I don't care, I don't know ... I know too much! On Incompleteness and Undecidedness in Abstract Argumentation

Pietro Baroni[1], Massimiliano Giacomin[1], and Beishui Liao[2]

[1] Dip. Ingegneria dell'Informazione, Univ. of Brescia, Brescia, Italy
{pietro.baroni,massimiliano.giacomin}@unibs.it
[2] Center for the Study of Language and Cognition, Zhejiang Univ., Hangzhou, China
baiseliao@zju.edu.cn

Abstract. Incompleteness and undecidedness are pervasively present in human reasoning activities and make the definition of the relevant computational models challenging. In this discussion paper we focus on one such model, namely abstract argumentation frameworks, and examine several flavours of incompleteness and undecidedness thereof, by providing a conceptual analysis, a critical literature review, and some new ideas with pointers to future research.

Keywords: Argumentation frameworks, Argumentation semantics, Incompleteness, Undecidedness.

1 Introduction

In everyday life answering a question is not just a matter of choosing between "Yes" and "No". You may have no interest in giving a definite answer (whether you have it ready or not) or, for some justified reason, may be unable to produce it and prefer to take a less committed position. This variety of behaviors (and the reasoning underlying them) is a key feature of dialogues between human beings, and, in a sense, of human intelligence itself. Providing a formal counterpart to them is therefore a plus, if not a must, for any formal approach aiming at representing and/or supporting intelligent dialogical and/or inferential activities. This paper focuses on one such approach, namely *abstract argumentation*, and aims at providing a conceptual analysis, a critical literature review, and some new ideas with pointers to future research concerning the treatment of incompleteness and undecidedness in this context. In a nutshell, abstract argumentation focuses on the evaluation of the justification status of a set of (typically conflicting) arguments according to a given *argumentation semantics*. So it can be roughly regarded as a formal approach to answer, for every single argument, the question: "Is this argument acceptable?" Analyzing and discussing which answers are available beyond "Yes" and "No" is the subject of this work.

The paper is organized as follows. Section 2 recalls the necessary background concepts, Section 3 is devoted to partial evaluations in argumentation semantics, Section 4 deals with different forms of undecidedness, and Section 5 concludes.

T. Eiter et al. (Eds.): Brewka Festschrift, LNAI 9060, pp. 265–280, 2015.

2 Background

This work lies in the context of Dung's theory [15] of abstract argumentation frameworks (AFs), whose definition is recalled below.

Definition 1. *An* argumentation framework *(AF) is defined as a pair* $\langle \mathcal{A}, \rightarrow \rangle$ *in which* $\mathcal{A}$ *is a set of* arguments *and* $\rightarrow \subseteq \mathcal{A} \times \mathcal{A}$ *describes the* attack relation *between arguments in* $\mathcal{A}$*, so that* $(\alpha, \beta) \in \rightarrow$ *(also denoted as* $\alpha \rightarrow \beta$*) indicates that the argument* α attacks *the argument* β*. For a set* $S \subseteq \mathcal{A}$*, the attackers of* S *are defined as* $S^{\leftarrow} = \{\alpha \in \mathcal{A} \mid \exists \beta \in S : \alpha \rightarrow \beta\}$ *and the attackees of* S *are defined as* $S^{\rightarrow} = \{\alpha \in \mathcal{A} \mid \exists \beta \in S : \beta \rightarrow \alpha\}$*.*

In Dung's theory arguments are abstract entities, whose nature and structure are not specified, as the formalism is focused only on the representation of their conflicts. Given an AF, a basic problem consists in determining the conflict outcome, namely assigning a justification status to arguments. An *argumentation semantics* can be conceived, in broad terms, as a formal way to answer this question.

Two main approaches to semantics definitions have been adopted in the literature (see [5] for a review). In the *extension-based* approach the "outcome" of an argumentation semantics when applied to a given AF is a set of *extensions*, where each extension is a set of arguments considered to be jointly acceptable.

Definition 2. *Given an AF* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$*, an extension-based semantics* σ *associates with* $\mathcal{F}$ *a subset of* $2^{\mathcal{A}}$*, denoted as* $\mathcal{E}_\sigma(\mathcal{F})$*.*

In the *labelling-based* approach the "outcome" is a set of labellings, where a labelling is the assignment to each argument of a label taken from a fixed set.

Definition 3. *Let* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$ *be an AF and* Λ *a set of labels. A* $\Lambda-$labelling *of* $\mathcal{F}$ *is a total function* $L : \mathcal{A} \longrightarrow \Lambda$*. The set of all* $\Lambda-$labellings *of* $\mathcal{F}$ *is denoted as* $\mathfrak{L}(\Lambda, \mathcal{F})$*.*

Definition 4. *Given an AF* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$ *and a set of labels* Λ*, a labelling-based semantics* σ *associates with* $\mathcal{F}$ *a subset of* $\mathfrak{L}(\Lambda, \mathcal{F})$*, denoted as* $\mathcal{L}_\sigma(\mathcal{F})$*.*

Some observations concerning the relationships between the labelling and extension-based approaches are worth making. First, as set membership can be expressed in terms of a binary labelling, e.g. with $\Lambda = \{\in, \notin\}$, the extension-based approach can be regarded as a special case of the general labelling-based approach. The latter is therefore potentially more expressive under a suitable choice of Λ. It has however to be noted that the almost universally adopted choice for Λ in the literature, namely the set $\Lambda^{iou} \triangleq \{\mathtt{in}, \mathtt{out}, \mathtt{und}\}$, has exactly the same expressiveness as the extension-based approach.

To see this, and also to introduce some concepts useful in the sequel, let us give some comments on the common intuitions underlying the two approaches.

First, coming back to the question about the justification status of arguments, each extension and each Λ^{iou}-labelling can be regarded as one of the possible (according to the semantics at hand) answers to the question. Indeed, an extension E identifies as justified arguments the members of E itself while a labelling identifies as justified arguments those that are labelled in. Further, arguments not included in E can be partitioned between those that are attacked by some member of E and those that are not. The former ones can be regarded as definitely rejected, and correspond to those labelled out, while the latter ones are in a sort of intermediate status between acceptance and rejection, and correspond to those labelled und. On this intuitive basis, a one-to-one formal correspondence between extensions and Λ^{iou}-labellings can be defined, which has been shown to hold for the main semantics in the literature [5].

As this paper deals mainly with general notions rather than with semantics-specific properties, we don't need to go through the various argumentation semantics considered in the literature and, for the sake of exemplification, we recall only the definition of *complete* semantics (denoted as $\mathcal{CO}$) in the two approaches. The two definitions are indeed equivalent (using the correspondence mentioned above) though they may appear rather different at first glance.

In the extension-based approach, basically a complete extension is a set of arguments which has no conflicts inside, defends all its elements against external attacks, and includes all the arguments it defends.

Definition 5. *Given an AF $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$, a set $S \subseteq \mathcal{A}$ is* conflict-free *iff $\nexists \alpha, \beta \in S : \alpha \rightarrow \beta$. S* defends *an argument α iff $\forall \beta$ s.t. $\beta \rightarrow \alpha$ $\exists \gamma \in S : \gamma \rightarrow \beta$. The set of arguments defended by S in $\mathcal{F}$ is denoted as $\mathcal{D}_{\mathcal{F}}(S)$. A set S is a* complete *extension of $\mathcal{F}$, i.e. $S \in \mathcal{E}_{\mathcal{CO}}(\mathcal{F})$, iff S is conflict-free and $S = \mathcal{D}_{\mathcal{F}}(S)$.*

In the labelling-based approach, a complete labelling is such that every argument label satisfies some legality constraints taking into account the labels assigned to the attackers of the argument.

Definition 6. *Let L be a Λ^{iou}-labelling of an AF $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$.*

- *An **in**-labelled argument is* legally ***in*** *iff all its attackers are labelled **out**.*
- *An **out**-labelled argument is* legally ***out*** *iff it has at least one attacker that is labelled **in**.*
- *An **und**-labelled argument is* legally ***und*** *iff not all its attackers are labelled **out** and it doesn't have an attacker that is labelled **in**.*

L is a complete labelling, i.e. $L \in \mathcal{L}_{\mathcal{CO}}(\mathcal{F})$ iff every argument is legally labelled.

Given that a semantics provides, in general, many[1] alternative answers (in form of extensions or labellings) to the "argument justification question", it has to be remarked that a further step consists in deriving a "synthetic" justification status for each argument considering the whole set of extensions or labellings.

[1] Most literature semantics provide at least one extension/labelling for every AF, with the exception of *stable* semantics [15] for which the set of extensions/labellings may be empty. To avoid detailed precisations, inessential to the subject of this paper, we assume non-empty sets of extensions/labellings in the following.

In the extension-based approach the two simplest ways to obtain this synthesis basically consist in the set-theoretical operations of intersection and union, leading respectively to the notions of *skeptical* and *credulous* justification, which have an obvious counterpart in the labelling-based approach.

Definition 7. *Given an extension-based semantics* σ *and an AF* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$*, an argument* α *is* skeptically justified *iff* $\forall E \in \mathcal{E}_\sigma(\mathcal{F})$ $\alpha \in E$*; an argument* α *is* credulously justified *iff* $\exists E \in \mathcal{E}_\sigma(\mathcal{F}) : \alpha \in E$*. Given a labelling-based semantics* σ *and an AF* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$*, an argument* α *is* skeptically justified *iff* $\forall L \in \mathcal{L}_\sigma(\mathcal{F})$ $L(\alpha) = \mathtt{in}$*;* α *is* credulously justified *iff* $\exists L \in \mathcal{L}_\sigma(\mathcal{F}) : L(A) = \mathtt{in}$*.*

On the basis of the quick review above, one may observe that traditional definitions in abstract argumentation are characterized by exhaustiveness (all the arguments in a framework are assigned a status by the semantics) and allow a unique form of undecidedness, corresponding to the `und` label. Both these constraints may turn out to be too rigid. On the one hand, exhaustiveness may be too demanding in practice, since it might be the case that providing an answer for all arguments is not always necessary. On the other hand, having a unique form of undecidedness may be regarded as poorly expressive, since the variety of cases where you don't have a definite answer may require a richer set of representation alternatives. We review and discuss approaches and ideas aiming at tackling these limitations in the next sections.

3 " I don't care! " Allowing for Incomplete Answers

Providing exhaustive answers is neither always a goal nor a necessity. In particular, there are several reasons why one may prefer not to evaluate the status of every argument. In the context of the reasoning or dialogical activity where the argumentation process is embedded, typically the actual goal is just assessing the status of a restricted subset of arguments, regarded as more important than the others[2], which are considered in the evaluation only if necessary for the main goal. Further, considering a dynamic context, where arguments and attacks may be modified on the fly, it may be the case that some parts of the framework are more subject to change than others and one may prefer, if possible, to restrict his/her evaluation to those parts of the framework which are regarded as "more stable", deferring the evaluation of other parts to a later moment and so avoiding to produce judgments probably needing a revision very soon. Two (not disjoint) motivations for partial evaluations emerge from these examples:

- saving computational resources by avoiding useless (i.e. unnecessary or too ephemeral) evaluations;
- even if computational resources are not an issue, avoiding to express positions beyond what is required (taking into account the goals and/or the dynamics of the process) according to a general criterion of *cautiousness* or *minimal commitment*.

[2] See, for instance, the notion of *desired set* introduced by Baumann and Brewka [10] in the context of the problem of argument enforcement.

Both motivations call for identifying some technically sound form of partial evaluation, though with different nuances that may have an effect on what kind of soundness is required. We review some literature approaches to *partial* argumentation semantics in the following subsections.

3.1 Using a *don't-care* Label

In [19] a labelling-based approach using the set of four labels $\Lambda^{JV} = \{+, -, \pm, \otimes\}$ is proposed. The first three labels correspond respectively to `in`, `out`, and `und` of the "traditional" Λ^{iou} set (the symbol $\pm$ indicating that both $+$ and $-$ are considered possible), while the fourth label[3] corresponds to a *don't-care* situation, namely to a non-assigned label. Indeed, a labelling L such that $L(\alpha) = \otimes$ for some argument α is called *partial* in [19], while a labelling L where $\nexists \alpha$ such that $L(\alpha) = \otimes$ will be called *total*.

In [19] the motivation for introducing a label corresponding to a *don't-care* situation is to have the possibility of not saying more than necessary, i.e. of not expressing any judgment concerning "arguments that are irrelevant or that do not interest the observer". This implies that, in principle, the choice of *don't-care* arguments is completely at the discretion of the agent carrying out the argumentation process. This freedom is however limited by the general legality constraints[4] on labellings based on the attack relation. In fact, according to [19, Definition 3] a (possibly partial) labelling L of an AF $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$ must satisfy the following conditions:

- $\forall \alpha \in \mathcal{A}$ if $L(\alpha) \in \{-, \pm\}$ then $\exists \beta \in \{\alpha\}^{\leftarrow}$ such that $L(\beta) \in \{+, \pm\}$;
- $\forall \alpha \in \mathcal{A}$ if $L(\alpha) \in \{+, \pm\}$ then $\forall \beta \in \{\alpha\}^{\leftarrow}$ $L(\beta) \in \{-, \pm\}$;
- $\forall \alpha \in \mathcal{A}$ if $L(\alpha) \in \{+, \pm\}$ then $\forall \beta \in \{\alpha\}^{\rightarrow}$ $L(\beta) \in \{-, \pm\}$.

While these rules do not mention explicitly the arguments labelled $\otimes$, they induce some constraints on them too. Intuitively, if you care about an argument, you should care also about some other arguments affecting or affected by it.

More precisely, $L(\alpha) = \otimes$ is possible only if the following conditions hold:

- $\forall \beta \in \{\alpha\}^{\leftarrow}$ $L(\beta) \in \{\otimes, -\}$;
- $\nexists \beta \in \{\alpha\}^{\rightarrow}$ such that $L(\beta) \in \{+, \pm\}$;
- $\forall \beta \in \{\alpha\}^{\rightarrow}$ if $L(\beta) = -$ then $\exists \gamma \in \{\beta\}^{\leftarrow} \setminus \{\alpha\}$ such that $L(\gamma) \in \{+, \pm\}$.

The first condition states that one can not abstain on an argument which has at least one attacker labelled $+$ or $\pm$. This evidences a sort of asymmetry in the approach of [19]: one can abstain on an argument that would otherwise be labelled $+$, but can not abstain on an argument that would otherwise be labelled $-$ or $\pm$. The second and third conditions concern the cases where one abstains on an argument α but not on (some of) the arguments attacked by α. More specifically, the second condition forbids any abstention on the attackers of an

[3] In [19] $\emptyset$ is used for the fourth label, we avoid the use of this overloaded symbol.

[4] Note that the legality constraints of [19] do not coincide with the "standard" legality constraints recalled in Definition 6.

argument labelled + or ±, while the third condition allows the abstention on an attacker of an argument β labelled − only if there is another attacker γ, labelled + or ±, justifying the label − of β.

The above rules imply in particular that "full carelessness" (i.e. a labelling L such that $L(\alpha) = \otimes$ for every argument α) is always possible, but partial carelessness is not arbitrary.

An additional observation concerns the relationship between abstention and the potential completions of a partial labelling L. A total labelling L' is a completion of a partial labelling L if $L(\alpha) \neq \otimes \Rightarrow L'(\alpha) = L(\alpha)$, i.e. if L' is obtained from L by replacing all and only the $\otimes$ labels with other labels (taking into account the legality constraints). In general, a legal partial labelling admits several different completions (and always admits at least one). It can be the case that a *don't-care* argument α gets the same label in all the possible completions of a partial labelling L, i.e. that the (only) legal label of α is univocally determined by the information carried by L. Still, according to [19], it is legal to abstain on α. This confirms that the $\otimes$ label does not correspond, *per se*, to any notion of indecision and is applicable to some arguments with an (implicitly) well determined label too.

While a further detailed discussion of the approach in [19] is beyond the scope of the present paper, the features discussed above will be enough to point out the basic differences with other notions of partial argumentation semantics reviewed in the next subsection.

3.2 Partial Semantics for Partial Computation

Given an *AF* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$, let $S \subset \mathcal{A}$ be a set of arguments which are of some interest for an agent (in a sense, they are the complement of the *don't-care* arguments mentioned in the previous section). In order to derive the justification status of the interesting arguments one may wonder whether it is necessary to preliminarily carry out a computation involving the whole *AF*, i.e. to first compute $\mathcal{E}_\sigma(\mathcal{F})$ or $\mathcal{L}_\sigma(\mathcal{F})$, or it is sufficient to carry out a partial computation involving only S and those other parts of the framework affecting the evaluation of the arguments in S. Given that most computational problems in abstract argumentation are intractable, reducing the set of arguments and attacks considered in the derivation of the desired outcomes is of great interest, since it may yield significant savings of computational resources. This calls for a suitable notion of partial semantics applicable to the restrictions of a framework.

This notion is also crucial in the area of argumentation dynamics, namely in contexts where the considered *AF* is subject to modifications over time. If these modifications affect only a part of the whole framework, there is the opportunity to reuse previously computed results concerning the part of the framework unaffected by modifications, instead of reevaluating the whole framework from scratch. Similar issues also arise from related investigation lines in abstract argumentation, like the study of incremental algorithms for argumentation dynamics [22,8] and of multi-sorted reasoning [24].

Defining a partial semantics involves dealing with two interplaying notions: on the one hand, one has to devise suitable ways to restrict a framework to subframeworks which are appropriate for the definition of partial semantics, on the other hand, one has to identify suitable semantics properties ensuring that the relation between local and global semantics evaluation is sound.

As to defining restrictions, given an AF $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$ and a set of arguments $S \subset \mathcal{A}$ a straightforward way to define the restriction of $\mathcal{F}$ to S is to suppress all arguments in $\mathcal{A} \setminus S$ and all attacks involving at least one argument not in S. Accordingly, the restriction $\mathcal{F}\downarrow_S$ of $\mathcal{F}$ to S is defined as $\mathcal{F}\downarrow_S = \langle S, \rightarrow \cap (S \times S) \rangle$.

This definition appears rather rough as it "cuts" all the links between S and other arguments: indeed it ignores $S^{\leftarrow} \cap (\mathcal{A} \setminus S)$ and $S^{\rightarrow} \cap (\mathcal{A} \setminus S)$. In spite of this, such definition turns out to be very useful under a suitable choice of S and of the semantics σ to be "partialized". First, one can simply choose S such that it is *unattacked*, namely such that $S^{\leftarrow} \cap (\mathcal{A} \setminus S) = \emptyset$. Second, one may focus on semantics featuring the *directionality* property [6], namely such that the evaluation of an unattacked set is not affected by the remaining parts of the framework. The relevant formal definitions are recalled below.

Definition 8. *Given an AF* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$*, a set* $S \subseteq \mathcal{A}$ *is* unattacked *iff* $S^{\leftarrow} \cap (\mathcal{A} \setminus S) = \emptyset$*. The set of unattacked sets of* $\mathcal{F}$ *is denoted as* $\mathcal{US}(\mathcal{F})$*.*

Definition 9. *An extension-based semantics* σ *satisfies the directionality criterion iff* $\forall \mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle, \forall S \in \mathcal{US}(\mathcal{F}), \mathcal{AE}_\sigma(\mathcal{F}, S) = \mathcal{E}_\sigma(\mathcal{F}\downarrow_S)$*, where* $\mathcal{AE}_\sigma(\mathcal{F}, S) \triangleq \{(E \cap S) \mid E \in \mathcal{E}_\sigma(\mathcal{F})\} \subseteq 2^S$*. A labelling-based semantics* σ *with label set* Λ *satisfies the directionality criterion iff* $\forall \mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle, \forall S \in \mathcal{US}(\mathcal{F}), \mathcal{AL}_\sigma(\mathcal{F}, S) = \mathcal{L}_\sigma(\mathcal{F}\downarrow_S)$*, where* $\mathcal{AL}_\sigma(\mathcal{F}, S) \triangleq \{L \cap (S \times \Lambda) \mid L \in \mathcal{L}_\sigma(\mathcal{F})\}$*.*

Under the above mentioned assumptions, a notion of partial semantics useful for partial and incremental computation has been introduced in [21]. Basically, given a set of arguments of interest S, the semantics evaluation is carried out on the restriction of $\mathcal{F}$ to the minimal unattacked set including S.

Definition 10. *Given an AF* $\mathcal{F} = \langle \mathcal{A}, \rightarrow \rangle$ *and a set of arguments* $S \subseteq \mathcal{A}$*, define* $rlvt_\mathcal{F}(S) = \min_\subseteq \{U \mid S \subseteq U \wedge U \in \mathcal{US}(\mathcal{F})\}$*. Given an extension-based (labelling-based) semantics* σ *satisfying the directionality criterion the partial semantics of* $\mathcal{F}$ *with respect to* S *is defined as* $\mathcal{E}_\sigma(\mathcal{F}\downarrow_{rlvt_\mathcal{F}(S)})$ $(\mathcal{L}_\sigma(\mathcal{F}\downarrow_{rlvt_\mathcal{F}(S)}))$*.*

The restriction to an unattacked set for a directional semantics has been (often implicitly) exploited as a starting point in works oriented towards incremental computation, like splitting argumentation frameworks [9], the division-based method [22] for argumentation dynamics and the decomposition-based approach [20]. In these contexts a further step towards a richer notion of partial semantics is made by considering the restriction to a set S which is not unattacked and receives some fixed influence from outside, formally this amounts to remove the assumption that $S^{\leftarrow} \cap (\mathcal{A} \setminus S) = \emptyset$, while still ignoring $S^{\rightarrow} \cap (\mathcal{A} \setminus S)$.

This has led to various notions of conditioned AF in the literature, where basically a conditioned AF is a framework receiving some attacks from a conditioning AF. In general, the conditioned and conditioning frameworks are obtained by partitioning a global framework according to some criterion. For instance, in a dynamic context, the conditioning framework corresponds to the part of the original framework which is not affected by a modification, so that previous computation results concerning this part can be reused for the new semantics evaluation concerning the affected part, corresponding to the conditioned framework. We recall here the relevant definitions from [22].

Definition 11. *Given an AF* $\mathcal{F}_1 = \langle \mathcal{A}_1, \rightarrow_1 \rangle$*, a conditioned AF with respect to* $\mathcal{F}_1$ *is a tuple* $\mathcal{CAF} = (\langle \mathcal{A}_2, \rightarrow_2 \rangle, (\mathcal{C}(\mathcal{A}_1), \mathcal{I}_{(\mathcal{C}(\mathcal{A}_1), \mathcal{A}_2)}))$ *in which*

- $\mathcal{F}_2 = \langle \mathcal{A}_2, \rightarrow_2 \rangle$ *is an AF that is conditioned by* $\mathcal{C}(\mathcal{A}_1)$ *in which* $\mathcal{A}_2 \cap \mathcal{A}_1 = \emptyset$*;*
- $\mathcal{C}(\mathcal{A}_1) \subseteq \mathcal{A}_1$ *is a nonempty set of arguments (called conditioning arguments) that have interactions with arguments in* $\mathcal{A}_2$*, i.e.,* $\forall \alpha \in \mathcal{C}(\mathcal{A}_1), \exists \beta \in \mathcal{A}_2$*, such that* $(\alpha, \beta) \in \mathcal{I}_{(\mathcal{C}(\mathcal{A}_1), \mathcal{A}_2)}$*;*
- $\mathcal{I}_{(\mathcal{C}(\mathcal{A}_1), \mathcal{A}_2)} \subseteq \mathcal{C}(\mathcal{A}_1) \times \mathcal{A}_2$ *is the set of interactions from the arguments in* $\mathcal{C}(\mathcal{A}_1)$ *to the arguments in* $\mathcal{A}_2$*.*

Semantics directionality still plays a crucial role in this context: the idea is that extension (labelling) computation in $\mathcal{F}_2$ depends on $\mathcal{F}_1$ but not vice versa (since $\mathcal{F}_1$ does not receive attacks from $\mathcal{F}_2$), hence one can use the extensions (labellings) of $\mathcal{F}_1$ as fixed conditions to determine the extensions (labellings) of $\mathcal{F}_2$ (the reader is referred to [22] for details). It must however also be stressed that in this enriched context directionality alone is no more sufficient to ensure that local semantics definitions at the local level are coherent with those at the global level, i.e. that combining the results of local evaluations one obtains the same outcomes of global evaluation. In particular, the role of the SCC-recursiveness property [7] in this context has been pointed out in [8].

Recently, a further generalization in the study of partial argumentation semantics has been achieved [3,4] by considering arbitrary partitions of an *AF* into subframeworks that, differently from the cases reviewed above, can be in a relation of mutual dependence[5]. In this context, a partition induces a set of subframeworks, each of which can be regarded as an *AF* receiving inputs (through some attacks) from other subframeworks and in turn feeding inputs to other subframeworks through other attacks. Modeling each of these subframeworks as an *argumentation framework with input*, it has been possible to identify a *canonical local function* [4] representing the counterpart at the local level of the semantics definition at a global level, under very mild requirements satisfied by most argumentation semantics in the literature.

It turns out however that combining the outcomes of the canonical local function of a semantics σ applied to the subframeworks does not always yield the same results obtained by applying σ at the global level. In other words, not every semantics is *decomposable* with respect to arbitrary partitions of an *AF*: this result poses a theoretical limit to the possibility of defining a partial notion of semantics preserving the same meaning as a global one. Accordingly, an interesting issue consists in identifying some restricted classes of partitions (e.g.

[5] The use of arbitrary partitions is called parameterized splitting in [11].

those based on the graph-theoretical notion of strongly connected components) where decomposability is recovered (the reader is referred to [4] for some relevant results). A further research direction than can benefit from a generalized notion of local evaluation is *multi-sorted argumentation* [24], namely the study of the application of different semantics to different parts of a framework.

3.3 Discussion

While the approach in [19] represents explicitly the notion of *don't-care* arguments with a specific label, all the approaches reviewed in section 3.2 use some restriction of the framework to focus attention on some set of arguments which, for some reasons, deserves to be considered separately. Both *don't-care* arguments and (most of) the restriction mechanisms have to obey some constraints and, to the best of our knowledge, their relations have not been investigated yet in the literature. As a preliminary observation it can be noted that constraints on *don't-care* arguments take directly into account the effect that ignoring an argument has on other arguments, while the restriction mechanisms typically considered in the literature take this effect into account indirectly through some graph-theoretical properties (e.g. the one of being an unattacked set). It follows that contraints referred to restriction mechanisms can be more limiting than those expressed in terms of *don't-care* arguments. To exemplify this, consider the simple framework $\mathcal{F}_1 = \langle\{\alpha, \beta, \gamma\}, \{(\alpha, \gamma)(\beta, \gamma)\}\rangle$. Here the status of γ can be determined by considering only one of its attackers (indifferently α or β) and in fact the following labellings are legal according to [19]: $L_1 = \langle(\alpha, +), (\beta, \otimes), (\gamma, -)\rangle$, $L_2 = \langle(\alpha, \otimes), (\beta, +), (\gamma, -)\rangle$, which means (correctly) that one can focus on either $\mathcal{F}_1 \downarrow_{\{\alpha,\gamma\}}$ or $\mathcal{F}_1 \downarrow_{\{\beta,\gamma\}}$ without losing any information about the status of γ. However, $\{\alpha, \gamma\}$ and $\{\beta, \gamma\}$ are not unattacked sets in $\mathcal{F}_1$, thus none of the restriction mechanisms considered in subsection 3.2 would allow this: they would either force the inclusion of the missing argument or take into account it as an input, while (in this specific case) this is, in fact, unnecessary. It must be said however that constraints concerning *don't-care* arguments concern local attack relations only, while the restriction mechanisms provide a direct way to select suitable partitions of a framework at a global level.

This suggests that combining *don't-care* arguments with restriction mechanisms may yield more advanced notions of partial semantics with respect to the state of the art. This appears a very interesting direction of future research: in particular, this combined approach may gain additional efficiency improvements by providing better solutions to the problem of identifying the minimal amount of computation sufficient to ensure that the status of a given set of interesting arguments is the same as the one resulting from a computation over the whole framework. In this perspective interesting relations may be drawn with the notions of *argumentation multipoles* [4] and of *critical sets* [17].

As a final note, since partial semantics notions can be considered also in extensions of the traditional Dung's AFs (in particular in Abstract Dialectical Frameworks [14]), considering the use of *don't-care* arguments in these extended formalisms represents another interesting line of future work.

4 " I don't know . . . I know too much " Variations of Undecidedness

4.1 Undecidedness Is Not All the Same

Even if you care about an argument, you may be unable to assign it a definite acceptance status (`in` or `out` using the Λ^{iou} label set, + or − using Λ^{JV}) and must be content with an intermediate status (`und` or ± respectively) representing some form of indecision. In both Λ^{iou} and Λ^{JV} the intermediate label is meant to represent every form of indecision, but one might observe that the reasons to be undecided can be rather different.

On the one hand, one may be undecided because s/he has no enough information to express a definite judgment and needs to wait for further information to arrive. For instance, if asked about whether it will rain tomorrow, you may have no hint at all, reply "I don't know", and then look for weather forecasts on the web. After surfing several weather web sites, however, you are not guaranteed to have gained a definite position, because some of them may promise a sunny day, while others presage thunderstorms. In this case, your indecision is still there but has changed nature since it is due to contradiction rather than to ignorance. Indeed, as suggested in [16], your reply should now be "I know too much", since you got an excess of (inconsistent) information.

Distinguishing these two kinds of indecision is the cornerstone of Belnap-Dunn (BD) four-valued logic [12]. BD-logic is based on the assumption that an information-providing agent has two basic moves available (namely asserting that a given statement is true or asserting that it is false) and that the status of a statement then results from the union of all the moves concerning it. So, if no move at all has been done, its status corresponds to indecision by ignorance (`N`: "neither told true nor told false"), if only positive or negative moves have been done the statement has a definite status (`T`: "told true" or `F`: "told false" respectively), if both positive and negative moves have been done, one gets indecision by contradiction (`B`: "both told true and told false"). The set of BD truth values is then $\Lambda^{BD} = \{\mathtt{N}, \mathtt{T}, \mathtt{F}, \mathtt{B}\}$.

The use in abstract argumentation of a set of labels $\Lambda^4 = \{\mathtt{none}, \mathtt{in}, \mathtt{out}, \mathtt{both}\}$ corresponding to the four truth values in Λ^{BD} has recently been proposed by Arieli [1], in the context of a conflict-tolerant approach to semantics definition, where the requirement of conflict-freeness for extensions/labellings is relaxed, in order to achieve non-conventional results in the handling of attack loops.

In this way a correspondence between extensions and Λ^4-labellings is obtained as follows. Given an extension E and an argument α:

– α is labelled `in` iff $\alpha \in E \wedge \alpha \notin E^{\rightarrow}$;
– α is labelled `out` iff $\alpha \notin E \wedge \alpha \in E^{\rightarrow}$;
– α is labelled `none` iff $\alpha \notin E \wedge \alpha \notin E^{\rightarrow}$;
– α is labelled `both` iff $\alpha \in E \wedge \alpha \in E^{\rightarrow}$.

Note that the fourth case is possible only if E is not conflict-free. While there is a formal correspondence and some intuitive analogy between the four labels in Λ^4 and the four truth values of BD-logic, it has to be remarked that they are conceptually different as they lie at different stages of the reasoning process.

The BD-model can be regarded as basically consisting of three phases:

1. *assertion production*: where agents make assertions by associating truth values to propositions;
2. *aggregation*: where different assertions concerning the same sentence are "put together" (by a simple union operation) yielding a four-valued labelling of propositions;
3. *use of aggregation outcomes*: where reasoning about labelled propositions is carried out (e.g. given two propositions p_1, labelled B, and p_2, labelled N, BD-logic specifies the truth value of $p_1 \vee p_2$, $p_1 \wedge p_2$, and so on ...).

To draw a comparison, also argumentation-based reasoning, called AB-model in the following, can be schematized in three phases:

1. *argument production*: where agents produce (possibly conflicting) arguments each supporting some conclusion;
2. *conflict management*: where semantics evaluation is applied to the set of arguments and attacks yielding a set of labellings of arguments;
3. *use of conflict outcomes*: where argument conclusions are evaluated on the basis of argument labels (note that the same conclusion can be supported by many arguments) and further reasoning is possibly carried out based on these evaluations.

The two models feature several structural similarities. Assertions in the BD-model can be regarded as a special kind of arguments following the generic scheme "If an agent tells that a given proposition p has a truth value v then there is a reason to believe that p has the truth value v." Hence the assertions correspond to the conclusions of the arguments. The aggregation in the BD-model can be regarded as a special kind of conflict management. The basic idea is that conflict arises when different truth values are asserted for the same proposition and that every conflict gives rise to "indecision by contradiction". In terms of Dung's theory, this amounts to consider the special case where only symmetric conflicts are present and to adopt a sceptical semantics (in particular the *grounded* semantics) for the evaluation of arguments. This in particular implies that only one labelling of arguments exists where all non conflicting arguments are accepted and all conflicting arguments are undecided. Then the conclusions of the accepted arguments get exactly the truth value that was asserted in the first phase, while the conclusions of conflicting arguments get the B value and the propositions not supported by any assertion/argument keep the N value. Due to its simplicity, conflict management is left implicit in the BD-model, which, in the aggregation phase, jumps directly to the assignment of truth values to propositions. Differently, conflict management between arguments is the focus of abstract argumentation theory, where semantics evaluation concerns assigning labels to *arguments* not to *conclusions*, while the step of evaluating conclusions and reasoning about the outcomes, namely the third process phase, is completely left out of the theory. Differently, this last phase is the main subject (not reviewed here) of the BD-model.

According to the analysis carried out above, we can identify three different "labelling" activities during the reasoning process.

First, *propositions* are labelled with *truth values*. This labelling is explicit in the assertion production phase of the BD-model, hence we will call these truth values *assertible values*. In the AB-model, this labelling corresponds to the contents of *argument conclusions* in the phase of argument production. These aspects are abstracted away and hence left implicit in abstract AFs.

Second, for the sake of conflict resolution, *arguments* are labelled with *acceptance values*. This activity is explicit in the conflict management phase of the AB-model (where acceptance values are called labels *tout court*), while it is left implicit in the BD-model.

Third, taking into account the results of conflict resolution, *argument conclusions*, i.e. *propositions*, are labelled with *aggregated conflict outcome values*. These values are produced in the aggregation phase in the BD-model, whose four-valued logic specifies then how to reason with them. On the other hand, in the AB-model they are regarded as a by-product of argument evaluation and, to the best of our knowledge, reasoning with them has received, by far, lesser attention in the literature.

The analysis carried out above evidences first of all the different nature of the uses of the "same" four values in the BD-model and in [1]. In the BD-model they are associated with *propositions/argument conclusions* and represent *aggregated conflict outcome values*, in [1] they are associated with *arguments* and represent *acceptance values*.

More interestingly, it points out some opportunities of cross-fertilization between these research areas.

On the one hand, the BD-model provides an advanced logic for reasoning about aggregated conflict outcomes which, to the best of our knowledge, has no parallel in the argumentation literature and could be used as a starting point to fill this significant gap in existing models of argumentation-based reasoning processes. Further, the use of a richer set of labels than Λ^{iou}, like Λ^4 in [1], promises a significant increase in the expressiveness (but also complexity) of labelling-based argumentation semantics, whose implications can be regarded as a largely unexplored research avenue. Moreover, since the notion of *don't-care* arguments encompassed in the Λ^{JV} set of labels is "orthogonal" to the distinction between `none` and `both` encompassed by the Λ^4 set, one might investigate the combination of the two ideas by considering a set of labels $\Lambda^5 = \{\texttt{none}, \texttt{in}, \texttt{out}, \texttt{both}, \otimes\}$.

On the other hand, the BD-model, initially conceived for the management of inconsistent inputs by a computer system [12] and recently considered as an approach to address the problem of inconsistent information on the Web [16], appears to rely on a very simple implicit argumentation model, using just one argumentation scheme for assertion production and an implicit skeptical semantics for argument acceptance evaluation. As a variety of more articulated models for argument construction [13] and evaluation [5] are available in the literature, using them to enrich the BD-model is a natural direction of investigation.

Leaving the development of these suggestions to future consideration, we focus in next subsection on another more fundamental issue concerning the modelling of undecidedness.

4.2 Epistemological Undecidedness

As discussed in the previous subsection, the BD-model assumes that the basic assertions an agent can make are binary: the set $\mathcal{AV}$ of *assertible values* is $\mathcal{AV} = \{\mathtt{T}, \mathtt{F}\}$ and the four labels in Λ^{BD} arise from the aggregation of multiple moves (or no move at all). In fact they correspond to all the possible subsets (including the empty one) of the set $\{\mathtt{T}, \mathtt{F}\}$. This model appears to be based on the assumption that the notion of a definite belief (corresponding to T and F) is more basic than the one of undecidedness, which is a derived concept: something which is undecided could (and, in a sense, should) be T or F in the end, but the lack of information or the presence of unresolved contradictory information prevents a more definite position.

It may be observed however that this modeling stance is somehow restricted and could be generalized. From a purely formal point of view, one may consider the case where the agents making the basic assertions adopt a richer set of assertible values. Following the BD-model scheme, this would give rise in turn to a richer set of *aggregated conflict outcome values*, since they correspond to the elements of $2^{\mathcal{AV}}$. To motivate this extension from a conceptual point of view, one may suggest the existence of an additional, more "fundamental", case of undecidedness, called *epistemological undecidedness* in the sequel. To provide a case for this, consider again the example of the weather forecast and suppose that the location you are interested in lies in a region with a specially complex geography, such that no existing weather forecast model is applicable. Then, you are undecided about whether tomorrow will be sunny (indeed you have good fundamental reasons to be so) and this indecision is rather different from the ones considered above. First, it clearly does not arise from contradictory information: you can not certainly say "I know too much". Second, even if it bears some superficial resemblance with the case "I don't know" represented by the truth value N in BD-logic, it is really different. The truth value N is meant to represent absence of information, i.e. no move at all by an agent, and can not conflict with a subsequent move: for instance if another agent makes a positive assertion then N is directly superseded by T. Epistemological undecidedness, instead, relies on some information and corresponds to a kind of move not encompassed by the models reviewed above: it may be represented by an additional assertible value U! corresponding to the intuitive answer: "I know that it is impossible to know". As a consequence, epistemological undecidedness can actually conflict with moves of other kinds. In the example, if one says that tomorrow will be sunny, your position will not be superseded and you may object to this assertion, even if you don't assert that it will be rainy.

One might think that objections based on epistemological undecidedness are analogous to the undercutting attacks exemplified by the famous Pollock's "red light" example [23]. In a nutshell, since an object looks red to you, you derive that it is red, but when you learn that the object is under a red light, your derivation is undercut and your reason to believe that the object is red is defeated, while still leaving open the possibility that it is actually red. In both epistemological undecidedness and Pollock's undercut, an objection is raised not by asserting the contrary of a given statement but providing reasons to leave it undecided. However there is a basic difference in the reasons of being undecided in the two cases. In Pollock's example the reason to be undecided is specific to the way the conclusion that the object is red has been derived. Knowing that the object is under a red light does not imply that you can not know whether the object is red, but only that you can not get to know it by looking at the object. For instance, if you have an old picture of the object under normal light, you get a new argument for which this specific undercut is no more effective (while, of course, other undercuts may arise). Thus Pollock's undercut is coherent with the view that indecision is due to the unability to definitely accept or reject a statement, while epistemological undecidedness means that you have reasons to regard a topic as unknowable independently of the way different positions can be derived. If you have reasons to believe that there is no way to forecast weather in a given location, then you are in conflict with any weather forecast, independently of the way it is derived. In this sense, an attack based on epistemological undecidedness can be seen as an additional form of rebut. The standard notion of rebut is based on the set of assertible values $\{\mathtt{T}, \mathtt{F}\}$ and a rebutting attack arises when different values are asserted for the same sentence, independently of the way they are derived. When extending the set of assertible values to $\mathcal{AV}^3 = \{\mathtt{T}, \mathtt{F}, \mathtt{U!}\}$, the notion of rebut remains the same, i.e. that different values are asserted for the same sentence, independently of the way they are derived, but there is a larger variety of rebut situations: not just T vs. F but also, U! vs. T, U! vs. F, and possibly even a three-way duel[6] U! vs. T vs. F.

A research agenda to encompass epistemological undecidedness into a BD-inspired AF model can then be drafted.

First, a suitable argument generation logic encompassing the extended set of assertible values $\mathcal{AV}^3 = \{\mathtt{T}, \mathtt{F}, \mathtt{U!}\}$ has to be investigated. For the sake of exploring the implications of the adoption of $\mathcal{AV}^3$ at a more abstract level one could consider a simple BD-like model where agents can make three kinds of assertions about a sentence.

Second, an abstract framework to represent the attack relations between arguments has to be identified. Traditional AFs encompass a unique kind of binary attack relation, but its expressiveness is probably insufficient in the extended context. First, one may wonder whether attacks involving arguments based on U! assertions against arguments involving arguments based on T or F assertions should be classified and treated differently from "traditional" attacks involving T vs. F assertions. Further, the distinction between rebutting and undercutting

[6] Like in the classic non-classical western movie *The Good, the Bad and the Ugly.*

attacks may need to be reassessed in this context. As a bottom ground for this representation, one may consider again a BD-like model where all attacks are symmetric and only unattacked arguments are accepted. A further research jump would be to consider the above issues in the contexts of abstract dialectical frameworks [14], where generic influence relations among arguments, rather than just attacks, are considered.

Finally, turning to reasoning about conflict outcomes for the propositions of interest, an extended logic would be needed for the *aggregated conflict outcome values*, which, following the line of the BD-model, might correspond to the element of $2^{\mathcal{AV}^3}$. To this purpose existing studies on bilattice-based generalizations of BD-logic could be taken as starting point [18,2].

5 Conclusions

We believe that sketching a few fluid research directions for the future is a suitable way to celebrate many solid research results achieved in the past. In this spirit, in this work we have analyzed and discussed some "non-mainstream" aspects of the treatment of incompleteness and undecidedness in argumentation, with the aim of posing questions rather than of giving answers. Whether these and similar matters represent just theoretical curiosities or will somehow contribute to narrow the gap between human reasoning and its formal models is an issue for next generations of researchers. For sure, their work will profit from the rich and still increasing conceptual and technical asset built by outstanding researchers like Gerhard Brewka, to whom the book including this chapter is dedicated.

Acknowledgments. The authors thank the anonymous reviewer for his/her helpful comments. This work was conceived and developed during a visit of prof. Beishui Liao to the University of Brescia supported partially by the National Natural Science Foundation of China under grant No. 61175058 and Zhejiang Provincial Natural Science Foundation of China under grant No. LY14F030014.

References

1. Arieli, O.: On the acceptance of loops in argumentation frameworks. J. of Logic and Computation (to appear, 2014)
2. Arieli, O., Avron, A.: The value of the four values. Artif. Intell. 102, 97–141 (1998)
3. Baroni, P., Boella, G., Cerutti, F., Giacomin, M., van der Torre, L.W.N., Villata, S.: On input/output argumentation frameworks. In: Proc. of the 4th Int. Conf. on Computational Models of Argument (COMMA 2012), pp. 358–365 (2012)
4. Baroni, P., Boella, G., Cerutti, F., Giacomin, M., Torre, L.W.N.v.d., Villata, S.: On the input/output behavior of argumentation frameworks. Artif. Intell. 217, 144–197 (2014)
5. Baroni, P., Caminada, M., Giacomin, M.: An introduction to argumentation semantics. Knowledge Engineering Review 26(4), 365–410 (2011)

6. Baroni, P., Giacomin, M.: On principle-based evaluation of extension-based argumentation semantics. Artif. Intell. 171(10/15), 675–700 (2007)
7. Baroni, P., Giacomin, M., Guida, G.: SCC-recursiveness: a general schema for argumentation semantics. Artif. Intell. 168(1-2), 165–210 (2005)
8. Baroni, P., Giacomin, M., Liao, B.: On topology-related properties of abstract argumentation semantics. A correction and extension to *Dynamics of argumentation systems: A division-based method*. Artif. Intell. 212, 104–115 (2014)
9. Baumann, R.: Splitting an argumentation framework. In: Delgrande, J.P., Faber, W. (eds.) LPNMR 2011. LNCS, vol. 6645, pp. 40–53. Springer, Heidelberg (2011)
10. Baumann, R., Brewka, G.: Expanding argumentation frameworks: Enforcing and monotonicity results. In: Proc. of the 3rd Int. Conf. on Computational Models of Argument (COMMA 2010), pp. 75–86 (2010)
11. Baumann, R., Brewka, G., Dvořák, W., Woltran, S.: Parameterized splitting: A simple modification-based approach. In: Erdem, E., Lee, J., Lierler, Y., Pearce, D. (eds.) Correct Reasoning. LNCS, vol. 7265, pp. 57–71. Springer, Heidelberg (2012)
12. Belnap, N.D.: How a computer should think. In: Ryle, G. (ed.) Contemporary aspects of philosophy, pp. 30–56. Oriel Press (1977)
13. Besnard, P., Garcia, A., Hunter, A., Modgil, S., Prakken, H., Simari, G., Toni, F.: Special issue: Tutorials on structured argumentation. Argument & Computation 5(1) (2014)
14. Brewka, G., Woltran, S.: Abstract dialectical frameworks. In: Proc. of the 12th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2010), pp. 102–111 (2010)
15. Dung, P.M.: On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. Artif. Intell. 77(2), 321–357 (1995)
16. Dunn, J.M.: Contradictory information: Too much of a good thing. J. of Philosophical Logic 39, 425–452 (2010)
17. Gabbay, D.M.: Fibring argumentation frames. Studia Logica 93(2-3), 231–295 (2009)
18. Ginsberg, M.L.: Multivalued logics: a uniform approach to reasoning in AI. Computer Intelligence 4, 256–316 (1988)
19. Jakobovits, H., Vermeir, D.: Robust semantics for argumentation frameworks. J. of Logic and Computation 9(2), 215–261 (1999)
20. Liao, B.: Toward incremental computation of argumentation semantics: A decomposition-based approach. Ann. Math. Artif. Intell. 67(3-4), 319–358 (2013)
21. Liao, B., Huang, H.: Partial semantics of argumentation: basic properties and empirical results. J. of Logic and Computation 23(3), 541–562 (2013)
22. Liao, B., Jin, L., Koons, R.C.: Dynamics of argumentation systems: A division-based method. Artif. Intell. 175(11), 1790–1814 (2011)
23. Pollock, J.: How to reason defeasibly. Artif. Intell. 57, 1–42 (1992)
24. Rienstra, T., Perotti, A., Villata, S., Gabbay, D.M., van der Torre, L.: Multi-sorted argumentation. In: Modgil, S., Oren, N., Toni, F. (eds.) TAFA 2011. LNCS, vol. 7132, pp. 215–231. Springer, Heidelberg (2012)