Positiveness and Skepticism in Abstract Argumentation: A First Approach

Pietro Baroni[1], Federico Cerutti[1,2] and Massimiliano Giacomin[1,*]

[1]Department of Information Engineering (University of Brescia), Italy
[2]Cardiff University, UK

Abstract

Starting from previous work concerning the notions of consistency and reinstatement in an abstract labelling setting, we propose an enhanced model where the above mentioned notions are integrated with the one of skepticism. Skepticism makes it possible to identify the labellings prescribed by argumentation semantics involving global requirements among those satisfying a combination of local consistency and reinstatement requirements. We then consider the issue of determining argument justification by synthesizing an evaluation labelling from those prescribed by the semantics, and we analyze the evaluation function most commonly used in the literature against a number of possible desiderata. Overall, we obtain a general model, based on the foundational notions of positiveness and skepticism, able to capture a variety of instances of different reasoning stages in abstract argumentation.

Keywords

Consistency, Reinstatement, Skepticism, Argumentation semantics, Argument justification

1. Introduction

In the context of Dung's abstract argumentation [1], argumentation semantics deal with conflicts between arguments by identifying a set of extensions or labellings indicating which sets of arguments are collectively acceptable and providing the basis to assess the justification status of each argument. In [2] it has been shown that traditional semantics notions can be equivalently expressed in terms of two dual properties, called consistency and reinstatement. The idea behind consistency is that conflicting arguments should not be accepted together. Since the labels assigned to arguments reflect different *positiveness* degrees, consistency can be modeled by means of a binary relation aimed at forbidding the pairs of labels that correspond to a simultaneous excess of positiveness. On the other hand, reinstatement corresponds to a sort of completeness condition aimed at avoiding abstention from justifying arguments. This can be modelled by a binary relation between labels that allows a negative label for an argument only if there is a sufficiently positive label assigned to one of its attackers.

As proved in [2], some combinations of specific instances of generalized consistency and reinstatement properties correspond to different kinds of labelling, including conflict-free,

AI^3 2023: 7th Workshop on Advances in Argumentation in Artificial Intelligence

*Corresponding author.
pietro.baroni@unibs.it (P. Baroni); federico.cerutti@unibs.it (F. Cerutti); massimiliano.giacomin@unibs.it (M. Giacomin)
0000-0001-5439-9561 (P. Baroni); 0000-0003-0755-0358 (F. Cerutti); 0000-0003-4771-4265 (M. Giacomin)

admissible, complete and stable labellings. However, other labellings cannot be characterized by just the local constraints enforcing consistency and reinstatement, since their definition explicitly or implicitly involves conditions at global level. For instance, the grounded labelling corresponds to the complete labelling minimizing the set of positively assessed arguments, while the preferred labellings maximize them. To fill this gap, we resort in this paper to a formal notion of *skepticism*, based on the observation that grounded and preferred semantics can be put in correspondence with different attitudes in justifying arguments. Our analysis on the role of skepticism is then extended to the assessment of argument justification status, which is derived from the extensions or labellings prescribed by the adopted argumentation semantics.

Altogether, this paper aims to provide a first framework where positiveness and skepticism are integrated, drawing some conceptual considerations and identifying some perspectives for further work.

The paper is organized as follows. Section 2 provides some background on the notion of consistency and reinstatement, while Section 3 proposes a model where they are integrated with a notion of skepticism, thus allowing the characterization of grounded, preferred and semi-stable semantics. Section 4 deals with argument justification in the proposed unifying setting. A number of intuitive desiderata based on the notions of positiveness and skepticism is proposed, and the traditional approach to assess argument justification is analyzed against them, pointing out several limitations. Finally, Section 5 draws some concluding remarks and discusses future avenues of research.

2. Consistency and reinstatement

To provide a general characterization of labelling-based assessments of entities of various kind, we have introduced in [3] a three-layer model that is briefly described below.

At the top level, the notion of assessment classes provides a reference point to characterize different assessment labels. These classes have an underlying order reflecting a level of positiveness of the assessment, with $c_1 \leq c_2$ meaning that c_2 corresponds to an at least as positive assessment as c_1.

Definition 1. *A set of assessment classes (abbreviated as sac(s) in the following) is a set C equipped with a total order $\leq$ (i.e. a reflexive, transitive and antisymmetric relation such that any two elements are comparable) and including a maximum and a minimum element (i.e. an element $c \in C$ such that $\forall c' \in C$ it holds that $c' \leq c$ or $c \leq c'$, respectively) which are assumed to be distinct.*

At an intermediate level, assessment labels are taken from a predefined set and classified on the basis of a sac, thus inheriting the relevant positiveness degree.

Definition 2. *Given a set of assessment classes C, a C-classified set of assessment labels is a set Λ equipped with a total function $C_\Lambda : \Lambda \to C$. The total preorder induced on Λ by C_Λ, also called positiveness preorder, will be denoted by $\preceq$ where $\lambda_1 \preceq \lambda_2$ iff $C_\Lambda(\lambda_1) \leq C_\Lambda(\lambda_2)$. As usual, $\lambda_1 \prec \lambda_2$ will denote $\lambda_1 \preceq \lambda_2$ and $\lambda_2 \not\preceq \lambda_1$.*

The fact that $\preceq$ is a total preorder is shown in [2]. We will abbreviate the term 'set(s) of assessment labels' as sal(s) and omit 'C-classified', when C is not ambiguous. Also, to distinguish preorders referring to different sals, given a sal Λ we will denote the relevant preorder as $\preceq_\Lambda$.

At the bottom level, a generic set of entities is considered, with the entities related by an intolerance relation, indicating who cannot stand whom (as an example, this relation might coincide with classical negation if the considered language is equipped with it). These entities can be assessed by the usual notion of labelling, i.e. assigning each entity a label.

Definition 3. *Given a set S, an intolerance relation on S is a binary relation $int \subseteq S \times S$, where $(s_1, s_2) \in int$ indicates that s_1 is intolerant of s_2 and will be denoted as $s_1 \odot s_2$, while $(s_1, s_2) \notin int$ will be denoted as $s_1 \ominus s_2$.*

Definition 4. *Given a sal Λ and a set S, a Λ-labelling of S is a function $L : S \to \Lambda$. Given a Λ-labelling L of S and a label $\lambda \in \Lambda$, we define $\lambda(L) = \{s \in S \mid L(s) = \lambda\}$.*

Labellings are typically required to satisfy two dual properties. On the one hand, in order to satisfy *consistency* two elements which cannot stand each other should not be assigned labels which are 'too positive' altogether. On the other hand, to avoid unjustified overly negative evaluations, labellings should satisfy *reinstatement*, i.e. a too negative label should not be assigned to an element unless another intolerant element is assigned a sufficiently positive label. The corresponding violations at the level of the labellings are modelled by distinct relations, namely an incompatibility relation and a reinstatement violation relation on assessment labels, induced by two corresponding relations on assessment classes.

Definition 5. *Given a sac C, an incompatibility relation on C is a relation $inc \subseteq C \times C$, where $(c_1, c_2) \in inc$ indicates that c_1 is incompatible with c_2 and will be denoted as $c_1 \boxdot c_2$, while $(c_1, c_2) \notin inc$ will be denoted as $c_1 \boxminus c_2$. Given a C-classified sal Λ, we define the induced incompatibility relation $inc' \subseteq \Lambda \times \Lambda$ as follows: for every $\lambda_1, \lambda_2 \in \Lambda$, $(\lambda_1, \lambda_2) \in inc'$ iff $(C_\Lambda(\lambda_1), C_\Lambda(\lambda_2)) \in inc$. With a little abuse of notation we will also denote $(\lambda_1, \lambda_2) \in inc'$ as $\lambda_1 \boxdot \lambda_2$, and analogously for $\lambda_1 \boxminus \lambda_2$.*

Definition 6. *Given a sac C, a reinstatement violation relation on C is a relation $rv \subseteq C \times C$, where $(c_1, c_2) \in rv$ indicates that c_1 is not sufficiently positive to justify c_2 and will be denoted as $c_1 \overline{\boxdot} c_2$, while $(c_1, c_2) \notin rv$ will be denoted as $c_1 \overline{\boxminus} c_2$. Given a C-classified sal Λ, we define the induced reinstatement violation relation $rv' \subseteq \Lambda \times \Lambda$ as follows: for every $\lambda_1, \lambda_2 \in \Lambda$, $(\lambda_1, \lambda_2) \in rv'$ iff $(C_\Lambda(\lambda_1), C_\Lambda(\lambda_2)) \in rv$. With a little abuse of notation we will also denote $(\lambda_1, \lambda_2) \in rv'$ as $\lambda_1 \overline{\boxdot} \lambda_2$, and analogously for $\lambda_1 \overline{\boxminus} \lambda_2$.*

Some rather natural properties can be identified for incompatibility and, in a dual manner, for reinstatement violation relations on C.

Definition 7. *Given a sac C, let inc be an incompatibility relation on C. We say that inc is well-founded if it satisfies the following properties:*

- *inc is* monotonic, *i.e. given $c_1, c_2 \in C$ such that $c_1 \boxdot c_2$, for every pair $c_1', c_2' \in C$ such that $c_1 \leq c_1'$ and $c_2 \leq c_2'$ it holds that $c_1' \boxdot c_2'$*

- *inc is non empty, i.e. inc* $\neq \emptyset$
- $\forall c_1 \in C, \exists c_2 \in C$ *such that* $c_1 \underline{\boxminus} c_2$ *and* $\exists c_3 \in C$ *such that* $c_3 \underline{\boxminus} c_1$

Definition 8. *Given a sac C, let rv be a reinstatement violation relation on C. We say that rv is well-founded iff it satisfies the following properties:*

- *rv is* dually monotonic, *i.e. given* $c_1, c_2 \in C$ *such that* $c_1 \overline{\boxdot} c_2$, *for every pair* $c_1', c_2' \in C$ *such that* $c_1' \leq c_1$ *and* $c_2' \leq c_2$ *it holds that* $c_1' \overline{\boxdot} c_2'$
- *rv is non empty, i.e. rv* $\neq \emptyset$
- $\forall c_1 \in C, \exists c_2 \in C$ *such that* $c_1 \overline{\boxminus} c_2$ *and* $\exists c_3 \in C$ *such that* $c_3 \overline{\boxminus} c_1$

The first property of each definition reflects the intuitions underlying the dual concepts of consistency and reinstatement. In particular, inconsistency arises from a sort of 'excess of simultaneous positiveness' in the assessment of some elements linked by intolerance, while reinstatement violation is due to an 'excess of cautiousness' in assigning positive labels. The second and third properties are the same in both definitions, and require that the intolerance relation between elements of S is not void of any effect and that each label is attainable, respectively.

The following definition formalizes the notions of consistent and inconsistent labelling.

Definition 9. *Given a set S equipped with an intolerance relation int, a sac C equipped with an incompatibility relation inc, and a C-classified sal* Λ, *a* Λ*-labelling L of S is int-inc-inconsistent iff*

$$\exists s_1, s_2 \in S \text{ such that } s_1 \odot s_2 \text{ and } L(s_1) \underline{\boxdot} L(s_2) \tag{1}$$

Conversely, we say that a labelling is int-inc-consistent if it is not int-inc-inconsistent, i.e.

$$\forall s_1, s_2 \in S \text{ such that } s_1 \odot s_2, \text{ it holds that } L(s_1) \underline{\boxminus} L(s_2) \tag{2}$$

The following definition dually introduces the notions of reinstatement compliant and uncompliant labelling. Note that a special condition is required for *initial* elements of S, i.e. elements s_2 of S such that there are no elements s_1 with $s_1 \odot s_2$ (the reader is referred to [2] for further details and explanations).

Definition 10. *Given a set S equipped with an intolerance relation int, a sac C equipped with a reinstatement violation relation rv, and a C-classified sal* Λ, *a* Λ*-labelling L of S is int-rv-uncompliant iff*

$$\exists s_2 \in S : \begin{cases} \min(C) \overline{\boxdot} C_\Lambda(L(s_2)) & \text{if } s_2 \text{ is initial} \\ \forall s_1 \in S \text{ such that } s_1 \odot s_2 \text{ it holds that } L(s_1) \overline{\boxdot} L(s_2) & \text{otherwise} \end{cases} \tag{3}$$

Conversely, we say that a labelling is int-rv-compliant if it is not int-rv-uncompliant, i.e.

$$\forall s_2 \in S \begin{cases} \min(C) \overline{\boxminus} C_\Lambda(L(s_2)) & \text{if } s_2 \text{ is initial} \\ \exists s_1 \in S \text{ such that } s_1 \odot s_2 \text{ and } L(s_1) \overline{\boxminus} L(s_2) & \text{otherwise} \end{cases} \tag{4}$$

The proposed model is able to capture the labelling-based version of Dung's semantics [1] as one of its instances. As well known, in Dung's theory an argumentation framework represents a set of arguments and the relevant conflicts.

Definition 11. *An argumentation framework is a pair $AF = (\mathscr{A}, \rightarrow)$ where $\mathscr{A}$ is a set of arguments and $\rightarrow \subseteq \mathscr{A} \times \mathscr{A}$ is a binary relation of attack between them. Given an argument $\alpha \in \mathscr{A}$, we denote as α^- the set $\{\beta \in \mathscr{A} \mid (\beta, \alpha) \in \rightarrow\}$.*

Given an abstract argumentation framework $AF = (\mathscr{A}, \rightarrow)$, we assume $S = \mathscr{A}$ and the intolerance relation coinciding with the attack relation, i.e. $\alpha \odot \beta$ iff $\alpha \in \beta^-$. We use the sal $\Lambda^{\mathrm{IOU}} = \{\mathtt{in}, \mathtt{out}, \mathtt{und}\}$ and the tripolar sac $C^3 = \{\mathsf{pos}, \mathsf{mid}, \mathsf{neg}\}$ with $\mathsf{neg} \leq \mathsf{mid} \leq \mathsf{pos}$. Furthermore, we adopt the classification $C^3_{\Lambda^{\mathrm{IOU}}} = \{(\mathtt{in}, \mathsf{pos}), (\mathtt{out}, \mathsf{neg}), (\mathtt{und}, \mathsf{mid})\}$, i.e. in corresponds to a definitely positive assessment, out to a definitely negative assessment, and und to an intermediate situation.

The main labelling-based semantics corresponding to the extension-based semantics introduced in [1] are defined below (the reader is referred to [4] for an extensive illustration.

Definition 12. *Let L be a labelling of an argumentation framework $AF = (\mathscr{A}, \rightarrow)$. L is* conflict-free *iff for each $\alpha \in \mathscr{A}$, if $L(\alpha) = \mathtt{in}$ then $\nexists \beta \in \alpha^- : L(\beta) = \mathtt{in}$, and if $L(\alpha) = \mathtt{out}$ then $\exists \beta \in \alpha^- : L(\beta) = \mathtt{in}$; it is* admissible *iff for each $\alpha \in \mathscr{A}$, if $L(\alpha) = \mathtt{in}$ then $\forall \beta \in \alpha^-, L(\beta) = \mathtt{out}$, and if $L(\alpha) = \mathtt{out}$ then $\exists \beta \in \alpha^- : L(\beta) = \mathtt{in}$; it is* complete *if it is admissible and for each $\alpha \in \mathscr{A}$ it holds that if $L(\alpha) = \mathtt{und}$ then $\nexists \beta \in \alpha^- : L(\beta) = \mathtt{in}$ and $\exists \beta \in \alpha^- : L(\beta) = \mathtt{und}$; it is* stable *if it is complete and $\nexists \alpha \in \mathscr{A} : L(\alpha) = \mathtt{und}$; it is* preferred *if it is complete and the set of arguments labelled* in *by L is maximal (w.r.t. $\subseteq$) among all complete labellings; it is* grounded[1] *if it is complete and the set of arguments labelled* in *by L is minimal (w.r.t. $\subseteq$) among all complete labellings; it is* semi-stable *if it is complete and the set of arguments labelled* und *by L is minimal (w.r.t. $\subseteq$) among all complete labellings.*

It is shown in [2] that conflict-free, admissible, complete and stable labellings correspond to different choices for the incompatibility and reinstatement violation relations. In particular, letting $\underline{\mathrm{inc}}_{C^3} = \{(\mathsf{pos}, \mathsf{pos})\}$, $\mathrm{inc}^a_{C^3} = \{(\mathsf{pos}, \mathsf{pos}), (\mathsf{mid}, \mathsf{pos})\}$, $\mathrm{inc}^c_{C^3} = \{(\mathsf{pos}, \mathsf{pos}), (\mathsf{pos}, \mathsf{mid}), (\mathsf{mid}, \mathsf{pos})\}$, $\overline{\mathrm{inc}}_{C^3} = \{(\mathsf{pos}, \mathsf{pos}), (\mathsf{pos}, \mathsf{mid}), (\mathsf{mid}, \mathsf{pos}), (\mathsf{mid}, \mathsf{mid})\}$, $\mathrm{rv}^{cf}_{C^3} = \{(\mathsf{neg}, \mathsf{neg}), (\mathsf{mid}, \mathsf{neg})\}$, and $\mathrm{rv}^c_{C^3} = \{(\mathsf{neg}, \mathsf{neg}), (\mathsf{neg}, \mathsf{mid}), (\mathsf{mid}, \mathsf{neg})\}$, we have the following proposition.

Proposition 1. *The set of conflict-free labellings coincides with the set of labellings which are $\rightarrow$-$\underline{inc}_{C^3}$-consistent and $\rightarrow$-$rv^{cf}_{C^3}$-compliant. The set of admissible labellings coincides with the set of labellings which are $\rightarrow$-$inc^a_{C^3}$-consistent and $\rightarrow$-$rv^{cf}_{C^3}$-compliant. The set of complete labellings coincides with the set of labellings which are $\rightarrow$-$inc^c_{C^3}$-consistent and $\rightarrow$-$rv^c_{C^3}$-compliant. The set of stable labellings coincides with the set of labellings which are $\rightarrow$-$\overline{inc}_{C^3}$-consistent and $\rightarrow$-$rv^c_{C^3}$-compliant.*

[1] It is known from [1] that the gounded labelling is unique.

3. The role of skepticism in labelling identification

As shown in the previous section, the conflict-free, admissible, complete and stable labellings can be identified as those satisfying specific compatibility and reinstatement conditions, corresponding to local constraints each involving a single argument and its attackers. However, a global criterion requiring comparisons between different labellings is needed to identify other kinds of labellings. For instance, the grounded labelling is defined as the complete labelling minimizing $\mathtt{in}$-labelled arguments, while preferred labellings maximize $\mathtt{in}$-labelled arguments. In a sense, the grounded labelling corresponds to a skeptical attitude, while preferred labellings enforce a less conservative attitude with more arguments possibly accepted.

The notion of skepticism between semantics in abstract argumentation has been introduced in [5]. Here we generalize the main concepts introduced in [5, 6] to the general model described above. First, we assume that the sac is partially preordered according to skepticism and that, similarly to the positiveness degree, this preorder is inherited by assessment labels.

Definition 13. *Given a sac C, we denote as $\leq_S$ a preorder (i.e. a reflexive and transitive relation) on C, intended to reflect the relevant skepticism degree (and thus we will sometimes refer to it as skepticism preorder). Given a C-classified set of assessment labels Λ, the skepticism preorder induced on Λ by C_Λ is denoted as $\preceq_S$ where $\lambda_1 \preceq_S \lambda_2$ iff $C_\Lambda(\lambda_1) \leq_S C_\Lambda(\lambda_2)$. As usual, $\lambda_1 \prec_S \lambda_2$ will denote $\lambda_1 \preceq_S \lambda_2$ and $\lambda_2 \not\preceq_S \lambda_1$.*

The fact that $\preceq_S$ is a preorder can be easily proved.

Intuitively, $\lambda_1 \preceq_S \lambda_2$ means that assigning the label λ_1 to an element $s \in S$ represents a "less decided" choice about its justification w.r.t. assigning λ_2 to s. It is worth pointing out that $\preceq_S$ must be clearly distinguished from $\preceq$, which instead reflects the positiveness degree of the labels. In particular, for two labels λ_1 and λ_2 corresponding to definite acceptance and definite rejection, i.e. associated to $\min(C)$ and $\max(C)$ respectively, it typically holds $\lambda_1 \prec \lambda_2$ but they reflects antithetical choices about the justification of an element and thus they can be incomparable w.r.t. skepticism.

As the tripolar sac $C^3 = \{\mathsf{pos}, \mathsf{mid}, \mathsf{neg}\}$ and the relevant sal $\Lambda^{\mathrm{IOU}} = \{\mathtt{in}, \mathtt{out}, \mathtt{und}\}$ adopted in abstract argumentation, we assume $\leq_S$ including the relations $\mathsf{mid} \leq_S \mathsf{pos}$, $\mathsf{mid} \leq_S \mathsf{neg}$, $\mathsf{mid} \leq_S \mathsf{mid}$, $\mathsf{neg} \leq_S \mathsf{neg}$, and $\mathsf{pos} \leq_S \mathsf{pos}$. Accordingly, it holds $\mathtt{und} \prec_S \mathtt{in}$ and $\mathtt{und} \prec_S \mathtt{out}$, while $\mathtt{in}$ and $\mathtt{out}$ are incomparable w.r.t. $\preceq_S$.

The skepticism preorder between labels can naturally be extended to labellings as follows.

Definition 14. *Let Λ be a sal equipped with a skepticism preorder $\preceq_S$ and let S be a set. The skepticism preorder between the Λ-labellings of S is denoted as $\preceq^L_S$, where $L_1 \preceq^L_S L_2$ iff $\forall s \in S$ $L_1(s) \preceq_S L_2(s)$.*

Intuitively, a labelling L_1 is less committed w.r.t. another labelling L_2 if L_1 makes a less committed choice w.r.t. L_2 for each of the elements of S, while an incomparable choice for even a single element makes the labellings L_1 and L_2 incomparable.

Turning to abstract argumentation, it is easy to see that $L_1 \preceq^L_S L_2$ iff $\mathtt{in}(L_1) \subseteq \mathtt{in}(L_2)$ and $\mathtt{out}(L_1) \subseteq \mathtt{out}(L_2)$. The skepticism preorder $\preceq^L_S$ allows one to identify the grounded and preferred labellings among complete labellings.

Proposition 2. *Assuming $\leq_S$ as described above, the set of preferred labellings coincides with the set of complete labellings that are maximal w.r.t. $\preceq_S^L$, and the grounded labelling coincides with the complete labelling that is minimal w.r.t. $\preceq_S^L$.*

Proof: It has been shown in [7] that, given two complete labellings L_1 and L_2, $\mathtt{in}(L_1) \subseteq \mathtt{in}(L_2)$ iff $\mathtt{out}(L_1) \subseteq \mathtt{out}(L_2)$. This entails that if L_1 and L_2 are complete labellings then $L_1 \preceq_S^L L_2$ iff $\mathtt{in}(L_1) \subseteq \mathtt{in}(L_2)$. The results then follow from the definitions of preferred and grounded labellings. □

As to semi-stable labellings, there are (at least) two ways to characterize them by maximizing labellings according to a skepticism relation.

First, grounded on the idea that the definition of semi-stable labellings requires the union of in-labelled and out-labelled arguments to be maximized (and thus there is no distinction between them) we could modify the basic skepticism relation between labels by considering in and out comparable, as having the same highest level of commitment. Thus, we can assume the preorder $\leq'_S$ instead of $\leq_S$, where $\leq'_S = \leq_S \cup \{(\mathrm{pos}, \mathrm{neg}), (\mathrm{neg}, \mathrm{pos})\}$. Accordingly, it holds $\mathtt{und} \prec'_S \mathtt{in}$, $\mathtt{und} \prec'_S \mathtt{out}$, $\mathtt{out} \preceq'_S \mathtt{in}$ and $\mathtt{in} \preceq'_S \mathtt{out}$. We can then characterize semi-stable labellings in a similar way as preferred labellings.

Proposition 3. *Assuming $\leq'_S$ as the skepticism preorder between sacs, the set of semi-stable labellings coincides with the set of complete labellings that are maximal w.r.t. $\preceq_S^L$.*

Proof: First, notice that given two labellings L_1 and L_2, $L_1 \preceq_S^L L_2$ iff $\forall s \in S : L_2(s) = \mathrm{und}$ it holds that $L_1(s) = \mathrm{und}$, i.e. iff $\mathrm{und}(L_2) \subseteq \mathrm{und}(L_1)$. Accordingly, the set of complete labellings that are maximal w.r.t. $\preceq_S^L$ coincides with the set of complete labellings that minimize the set of arguments labelled und, i.e. with the set of semi-stable labellings. □

Alternatively, it might be observed that giving up incomparability between in and out is not necessary: we can adopt a different way to extend the skepticism preorder between labels to the labellings of *S*, enforcing a weaker relation w.r.t. $\preceq_S^L$.

Definition 15. *Let Λ be a sal equipped with a skepticism preorder $\preceq_S$ and let S be a set. The relation $\preceq'^L_S$ between the Λ-labellings of S is defined as follows. Given two Λ-labellings L_1 and L_2 of S, $L_1 \preceq'^L_S L_2$ iff $\forall s \in S$, $L_2(s) \not\prec_S L_1(s)$.*

It has to be remarkes that, in general $\preceq'^L_S$ is not a preorder, since it does not satisfy transitivity. For instance, consider a sal $\Lambda = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ and an induced skepticism preorder $\preceq_S$ such as $\lambda_1 \prec_S \lambda_2 \prec_S \lambda_3$, $\lambda_1 \prec_S \lambda_4$, λ_2 and λ_4 incomparable, λ_3 and λ_4 incomparable (this can be induced for instance by a *sac* such that each label corresponds exactly to a class, with a preorder between assessment classes corresponding to $\preceq_S$). Consider a singleton $S = \{s\}$ and three labellings L_1, L_2 and L_3 such that $L_1(s) = \lambda_3$, $L_2(s) = \lambda_4$ and $L_3(s) = \lambda_2$. We have $L_1 \preceq'^L_S L_2$ since $L_2(s) \not\prec_S L_1(s)$, and $L_2 \preceq'^L_S L_3$ since $L_3(s) \not\prec_S L_2(s)$. However, it is not the case that $L_1 \preceq'^L_S L_3$, since $L_3(s) \prec_S L_1(s)$.

Nevertheless, transitivity holds in case the skepticism preorder on Λ is *quasi total*, as specified in the following definition.

Definition 16. *Let Λ be a sal equipped with a skepticism preorder $\preceq_S$. We say that $\preceq_S$ is quasi total if there is a set $M \subseteq \Lambda$ such that*

- *$\forall \lambda_1, \lambda_2 \in \Lambda$ such that $\lambda_1 \notin M$ or $\lambda_2 \notin M$, it holds $\lambda_1 \preceq_S \lambda_2$ or $\lambda_2 \preceq_S \lambda_1$;*
- *$\forall \lambda_2 \in M, \nexists \lambda_1 \in \Lambda$ such that $\lambda_2 \preceq_S \lambda_1$.*

Note that if $\lambda_1, \lambda_2 \in M$ then, by the second point of Definition 16, $\lambda_1 \not\preceq_S \lambda_2$ and $\lambda_2 \not\preceq_S \lambda_1$, i.e. λ_1 and λ_2 are incomparable. Intuitively, we assume a set M of maximally committed labels (and thus labels in M are incomparable each other) such that all the other elements are less committed than them, while non maximal elements are comparable to each other.

It is easy to see that the partial order $\leq_S$ assumed for Λ^{IOU} is quasi total, with $M = \{\mathtt{in}, \mathtt{out}\}$.

The following proposition shows that $\preceq'^L_S$ is a preorder when $\preceq_S$ is quasi total.

Proposition 4. *Let Λ be a sal equipped with a skepticism preorder $\preceq_S$ and let S be a set. If $\preceq_S$ is quasi total, then the relation $\preceq'^L_S$ between the Λ-labellings of S as defined in Definition 15 is a preorder.*

Proof: We have to show that $\preceq'^L_S$ is reflexive and transitive.

As to reflexivity, consider a Λ-labelling L of S. Since $\preceq_S$ is reflexive, $\forall s \in S$ it holds that $L(s) \preceq_S L(s)$, thus $L(s) \not\prec_S L(s)$. According to Definition 15 it then holds $L \preceq'^L_S L$.

As to transitivity, consider three Λ-labellings L_1, L_2 and L_2 of S such that $L_1 \preceq'^L_S L_2$ and $L_2 \preceq'^L_S L_3$. We have to show that $L_1 \preceq'^L_S L_3$, i.e. according to Definition 15 that $\forall s \in S$, $L_3(s) \not\prec_S L_1(s)$. For each $s \in S$ we distinguish two cases.

First, if $L_1(s) \preceq_S L_2(s)$ and $L_2(s) \preceq_S L_3(s)$, by transitivity of $\preceq_S$ it holds that $L_1(s) \preceq_S L_3(s)$, thus $L_3(s) \not\prec_S L_1(s)$.

In the other case $L_1(s) \not\preceq_S L_2(s)$ or $L_2(s) \not\preceq_S L_3(s)$, and since $L_1 \preceq'^L_S L_2$ and $L_2 \preceq'^L_S L_3$ we have that $L_1(s)$ and $L_2(s)$ are incomparable w.r.t. $\preceq_S$, or $L_2(s)$ and $L_3(s)$ are incomparable w.r.t. $\preceq_S$. By the first point of Definition 16, $\{L_1(s), L_2(s)\} \subseteq M$ or $\{L_2(s), L_3(s)\} \subseteq M$. In any case $L_2(s) \in M$, thus by the second point of Definition 16 $L_2(s) \not\preceq_S L_3(s)$, and since $L_2 \preceq'^L_S L_3$ we have that $L_2(s)$ and $L_3(s)$ are incomparable. By the first point of Definition 16 this can hold only if $L_3(s) \in M$. By the second point of Definition 16 it cannot then be the case that $L_3(s) \prec_S L_1(s)$, i.e. it holds $L_3(s) \not\prec_S L_1(s)$. $\square$

We can then show that, in the case of abstract argumentation, the relation $\preceq'^L_S$ (which is a preorder by Proposition 4) allows one to identify the semi-stable labellings among complete labellings.

Proposition 5. *The set of semi-stable labellings coincides with the set of complete labellings that are maximal w.r.t. $\preceq'^L_S$.*

Proof: Given two labellings L_1 and L_2, $L_1 \preceq'^L_S L_2$ iff $\forall s \in S$, $L_2(s) \not\prec_S L_1(s)$. If $L_2(s) \in \{\mathtt{in}, \mathtt{out}\}$ then it cannot be the case that $L_2(s) \prec_S L_1(s)$, while if $L_2(s) = \mathtt{und}$ then $L_2(s) \not\prec_S L_1(s)$ holds iff $L_1(s) = \mathtt{und}$. Summing up, $L_1 \preceq'^L_S L_2$ iff $\mathtt{und}(L_2) \subseteq \mathtt{und}(L_1)$. As a consequence, as in the proof of Proposition 3 the set of complete labellings that are maximal w.r.t. $\preceq^L_S$ coincides with the set of complete labellings that minimize the set of arguments labelled $\mathtt{und}$, i.e. with the set of semi-stable labellings. $\square$

4. Argument evaluation

As a subsequent step of argumentative reasoning (see e.g. [8]), the labellings prescribed by an argumentation semantics are typically used to evaluate argument justification and the various justification states can be represented as labels belonging to a predefined set Λ_e. Accordingly, this evaluation step can be modelled by a function which takes as input a set of Λ-labellings of a set S and returns as output a single Λ_e-labelling of S.

Definition 17. *Given two disjoint*[2] *sals* Λ *and* Λ_e, *an* evaluation function *from* Λ *to* Λ_e *is a mapping* evfun *which for every set* S *associates to each non-empty set of* Λ*-labellings of* S *a* Λ_e*-labelling of* S.

A particular case of evaluation function corresponds to deriving the evaluation label of each argument only from the labels assigned to the same argument by the Λ-labellings. The corresponding mapping is modelled by a *simple synthesis function.*

Definition 18. *Given two disjoint sals* Λ *and* Λ_e, *a simple synthesis function (ssf) from* Λ *to* Λ_e *is a mapping* $\mathsf{syn} : 2^{\Lambda} \setminus \{\emptyset\} \to \Lambda_e$. *The evaluation function derived from* syn, *denoted as* $\mathsf{evfun}^{\mathsf{syn}}$, *is defined, for every non-empty set of* Λ*-labellings* $\mathscr{L}$ *and for every* $s \in S$ *as*

$$\mathsf{evfun}^{\mathsf{syn}}(\mathscr{L})(s) = \mathsf{syn}(\mathscr{L}^{\downarrow}(s))$$

where $\mathscr{L}^{\downarrow}(s) \triangleq \{L(s) \mid L \in \mathscr{L}\}$.

In abstract argumentation, the most typical argument evaluation is based on three justification states. In particular, if the semantics prescribes a set $\mathscr{L}$ of Λ^{IOU}-labellings of a set of arguments $\mathscr{A}$, an argument $\alpha \in \mathscr{A}$ is:

- skeptically justified iff $\forall L \in \mathscr{L}\ L(\alpha) = \mathtt{in}$;
- credulously justified iff it is not skeptically justified [3] and $\exists L \in \mathscr{L} : L(\alpha) = \mathtt{in}$;
- not justified iff $\nexists L \in \mathscr{L} : L(\alpha) = \mathtt{in}$.

Accordingly, we consider a sal $\Lambda^{\text{AJ}} = \{\mathsf{SkJ}, \mathsf{CrJ}, \mathsf{NoJ}\}$ and a ssf syn_{AJ} from Λ^{IOU} to Λ^{AJ} defined, for every $\Lambda \subseteq \Lambda^{\text{IOU}}$ as follows:

- $\mathsf{syn}_{\text{AJ}}(\Lambda) = \mathsf{SkJ}$ if $\Lambda = \{\mathtt{in}\}$;
- $\mathsf{syn}_{\text{AJ}}(\Lambda) = \mathsf{CrJ}$ if $\Lambda \supsetneq \{\mathtt{in}\}$;
- $\mathsf{syn}_{\text{AJ}}(\Lambda) = \mathsf{NoJ}$ otherwise.

As to the classification of Λ^{AJ}, it is intuitive to assume $C^3_{\Lambda^{\text{AJ}}} = \{(\mathsf{SkJ}, \mathsf{pos}), (\mathsf{NoJ}, \mathsf{neg}), (\mathsf{CrJ}, \mathsf{mid})\}$. This way, it turns out that $\mathsf{NoJ} \prec \mathsf{CrJ} \prec \mathsf{SkJ}$ and $\mathsf{CrJ} \prec_S \mathsf{NoJ}$, $\mathsf{CrJ} \prec_S \mathsf{SkJ}$, NoJ and SkJ incomparable w.r.t. $\preceq_S$.

[2] We assume without loss of generality that Λ and Λ_e are disjoint. Since they are used in different stages of the reasoning process, it is always possible to adopt different 'names' for the labels in Λ and Λ_e.

[3] Traditionally credulous justification is regarded as including skeptical justification. We enforce disjoint notions so that argument justification can be properly modelled as a labelling.

It is interesting to define desiderata for evaluation functions based on the notions of positiveness and skepticism and also to consider the issue of preserving consistency and reinstatement properties across the evaluation step.

In order to express desiderata concerning positiveness, we first define a positiveness ordering between labels belonging to different *sals*, provided that they are classified with the same assessment class. In particular, we will assume that both Λ and Λ_e are C-classified, where C is a sac that provides a reference structure to compare labels in $\Lambda \cup \Lambda_e$.

Definition 19. *Let C be a sac, and let Λ and Λ_e be two disjoint C-classified sals. The total preorder induced on $\Lambda \cup \Lambda_e$ by C_Λ and C_{Λ_e}, also called positiveness preorder, is denoted as $\preceq$ where $\lambda_1 \preceq \lambda_2$ iff $C_{\Lambda_1}(\lambda_1) \leq C_{\Lambda_2}(\lambda_2)$, where each of Λ_1, Λ_2 is either Λ or Λ_e, $\lambda_1 \in \Lambda_1$ and $\lambda_2 \in \Lambda_2$.*

It should be noted that in case $\Lambda = \Lambda_e$ the definition corresponds to the relation $\preceq$ as per Definition 2. The proof that $\preceq$ is a preorder is the same.

We can now introduce a positiveness order between labellings and sets of labellings, possibly based on different sals.

Definition 20. *Let C be a sac, and let Λ and Λ_e be two disjoint C-classified sals. Let S be a set. The positiveness preorder between the Λ-labellings and Λ_e-labellings of S is denoted as $\preceq^L$, where $L_1 \preceq^L L_2$ iff $\forall s \in S$, $L_1(s) \preceq L_2(s)$. The positiveness preorder between the sets of Λ-labellings and Λ_e-labellings of S is denoted as $\preceq^{SL}$, where $\mathscr{L}_1 \preceq^{SL} \mathscr{L}_2$ iff $\forall L \in \mathscr{L}_1\ \exists L' \in \mathscr{L}_2$ such that $L \preceq^L L'$ and $\forall L' \in \mathscr{L}_2\ \exists L \in \mathscr{L}_1$ such that $L \preceq^L L'$.*

Intuitively, a labelling L_1 is no more positive w.r.t. another labelling L_2 if L_1 makes a no more positive choice w.r.t. L_2 for each of the elements of S. The idea of the $\preceq^{SL}$ relation between sets of labellings is that every labelling of $\mathscr{L}_1$ can be mapped into one at least as positive labelling of $\mathscr{L}_2$ and at the same time every labelling of $\mathscr{L}_2$ can be mapped into a no more positive labelling of $\mathscr{L}_1$. It is easy to see that both $\preceq^L$ and $\preceq^{SL}$ are preorders, taking into account that $\preceq$ is a preorder.

We can now express two possible desiderata for an evaluation function based on the positiveness preorder between sets of labellings. First, it is reasonable for an evaluation function to be monotonic with respect to this preorder.

Definition 21. *Let C be a sac, and let Λ and Λ_e be two disjoint C-classified sals. An evaluation function* evfun *from Λ to Λ_e is* well-behaved *iff for any set S and for any two non-empty sets of Λ-labellings $\mathscr{L}_1, \mathscr{L}_2$ of S such that $\mathscr{L}_1 \preceq^{SL} \mathscr{L}_2$, it holds that* $\mathsf{evfun}(\mathscr{L}_1) \preceq^L \mathsf{evfun}(\mathscr{L}_2)$.

Besides monotonicity, one might require the evaluation function to be reasonably bounded on the basis of the aggregated labellings.

Definition 22. *Let C be a sac, and let Λ and Λ_e be two disjoint C-classified sals. An evaluation function* evfun *from Λ to Λ_e is* faithful *iff for any set S and for any non-empty set $\mathscr{L}$ of Λ-labellings of S, $\exists L \in \mathscr{L}$ such that* $\mathsf{evfun}(\mathscr{L}) \preceq^L L$ *and $\exists L \in \mathscr{L}$ such that $L \preceq^L$* $\mathsf{evfun}(\mathscr{L})$.

In words, the result produced by evfun is neither strictly greater nor strictly lower than all labellings which are aggregated.

A counterpart of well-behaved and faithful evaluation functions can be introduced by considering the skepticism relation too. As in the case of positiveness, this first requires to consider labels belonging to different sals.

Definition 23. *Let C be a sac equipped with a skepticism preorder $\leq_S$, and let Λ and Λ_e be two disjoint C-classified sals. The skepticism preorder induced on $\Lambda \cup \Lambda_e$ by C_Λ and C_{Λ_e}, is denoted as $\preceq_S$ where $\lambda_1 \preceq_S \lambda_2$ iff $C_{\Lambda_1}(\lambda_1) \leq_S C_{\Lambda_2}(\lambda_2)$, where each of Λ_1, Λ_2 is either Λ or Λ_e, $\lambda_1 \in \Lambda_1$ and $\lambda_2 \in \Lambda_2$.*

We can then directly extend Definition 14 and Definition 15 to the case of different sals, introducing two skepticism relations between labellings.

Definition 24. *Let C be a sac equipped with a skepticism preorder $\leq_S$, and let Λ and Λ_e be two disjoint C-classified sals. Let S be a set. The skepticism preorder $\preceq^L_S$ and the relation $\preceq'^L_S$ between the Λ-labellings and Λ_e-labellings of S are defined as follows: $L_1 \preceq^L_S L_2$ iff $\forall s \in S\ L_1(s) \preceq_S L_2(s)$, and $L_1 \preceq'^L_S L_2$ iff $\forall s \in S$, $L_2(s) \not\prec_S L_1(s)$.*

The following definition introduces the skepticism relation between sets of labellings, assuming a skepticism preorder $\preceq^L_S$ between labellings as a parameter that can be instantiated in different ways, e.g. with the relation $\preceq^L_S$ or $\preceq'^L_S$ as in Definition 24.

Definition 25. *Let C be a sac equipped with a skepticism preorder $\leq_S$, and let Λ and Λ_e be two disjoint C-classified sals. Let S be a set and $\preceq^L_S$ a binary relation between the Λ-labellings and Λ_e-labellings of S. The skepticism preorder between the sets of Λ-labellings and Λ_e-labellings of S induced by $\preceq^L_S$ is denoted as $\preceq^{SL}_S$, where $\mathscr{L}_1 \preceq^{SL}_S \mathscr{L}_2$ iff $\forall L_2 \in \mathscr{L}_2\ \exists L_1 \in \mathscr{L}_1$ such that $L_1 \preceq^L_S L_2$.*

It should be noted that Definition 25 does not correspond to the way the positiveness preorder between labellings is extended to sets of labellings (Definition 20). In particular, for $\mathscr{L}_1 \preceq^{SL}_S \mathscr{L}_2$ to hold each labelling of $\mathscr{L}_2$ must be not less committed than at least a labelling of $\mathscr{L}_1$, while $\mathscr{L}_1$ can include labellings that are unrelated to those in $\mathscr{L}_2$. Intuitively, including additional labellings in $\mathscr{L}_1$ leaves open more possibilities as far as the argument justification is concerned, thus corresponding to a less decided assessment.

We can then introduce the notion of well-behaved evaluation function w.r.t. skepticism (which is parameterized w.r.t. the adopted skepticism relation between labellings $\preceq^L_S$ too).

Definition 26. *Let C be a sac, and let Λ and Λ_e be two disjoint C-classified sals. Let $\preceq^L_S$ be the adopted skepticism relation between labellings. An evaluation function* evfun *from Λ to Λ_e is* well-behaved *w.r.t. skepticism iff for any set S and for any two non-empty sets of Λ-labellings $\mathscr{L}_1, \mathscr{L}_2$ of S such that $\mathscr{L}_1 \preceq^{SL}_S \mathscr{L}_2$ (where $\preceq^{SL}_S$ is induced by the relation $\preceq^L_S$), it holds that* $\mathsf{evfun}(\mathscr{L}_1) \preceq^L_S \mathsf{evfun}(\mathscr{L}_2)$.

We then consider the counterpart of Definition 22 w.r.t. skepticism (again parameterized w.r.t. the adopted skepticism relation between labellings $\preceq^L_S$).

Definition 27. *Let C be a sac, and let Λ and Λ_e be two disjoint C-classified sals. Let $\preceq^L_S$ be the adopted skepticism relation between labellings. An evaluation function* evfun *from Λ to Λ_e is* faithful w.r.t. skepticism *iff for any set S and for any non-empty set $\mathscr{L}$ of Λ-labellings of S, $\exists L \in \mathscr{L}$ such that* $\text{evfun}(\mathscr{L}) \preceq^L_S L$ *and* $\exists L \in \mathscr{L}$ *such that* $L \preceq^L_S \text{evfun}(\mathscr{L})$.

Turning to consistency and reinstatement preservation, it appears desirable that the consistency and reinstatement properties of the original labellings are not lost in the derived justification labelling. Since consistency and reinstatement refers to a specific incompatibility and reinstatement violation relation, respectively, also preservation properties refer to them.

Definition 28. *Let C be a sac equipped with an incompatibility relation inc, and Λ and Λ_e be two C-classified sets of labels. An evaluation function* evfun *from Λ to Λ_e is* consistency preserving *according to*[4] *inc iff for any set S equipped with an intolerance relation int and any non-empty int-inc-consistent set $\mathscr{L}_1$ of Λ-labellings of S it holds that the labelling* $\text{evfun}(\mathscr{L}_1)$ *is int-inc-consistent.*

Definition 29. *Let C be a sac equipped with a reinstatement violation relation rv, and Λ and Λ_e be two C-classifed sets of labels. An evaluation function* evfun *from Λ to Λ_e is* reinstatement preserving *according to*[5] *rv iff for any set S equipped with an intolerance relation int and any non-empty int-rv-compliant set $\mathscr{L}_1$ of Λ-labellings of S it holds that the labelling* $\text{evfun}(\mathscr{L}_1)$ *is int-rv-compliant.*

It is then interesting to analyze the evaluation function $\text{evfun}^{\text{syn}_{AJ}}$ typically adopted in abstract argumentation against the requirements introduced so far.

As to the satisfied requirements, it has been shown in [3] that $\text{evfun}^{\text{syn}_{AJ}}$ is well-behaved and consistency preserving according to $\underline{\text{inc}}_{C^3}$, $\text{inc}^a_{C^3}$ and $\text{inc}^c_{C^3}$, while it is not according to $\overline{\text{inc}}_{C^3}$. A counterexample is provided below.

Example 1. *Consider the argumentation framework $AF = \langle\{\alpha, \beta\}, \{(\alpha, \beta), (\beta, \alpha)\}\rangle$ i.e. including a pair of mutually attacking arguments. The preferred labellings of AF are L_1 and L_2, where $L_1(\alpha) = \texttt{in}, L_1(\beta) = \texttt{out}$ and $L_2(\alpha) = \texttt{out}, L_2(\beta) = \texttt{in}$. The set $\mathscr{L} = \{L_1, L_2\}$ is $\rightarrow -\overline{inc}_{C^3} -$ consistent (this can also be inferred by Proposition 1 taking into account that L_1 and L_2 are also the stable labellings of AF). We have $\text{evfun}^{\text{syn}_{AJ}}(\{L\}) = L_e$, with $L_e(\alpha) = L_e(\beta) = \text{CrJ}$. However, L_e is not $\rightarrow -\overline{inc}_{C^3} -$ consistent, since* $(\text{mid}, \text{mid}) \in \overline{inc}_{C^3}$. *As a consequence,* $\text{evfun}^{\text{syn}_{AJ}}$ *is not consistency preserving according to $\overline{inc}_{C^3}$.*

Unfortunately, all of the other requirements are unsatisfied. The next example shows that $\text{evfun}^{\text{syn}_{AJ}}$ fails to satisfy faithfulness.

Example 2. *Consider the same argumentation framework $AF = \langle\{\alpha, \beta\}, \{(\alpha, \beta), (\beta, \alpha)\}\rangle$ of Example 1. The grounded semantics prescribes the unique labelling L such that $L(\alpha) = L(\beta) = \texttt{und}$. We*

[4]More precisely, consistency preservation should be defined w.r.t. a tuple $(C, C_\Lambda, C_{\Lambda_e}, \text{inc})$, since it also depends on how the labels of Λ and Λ_e are mapped into assessment classes. However, for ease of notation we focus on the incompatibility relation inc, since the mappings C_Λ and C_{Λ_e} are usually clear from the context.

[5]Similarly to the case of consistency preservation, reinstatement preservation should be defined w.r.t. $(C, C_\Lambda, C_{\Lambda_e}, \text{rv})$.

have then $\text{evfun}^{\text{syn}_{AJ}}(\{L\}) = L_e$*, with* $L_e(\alpha) = L_e(\beta) = \text{NoJ}$*. Now, according to* $C^3_{\Lambda IOU}$ *the label* und *corresponds to the assessment class* mid *and according to* $C^3_{\Lambda AJ}$ NoJ *corresponds to* neg*. Since it is not the case that* mid $\preceq$ neg*, it is also not the case that* $L \preceq^L L_e$*, violating the second condition of Definition 22.*

Moreover, $\text{evfun}^{\text{syn}_{AJ}}$ is neither well-behaved nor faithful w.r.t. skepticism, for both of the choices concerning the skepticism relation between labellings.

Example 3. *Consider a set S including a single element s, and two sets of labellings* $\mathscr{L}_1 = \{L'_1, L''_1\}$ *and* $\mathscr{L}_2 = \{L'_2, L''_2\}$*, with* $L'_1(s) = \texttt{out}$*,* $L''_1(s) = \texttt{und}$*,* $L'_2(s) = \texttt{in}$*,* $L''_2(s) = \texttt{und}$*. It can be checked that* $L''_1 \preceq^L_S L'_2$*,* $L''_1 \preceq^L_S L''_2$*,* $L''_1 \preceq'^L_S L'_2$*, and* $L''_1 \preceq'^L_S L''_2$*, thus* $\mathscr{L}_1 \preceq^{SL}_S \mathscr{L}_2$ *both with* $\preceq^{SL}_S$ *induced by* $\preceq^L_S$ *and with* $\preceq^{SL}_S$ *induced by* $\preceq'^L_S$*. On the other hand,* $\text{evfun}^{\text{syn}_{AJ}}(\mathscr{L}_1)(s) = \text{NoJ}$*,* $\text{evfun}^{\text{syn}_{AJ}}(\mathscr{L}_2)(s) = \text{CrJ}$*, thus it neither holds* $\text{evfun}^{\text{syn}_{AJ}}(\mathscr{L}_1) \preceq^L_S \text{evfun}^{\text{syn}_{AJ}}(\mathscr{L}_2)(s)$ *nor* $\text{evfun}^{\text{syn}_{AJ}}(\mathscr{L}_1) \preceq'^L_S \text{evfun}^{\text{syn}_{AJ}}(\mathscr{L}_2)(s)$*. This shows that* $\text{evfun}^{\text{syn}_{AJ}}$ *is not well-behaved w.r.t. skepticism.*

Example 4. *Consider again the argumentation framework and the labelling L of Example 2. It is easy to see that, letting again* L_e *be the labelling obtained by* $\text{evfun}^{\text{syn}_{AJ}}(\{L\})$*, neither* $L_e \preceq^L_S L$ *nor* $L_e \preceq'^L_S L$ *holds. Thus, the first condition of Definition 27 is violated, showing that* $\text{evfun}^{\text{syn}_{AJ}}$ *is not faithful w.r.t. skepticism.*

Finally, $\text{evfun}^{\text{syn}_{AJ}}$ is neither reinstatement preserving according to $\text{rv}^{cf}_{C^3}$ nor reinstatement preserving according to $\text{rv}^{c}_{C^3}$.

Example 5. *Consider the argumentation framework* $AF = \langle\{\alpha, \beta, \gamma\}, \{(\alpha, \beta), (\beta, \alpha), (\alpha, \gamma), (\beta, \gamma)\}\rangle$*. Here preferred and stable semantics prescribe two labellings, namely* L_1 *with* $L_1(\alpha) = \texttt{in}$*,* $L_1(\beta) = \texttt{out}$*,* $L_1(\gamma) = \texttt{out}$*, and* L_2 *with* $L_1(\alpha) = \texttt{out}$*,* $L_1(\beta) = \texttt{in}$*,* $L_1(\gamma) = \texttt{out}$*. Both of them are* $\rightarrow -rv^{cf}_{C^3} -$ *compliant and* $\rightarrow -rv^{c}_{C^3} -$ *compliant (this also derives from Proposition 1). Let* $\mathscr{L} = \{L_1, L_2\}$*. The synthesis produced by* $\text{evfun}^{\text{syn}_{AJ}}(\mathscr{L})$ *for* α *and* β *is* CrJ *and for* γ *is* NoJ *corresponding to the pair* (mid, neg) *which is forbidden by both* $rv^{cf}_{C^3}$ *and* $rv^{c}_{C^3}$*.*

5. Conclusions and perspectives

In this paper, we have introduced a first model for argument evaluation in Dung's abstract argumentation, based on the complementary notions of positiveness and skepticism characterizing the assessment labels. The model is graphically represented in Figure 5, where boxes denote reasoning steps and ovals represent (possibly singleton) sets of labellings, and with the parameters instantiating the reasoning steps reported at the top of the boxes. In particular, on the basis of the set of assessment labels Λ all of the possible labellings (i.e. the unconstrained labellings) of the input argumentation framework can be constructed. Then, unconstrained labellings are filtered by selecting those that satisfy the local constraints expressed by the incompatibility and reinstatement violation relations (see Section 2) and further possibly[6] filtered by maximizing or

[6]In the figure, 'no' indicates that no filter is applied. This can be the case e.g. of stable semantics, since stable labellings can be directly identified by local constraints filtering.

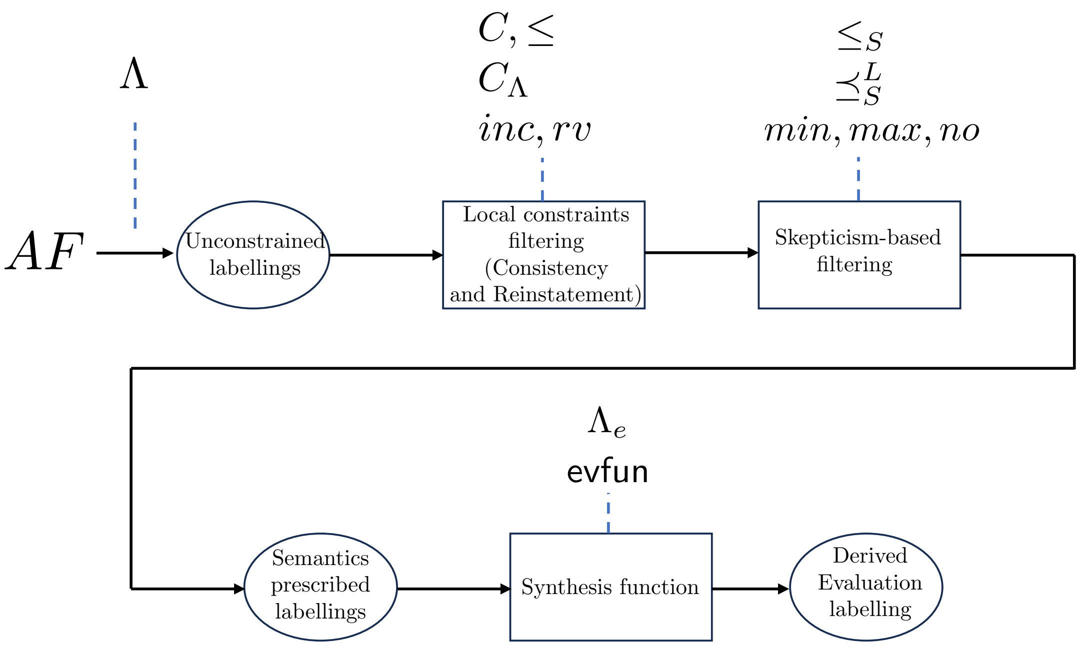

Figure 1: A graphical representation of the general model for argument evaluation.

minimizing w.r.t. a skepticism relation (see Section 3). Finally, an evaluation function is applied to determine an evaluation labelling (see Section 4). We remark that the proposed model is conceptual rather than being an implementation architecture, and it represents a first step open to further investigations, some of which are discussed below.

Starting from the final step, we have shown that $\mathsf{evfun}^{\mathsf{syn}_{\mathrm{AJ}}}$ fails to satisfy most of the desiderata introduced in Section 4. The counterexamples provided seem to suggest that a finer-grained set of evaluation labels would be needed w.r.t. to Λ^{AJ}. For instance, referring to Example 2 the reason why faithfulness is not satisfied is that $\mathsf{evfun}^{\mathsf{syn}_{\mathrm{AJ}}}$ does not distinguish between the case where an argument is out-labelled by all labellings and the case where the argument is und-labelled by at least one labelling. A distinction between these two cases would require an additional label in Λ^{AJ} to represent the latter case, that would be mapped to an assessment class characterized by an intermediate positiveness level. Similar considerations apply to Example 4 concerning skepticism. As to Example 5 concerning reinstatement preservation, it seems that $\mathsf{evfun}^{\mathsf{syn}_{\mathrm{AJ}}}$ fails to distinguish between the case where an argument is out-labelled because of an attacker which is in-labelled by all labellings, and the case of *floating defeat* where different attackers are in-labelled by different labellings. Distinguishing between these two cases might require both a finer-grained set of evaluation labels and an evaluation function which is not based on a simple synthesis function. In general, devising a set of assessment labels of Λ_e and an evaluation function evfun able to satisfy more desiderata is an interesting research issue.

Turning to skepticism-based and local constraints filtering, it would be interesting to consider argumentation semantics that are not directly captured in the current proposal. For instance, the ideal semantics [9] prescribes as the unique labelling the most committed labelling that is less committed than all preferred labellings. This can be captured by introducing a skepticism-based filter that receives in input an additional set of labellings w.r.t. the one to be filtered. Furthermore, some variations of the introduced notions can be considered, and different combinations of consistency, reinstatement and skepticism-based relations might be explored.

Another interesting issue concerns the set of assessment labels Λ adopted for the semantics prescribed labellings. While we focused on tripolar labellings, quadripolar labellings have also been introduced in the argumentation literature [10, 11] and will be considered in future work. Finally, real-valued labels and the use of infinite sets of labels could be considered, encompassing gradual forms of argumentation.

Acknowledgments

This work has been partially supported by the research project GNCS-INdAM CUP E55F22000270001, "Verifica Formale di Dibattiti nella Teoria dell'Argomentazione".

References

[1] P. M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games, Artif. Intell. 77 (1995) 321–357.

[2] P. Baroni, F. Cerutti, M. Giacomin, Generalizing consistency and reinstatement in abstract argumentation, in: Proc. of the 6th Workshop on Advances in Argumentation in Artificial Intelligence (AI^3 2022), 2022.

[3] P. Baroni, F. Cerutti, M. Giacomin, A generalized notion of consistency with applications to formal argumentation, in: Proc. of the 9th Int. Conf. on Computational Models of Argument (COMMA 2022), 2022, pp. 56–67.

[4] P. Baroni, M. Caminada, M. Giacomin, An introduction to argumentation semantics, Knowledge Engineering Review 26 (2011) 365–410.

[5] P. Baroni, M. Giacomin, G. Guida, Towards a formalization of skepticism in extension-based argumentation semantics, in: Proc. of the 4th Workshop on Computational Models of Natural Argument (CMNA 2004), 2004, pp. 47–52.

[6] P. Baroni, M. Giacomin, Skepticism relations for comparing argumentation semantics, International Journal of Approximate Reasoning 50 (2009) 854–866.

[7] M. Caminada, An algorithm for computing semi-stable semantics, in: Proc. of the 9th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU-07), 2007, pp. 222–234.

[8] P. Baroni, R. Riveret, Enhancing statement evaluation in argumentation via multi-labelling systems, J. Artif. Intell. Res. 66 (2019) 793–860.

[9] P. Dung, P. Mancarella, F. Toni, Computing ideal sceptical argumentation, Artificial Intelligence 171 (2007) 642–674.

[10] H. Jakobovits, D. Vermeir, Robust semantics for argumentation frameworks, J. of Logic and Computation 9 (1999) 215–261.

[11] O. Arieli, Conflict-free and conflict-tolerant semantics for constrained argumentation frameworks, J. Appl. Log. 13 (2015) 582–604.