnameier, Peirce-suit of truth – Why inference to the best explanation and abduction ought not to be confused, *Erkenntnis* 60 (2004) 75–105.
- [5] D. Gabbay, J. Woods, *The reach of abduction. A practical logic of cognitive systems*, North-Holland, 2005.
- [6] O. Arieli, C. Straßer, Sequent-based logical argumentation, *Argument & Computation* 6 (2015) 73–99.
- [7] A. Borg, Assumptive sequent-based argumentation, *Journal of Applied Logics – IfCoLog Journal of Logics and Their Applications* 7 (2020) 227–294.
- [8] G. Gentzen, Untersuchungen über das logische schließen I, II, *Mathematische Zeitschrift* 39 (1934) 176–210, 405–431.
- [9] P. M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games, *Artificial Intelligence* 77 (1995) 321–357.
- [10] P. Baroni, M. Caminada, M. Giacomin, Abstract argumentation frameworks and their semantics, in: P. Baroni, D. Gabbay, M. Giacomin, L. van der Torre (Eds.), *Handbook of Formal Argumentation*, volume 1, College Publications, 2018, pp. 159–236.
- [11] J. Meheus, L. Verhoeven, M. Van Dyck, D. Provijn, Ampliative adaptive logics and the foundation of logic-based approaches to abduction, *Logical and Computational Aspects of Model-Based Reasoning* 25 (2002) 39–71.
- [12] J. Meheus, D. Batens, A formal logic for abductive reasoning, *Logic Journal of the IGPL* 14 (2006) 221–236.
- [13] M. Denecker, A. Kakas, Abduction in logic programming, in: *Computational Logic: Logic Programming and Beyond*, LNCS 2407, Springer, 2002, pp. 402–436.
- [14] K. Cyras, A. Rago, E. Albin, P. Baroni, F. Toni, Argumentative XAI: A survey, in: *Proc. 30th IJCAI*, 2021, pp. 4392–4399.
- [15] D. Walton, C. Reed, F. Macagno, *Argumentation Schemes*, Cambridge University Press, 2008.
- [16] G. Boella, D. Gabbay, L. van der Torre, S. Villata, Meta-argumentation modelling I: Methodology and techniques, *Studia Logica* 93 (2009) 297–355.
- [17] P. Olmos, Abduction and comparative weighing of explanatory hypotheses: an argumentative approach, *Logic Journal of the IGPL* 29(4) (2021) 523–535.