Journal of Logic and Computation Advance Access published February 13, 2014

Beyond admissibility: accepting cycles in argumentation with game protocols for cogency criteria

GUSTAVO A. BODANZA and FERNANDO A. TOHMÉ, *Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina and Instituto de Investigaciones Económicas y Sociales del Sur, Bahía Blanca, CONICET. E-mail: bodanza@gmail.com; ftohme@gmail.com*

GUILLERMO R. SIMARI, *Universidad Nacional del Sur, Bahía Blanca, 8000, Argentina. E-mail: grsimari@gmail.com*

Abstract

In this article, we extend Dung's formal approach from *admissibility* to less demanding extension semantics allowing arguments in cycles of attacks. We present an acceptance criterion leading to the characterization of three semantics called *pairwise cogency*, *weak cogency* and *cyclic cogency*. Particular game-theoretic protocols allow us to identify winning strategies with extensions in different semantics. Furthermore, an algorithmic characterization of those games exhibits clearly how self-attacking or in odd-length cycles of attack can be rationally managed beyond the limits of admissibility.

Keywords: argumentation frameworks, argumentation games, extension semantics, defence criteria.

1 Introduction

Argumentation is a process in which arguments for and against some claim are put forward by contending parties. While it might adopt different patterns, a very general characterization of argumentation starts from a general set of arguments and a given 'attack' relation among them. In [10], Dung formalizes an *argumentation framework* as a tuple $AF = \langle \mathcal{A}, \gg \rangle$, which consists of a set of arguments $\mathcal{A}$ and the *attack relation* $\gg \subseteq \mathcal{A} \times \mathcal{A}$ such that, given $A, B \in \mathcal{A}$, $A \gg B$ means that A attacks B. For simplicity, we will only consider *finite* argumentation frameworks, i.e. frameworks based in a finite set $\mathcal{A}$.[1]

The main problem for argumentation frameworks is the determination of which arguments should be accepted in the framework. Acceptance, in turn, depends on the way in which arguments can be defended from attacks. The different ways in which this can be achieved lead to corresponding notions of acceptance. An acceptance criterion can be captured as an *extension semantics* $\mathcal{S}$, yielding, for every argumentation framework AF, a family $\mathcal{E}_{\mathcal{S}}(AF) \subseteq 2^{\mathcal{A}}$ of 'extensions' of AF. An argument is said *credulously* accepted under a semantics $\mathcal{S}$ if and only if it belongs to *some* extension $E \in \mathcal{E}_{\mathcal{S}}(AF)$, and is said to be *sceptically* accepted if and only if it belongs to *every* extension $E \in \mathcal{E}_{\mathcal{S}}(AF)$.

Argumentation semantics have been evaluated in different ways, usually in terms on how well they behave on canonical examples but also through their confrontation with human behaviour [20]. Baroni and Giacomin [3] presented a more systematic approach to the evaluation of extension semantics,

[1]From now on 'AF' will denote an arbitrary argumentation framework, unless the contrary is stated. Moreover, we will eventually refer to attacks among arguments and/or sets of arguments, overloading the symbol '$\gg$'.

doi:10.1093/logcom/exu004

based on how they satisfy certain principles. One of them is *admissibility*, a criterion introduced in Dung's seminal work and that most extension semantics in the literature satisfy. In [3] two versions of admissibility are distinguished, a weak one (every argument that belongs to an extension is defended by any argument within the extension) and a strong one (every argument that belongs to an extension is defended by *another* argument within the extension).

The study of extension semantics under other principles gained attention in recent years. Most of them are concerned with the treatment of loops in the attack relation and defining weaker forms of defence. One of the reasons for the interest on this line of research is the failure of admissibility in addressing loops in the attack relation. In fact, admissibility discards arguments attacked by self-attacking arguments and rejects arguments involved in odd-length cycles of length ≥ 3 (*ceteris paribus*). Arguably, $CF2$ [2] and *stage* [23] semantics are the best known amendments for these shortcomings, but other alternatives exist. In [4] and [5] some extension semantics can be found, based on a general criteria of defence called 'cogency'. Unlike admissibility, which is a *property* of sets of arguments, cogency is predicated as a relationship *between sets* of arguments. So, an admissible set of arguments is one that defends itself from *any* external attack, while a set of arguments S is at least as cogent as a set of arguments S' if S is admissible in the argumentation framework restricted to $S \cup S'$.

The notion of cogency, albeit based on the notion of admissibility, yields a class of acceptability criteria with different degrees of tolerance to the presence of loops in the attack relations. Furthermore, most usual proof-procedures for argumentation based on *labellings* [7, 15] and *dialogue games* [1, 8, 11, 12, 16, 21, 22, 24, 26] can be retooled for cogency.

Precisely, the goal of this article is to provide a dialogue game counterpart to cogency-based semantics (to our knowledge, this is the first approach of its kind for non-admissible extension semantics). We present a precise game-theoretical formulation of two-party dialogue games in the line of others in the literature, mainly [22]. Basic rules of the game allow to capture admissibility in terms of the class of arguments for which their proponent has winning strategies. We will show that winning strategies under specific protocols constraining the moves of the players allow to capture three different cogency-based semantics. In [6] we made a first approach towards the capture of pairwise and weak cogency; the present article is an extension of that work in which, in addition, cyclic cogency is addressed (in line with the concepts introduced in [4] and [5]) and sceptical decision games are defined for all the three criteria under specific conditions. Each protocol either adds new constraints on the legal moves of the players (pairwise cogency) or specifies the ways other protocols are invoked at different game rounds (weak and cyclic cogency). More precisely, the protocols for weak and cyclic cogency define a *meta-argumentation game* in which the players (proponent and opponent) play a series of (object level) argumentation games. Weak cogency is captured by the winning strategies of the proponent that end up advancing, if it exists, a better theory (in terms of cogency) than anyone the opponent can use against her. Cyclic cogency, in turn, is obtained by means of the winning strategies of the proponent that prescribe advancing, if they exist, a finite sequence of theories, each one better (in terms of cogency) than the previous one, starting at any theory the opponent can use against her and ending back at the proponent's initial theory. Finally, the sceptical decision problem is approached through the winning strategies of the proponent under a protocol that captures the arguments that belong to every set that maximally satisfies a given cogency criterion.

The article is organized as follows. In Section 2, we present several principles on which defence criteria can be based. On one hand, the admissibility-related criteria introduced in [3] are revisited; on the other hand, our cogency-related criteria are defined. In Section 3, the basics of a dialogue game are defined together with protocols for admissibility. This game is in essence the same which, under different protocols, captures in its winning strategies the different cogency criteria. This is shown in Section 4. Section 5 evaluates some known non-admissible semantics according to our cogency related principles. Final considerations close the paper in Section 6.

2 Defence criteria

In [3], Baroni and Giacomin have proposed several criteria for the evaluation of extension semantics. We will first analyse those related to the widely accepted notion of defence called *admissibility*. Additionally, we will introduce the *conflict-freeness* criterion, which is satisfied by all the extension semantics in the literature. After that, we will present the concept of cogency and its ensuing criteria: pairwise, weak and cyclic cogency.

2.1 Admissibility and conflict-freeness

Admissibility is considered in [3] as a general criterion of defence that can appear with two different degrees of strength: a weak version that we call 'defensibility' and a strong one called 'strong admissibility'.

Definition 2.1
For any argumentation framework $AF = \langle \mathcal{A}, \gg \rangle$ and $S \subseteq \mathcal{A}$, let $F(S) = \{A : \forall B(B \gg A \Rightarrow \exists C \in S \;\; C \gg B)\}$ ([10]'s characteristic function). A subset $T \subseteq \mathcal{A}$ is *defensible* iff T is such that:

$$T \subseteq F(T) \tag{1}$$

A semantics $\mathcal{S}$ satisfies the *defensibility criterion* iff for any argumentation framework AF and for every $E \in \mathcal{E}_{\mathcal{S}}(AF)$, E is defensible.

Another important criterion considered in [3] demands that all the extensions sanctioned by a semantics must be conflict-free. Although not strictly related with defence, all the extension semantics in the literature satisfy the following principle:

Definition 2.2
A set of arguments T is *conflict-free* iff $\forall A, B \in T \; (A \not\gg B)$. A semantics $\mathcal{S}$ satisfies the conflict-free criterion iff for any argumentation framework $AF = \langle \mathcal{A}, \gg \rangle$ and for every $E \in \mathcal{E}_{\mathcal{S}}(AF)$, E is conflict-free.

Dung [10] uses the term 'admissible' for sets of arguments which are both defensible and conflict-free. In what follows, we use the term *admissible* to denote sets that satisfy Dung's criteria. On the other hand, *defensibility* is applied to denote a property of extension semantics. While related, these uses are not identical since, while any semantics for which all its extensions are admissible satisfies the defensibility criterion, its converse is not necessarily true.

A stronger criterion of defence is obtained also in terms of classes of arguments.

Definition 2.3
Given a class of arguments $S \subseteq \mathcal{A}$, the set of arguments *strongly defended* by S is the set $sd(S) \subseteq \mathcal{A}$ such that $A \in sd(S)$ iff for all $B \in \mathcal{A}$, if $B \gg A$ then there exists some $C \in S - \{A\}$ such that $C \gg B$ and $C \in sd(S - \{A\})$.

Definition 2.4
For any argumentation framework $AF = \langle \mathcal{A}, \gg \rangle$, a subset $T \subseteq \mathcal{A}$ is *strongly admissible* iff T satisfies the following condition:

$$T \subseteq sd(T) \tag{2}$$

A semantics $\mathcal{S}$ satisfies the *strong admissibility* criterion iff for any argumentation framework AF and for every $E \in \mathcal{E}_{\mathcal{S}}(AF)$, E is strongly admissible.

All the semantics presented by Dung satisfy the defensibility criterion but only grounded semantics satisfies the strong admissibility criterion.

2.2 Cogency

There also exist extension semantics that go beyond admissibility. Examples are *stage* [23], *CF*2 [2], *sustainable*, *tolerant* [5] and *lax* [4] semantics. In the same way as [3] extract the admissibility-based principles from Dung's admissibility defence criterion, we will extract three defence criteria, weaker than admissibility, for sustainable, lax and tolerant semantics, respectively: *pairwise cogency*, *weak cogency* and *cyclic cogency*.

Informally, we call 'cogent' an argument that is accepted unless some of its attackers have better defences than itself. For example, an argument supported by a well-established scientific theory, e.g., the argument from special relativity theory that no particle can exceed the speed of light is not defeated by an argument drawn from experiments showing an 'anomaly', say neutrinos traveling faster than light, unless this evidence is in turn supported by a rival, stronger scientific theory. Cogency can be seen as a notion according to which admissibility exerts only a relative, contextual, authority.

Definition 2.5
Given an argumentation framework $AF = \langle \mathcal{A}, \gg \rangle$, and two subsets $S, S' \in 2^{\mathcal{A}}$, we say that S is *at least as cogent as* S', in symbols, $S \geq_{cog} S'$, iff S is admissible in the restricted argumentation framework $\langle \mathcal{A}, \gg_{|S \cup S'} \rangle$. We say that S is *strictly more cogent than* S', in symbols, $S >_{cog} S'$, iff $S \geq_{cog} S'$ and not $S' \geq_{cog} S$.

Up from $\geq_{cog}$ we obtain a new criterion of defence, weaker than admissibility.

Definition 2.6
For any argumentation framework $AF = \langle \mathcal{A}, \gg \rangle$, a subset $T \subseteq \mathcal{A}$ is *pairwise cogent* iff T is maximal w.r.t. $\geq_{cog}$, i.e.:

$$\forall S \subseteq \mathcal{A}\; S \not>_{cog} T \tag{3}$$

A semantics $\mathcal{S}$ satisfies the *pairwise cogency* criterion iff for any argumentation framework AF and for every $E \in \mathcal{E}_{\mathcal{S}}(AF)$, E is pairwise cogent.

Let $Cog(AF) = \{E \subseteq \mathcal{A} : E$ is pairwise cogent$\}$. It is easy to see that if E is admissible then $E \in Cog(AF)$ (moreover, $E \geq_{cog} S$ for every $S \subseteq \mathcal{A}$). *Sustainable* semantics, in particular, takes as extensions all the maximal (w.r.t. $\subseteq$) elements of $Cog(AF)$.

The salient behaviour of sustainable semantics is the avoidance of undesired interferences of self-attacking arguments.

Example 2.7 (Pollock's 'lottery paradox' paradox example [19])
Imagine a fair lottery with 1,000,000 tickets, so that each ticket has one in a million chances of winning. Given a particular ticket, we can tentatively conclude that it will not win, since its probability is very low. But the same can be concluded about every ticket, hence we can conclude that none of them will win. But given that the lottery is fair, one ticket must be drawn. This is called 'the lottery paradox'. Upon this contradiction a further contradiction can be obtained, namely that since none of the tickets will win, the lottery cannot be fair. That is, the argument of the lottery paradox, based upon the premise that the lottery is fair, attacks itself since one of its consequences contradicts its premise. This is what Pollock called the ' 'lottery paradox' paradox'. It can be modelled with the

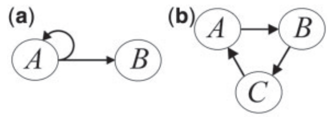

FIGURE 1. Two argumentation frameworks with odd-length cycles of attack.

argumentation framework $AF = \langle\{A,B\},\{(A,A),(A,B)\}\rangle$ (Figure 1a), where:

$A =$ 'there are good prima facie reasons to conclude that the lottery is fair, hence presumably the lottery is fair', and
$B =$ 'accepting argument A, for each ticket t_i, $1 \leq i \leq 1{,}000{,}000$, we can tentatively infer that t_i will not win, but this implies that no ticket will be drawn, hence the lottery is not fair'.

This framework has only one sustainable extension, $\{B\}$. Note that the only subsets satisfying (3) are $\emptyset$ and $\{B\}$, but the latter is the only maximal one. On the other hand, note that we have both $\{B\} \not>_{cog} \{A\}$ and $\{A\} \not>_{cog} \{B\}$. This explains why $\{B\} \in Cog(AF)$, even if it is not admissible.

Even weaker (but sensible) defence criteria can be conceived. Let us see the following example:

EXAMPLE 2.8
Assume you have to choose a school for your children and your candidates are s_1, s_2 and s_3. You evaluate them according to three criteria: proximity, tuition fee and social environment. A candidate is preferred to another if it is better with respect to most of the criteria. Assume now that after a one-to-one comparison you build these three arguments:

A: 's_2 is better than s_1 with respect to proximity, but s_1 is better than s_2 with respect to tuition fee and environment; then choose s_1';
B: 's_3 is better than s_2 with respect to environment, but s_2 is better than s_3 with respect to proximity and tuition fee; then choose s_2'
C: 's_1 is better than s_3 with respect to tuition fee, but s_3 is better than s_1 with respect to environment and proximity; then choose s_3'.

This situation is modeled as the argumentation framework $\langle\{A,B,C\}, \{(A,B), (B,C), (C,A)\}\rangle$ which is depicted in Figure 1b.

The above example shows a decision which is blocked up as the three arguments are involved in a odd-length cycle of attacks. Pairwise cogency (as well as admissibility) will deem all the three arguments out of every extension. So, according to our intuition, pairwise cogency is still a too stringent criterion in contexts of practical decisions as this one (readers familiarized with Social Choice Theory will see the resemblance with Condorcet's paradox). In our opinion, each argument here should belong to some extension as each one is arguably not worse defended than its attacker. The following criterion achieves the goal just requiring for a candidate that if there exists some strictly more cogent set then this should not be pairwise cogent (notice that in the example we have $\{A\} >_{cog} \{B\} >_{cog} \{C\} >_{cog} \{A\}$ and moreover none of this sets is pairwise cogent).

DEFINITION 2.9
For any argumentation framework $AF = \langle\mathcal{A},\gg\rangle$, a subset $T \subseteq \mathcal{A}$ is *weakly cogent* iff T satisfies the following condition:

$$\forall S \subseteq \mathcal{A}\ (S >_{cog} T \Rightarrow S \notin Cog(AF)) \tag{4}$$

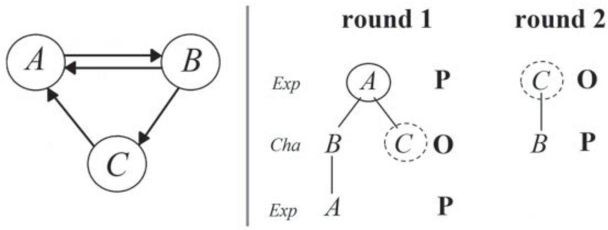

FIGURE 2. {C} is not a weakly cogent set. The game shows that C cannot be defended by the opponent.

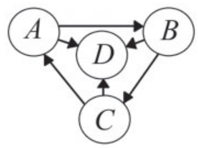

FIGURE 3. {D} is weakly cogent but not cyclically cogent.

A semantics $\mathcal{S}$ satisfies the *weak cogency* criterion iff for any argumentation framework AF and for every $E \in \mathcal{E}_{\mathcal{S}}(AF)$, E is weakly cogent.

We call *Lax* a semantics characterized as the class of maximal (w.r.t. $\subseteq$) subsets of arguments satisfying (4). Note that every weakly cogent set T is such that it is either pairwise cogent or for every $S >_{cog} T$ there exists T' satisfying $T' >_{cog} S$. It follows that every semantics that satisfies pairwise cogency also satisfies weak cogency, and if it satisfies weak cogency then it is conflict-free (note that if T is not conflict-free, then $\emptyset >_{cog} T$ and $\emptyset \in Cog(AF)$).

Going back to the last example, the lax extensions $\{A\}$, $\{B\}$ and $\{C\}$ coincide with those sanctioned by the $CF2$ semantics. Nevertheless, $CF2$ does not satisfy pairwise nor weak cogency. For instance, in the argumentation framework $\langle\{A,B,C\},\{(A,B),(B,A),(B,C),(C,A)\}\rangle$ (Figure 2, left), the $CF2$ semantics yields again the three extensions $\{A\}$, $\{B\}$ and $\{C\}$, but $\{C\}$ is not weakly cogent (since $\{B\} >_{cog} \{C\}$ and $\{B\} \in Cog(AF)$).

On the other hand, weak cogency may seem *too weak* with respect to other situations. Consider the following interpretation of the argumentation framework depicted in Figure 3.

EXAMPLE 2.10
A: 's_2 is better than s_1 with respect to proximity, but s_1 is better than s_2 with respect to tuition fee and environment; moreover, s_2 is better than s_4 in every respect. Then choose s_1';
B: 's_3 is better than s_2 with respect to environment, but s_2 is better than s_3 with respect to proximity and tuition fee; moreover, s_3 is better than s_4 in every respect. Then choose s_2';
C: 's_1 is better than s_3 with respect to tuition fee, but s_3 is better than s_1 with respect to environment and proximity; moreover, s_1 is better than s_4 in every respect. Then choose s_3';
D: 's_4 is acceptable by default, then choose s_4'.

Since neither A, B nor C belongs to a pairwise cogent set, $\{D\}$ is weakly cogent. But accepting D seems excessively naive, since $\{D\}$ is defenceless against the attack of all the remaining extensions. So, what we want here is a criterion so lax as to accept arguments that belong to odd-length cycles of attack but stringent enough as to avoid the acceptance of arguments like D. The *cyclic cogency* criterion captures the idea that arguments that are considered once and again in any sequence of defeats deserve acceptance.

DEFINITION 2.11
Let $S^{>} =_{df} \{S' \colon \text{if } S' \neq S \text{ then there exists a sequence } S >_{cog} \ldots >_{cog} S'\}$. For any argumentation framework $AF = \langle \mathcal{A}, \gg \rangle$, a subset $T \subseteq \mathcal{A}$ is a *cyclically cogent* set iff T satisfies the following condition:

$$\forall S \subseteq \mathcal{A}(S >_{cog} T \rightarrow S \in T^{>}) \tag{5}$$

A semantics $\mathcal{S}$ satisfies the *cyclical cogency* criterion iff for any argumentation framework AF and for every $E \in \mathcal{E}_{\mathcal{S}}(AF)$, E is a cyclically cogent set.

EXAMPLE 2.12
Consider again the argumentation framework depicted in Figure 3. $\{A\}$, $\{B\}$ and $\{C\}$ are cyclically cogent sets, but $\{D\}$ is not.

In the above example, the $CF2$ semantics also yields $\{A\}$, $\{B\}$ and $\{C\}$ as extensions and rejects $\{D\}$. Nevertheless, it does not satisfy the cyclic cogency criterion. In the argumentation framework shown in Figure 2 (left), note that $\{A\}$ and $\{C\}$ are $CF2$ extensions which are not cyclically cogent sets.

The following result connects conflict-free and cogency-related criteria (it will be useful later in section 4.4).

LEMMA 2.13
If S and S' are both $\begin{Bmatrix} \textit{pairwise} \\ \textit{weakly} \\ \textit{cyclic} \end{Bmatrix}$ cogent sets, then

$S \cup S'$ is a $\begin{Bmatrix} \textit{pairwise} \\ \textit{weakly} \\ \textit{cyclic} \end{Bmatrix}$ cogent set iff $S \cup S'$ is conflict-free.

PROOF. ($\rightarrow$) By contraposition. Note that if $S \cup S'$ is not conflict-free then $\emptyset >_{cog} S \cup S'$, which implies that $S \cup S'$ is neither pairwise, weakly nor cyclic cogent.
($\leftarrow$) The cases related to pairwise and weakly cogency are both proved by way of contradiction in a similar way, while that related to cyclic cogency follows immediately from the case of weakly cogency since every cyclic cogent set is also weakly cogent. Let $S \cup S'$ be conflict-free:

- *Pairwise cogency*. Assume that $T >_{cog} S \cup S'$. Then there exists some $A \in T$ such that for some $B \in S \cup S'$, $A \gg B$ and $S \cup S' \not\gg A$. But then either $A \gg S$ and $S \not\gg A$ or $A \gg S'$ and $S' \not\gg A$. Hence either $S \notin Cog(AF)$ or $S' \notin Cog(AF)$. Either case contradicts the main hypothesis.
- *Weak cogency*. Assume that $T >_{cog} S \cup S'$ and there exists no T' such that $T' >_{cog} T$. Then there exists some $A \in T$ such that for some $B \in S \cup S'$, $A \gg B$ and $S \cup S' \not\gg A$. But then either $A \gg S$ and $S \not\gg A$ or $A \gg S'$ and $S' \not\gg A$. Hence there exists no T' such that $T' >_{cog} T$ and either S is not weakly cogent or S' is not weakly cogent. Either case contradicts the main hypothesis.

■

3 Argumentation games over argumentation frameworks

We will introduce now the topic of how the defence criteria presented in the previous section can be captured by means of dialogue games. More precisely, we will define a general dialogue model, in line with others in the literature, and specific protocols for the admissibility criteria. In the next section, we will deal with our main goal which is to find specific protocol rules and their relationships with the cogency-based criteria. A salient contribution will be to show that all the protocols, both for admissibility and cogency criteria, can be seen in terms of variants of the same basic game.

The abstract character of argumentation frameworks calls for very general notions of acceptance. Here we enumerate the most common features of dialogical proof procedures:

– *Two party*: one of the players is the 'proponent' (**P**), who defends an argument, and the the other is the 'opponent' (**O**), who aims to defeat it.

– *Zero-sum game*: only one of the players wins.

– *Finiteness*: the number of arguments put forward by both **P** and **O** is finite.[2]

These features provide a very basic underlying framework of dialogue games ([16], etc.). Other, more comprehensive notions of acceptance can be defined on the basis of this model by adding further *protocol rules*, without changing the features of the original game. Next we redefine argumentation games in precise game-theoretical terms as two-players zero-sum games [18], following in general the formalism in [22].

Definition 3.1
An *argumentation game* on an argumentation framework $AF = \langle \mathcal{A}, \gg \rangle$ is a zero-sum extensive game in which:

(1) There are two players, i and $-i$, who play the roles of **P** and **O**, respectively.
(2) A *history* in the game is any sequence $A_0, A_1, A_2, \ldots, A_{2k}, A_{2k+1}, \ldots$ of choices of arguments in $\mathcal{A}$ made by the players in the game. A_{2k} corresponds to **P** and A_{2k+1} to **O**, for $k = 0, 1, \ldots$.
(3) At any history, A_0 is the argument that player **P** intends to defend.
(4) In a history the choices by a player i at a level $k > 0$ are $C_i(k) = \{A \in \mathcal{A} : \exists B \in C_{-i}(k-1), A \gg B\}$.
(5) A history of finite length K, $A_0, \ldots, A_K$, is *terminal* if A_K corresponds to player j ($j = i$ or $j = -i$) and $C_{-j}(K+1) = \emptyset$.
(6) Payoffs are determined at terminal histories: at $A_0, \ldots, A_K$, **P**'s payoff is 1 if K is even (i.e. **O** cannot reply to **P**'s last argument), and -1 otherwise. In turn, **O**'s payoff at $A_0, \ldots, A_K$ is 1 if K is odd and -1 otherwise.

A game in which **P** intends to defend an argument A can be represented by a *rooted tree*, in which A is the root. Each non-terminal node at level l consists of a history $A_0, \ldots, A_l$ and its children are all the histories $A_0, \ldots, A_l, A_{l+1}$. The terminal nodes are, of course, the terminal histories.[3]

Definition 3.2
A *strategy* for a player i is a function that assigns an element $A_{l+1} \in C_i(l)$ at each non-terminal history $A_0, \ldots, A_l$ where A_l corresponds to player $-i$. A strategy W of the player i in game in which i intends to defend argument A will be denoted by '$W(A)_i$'.

Notice that any pair of strategies, one for each player, in a game starting at A, $W(A)_i$ and $W(A)_{-i}$ determine a unique terminal history.

[2]All the ensuing results can be extended to *determined* infinite games [13], but we will keep our analysis in the finite realm since this extension involves topological questions far removed from the mathematical approach followed in the literature on argumentation [17].

[3]The usual custom is to identify each node with the *last* argument in the corresponding history. So for instance the node for $A_0, \ldots, A_l$ is denoted just A_l.

Definition 3.3

A strategy $W^*(A)_i$ of player i in a game in which $i = \mathbf{P}$ is said a *winning strategy* for **P** if for every strategy chosen by $-i$, $W(A)_{-i}$, the ensuing terminal history yields a payoff 1 for player i.[4]

If **P** has a winning strategy, it means that her initial argument can be defended against any possible attack. On the contrary, if **O** has a winning strategy, it means that **P** is defeated.

Notice that winning strategies for either **P** or **O** cannot be ensured to exist if the game tree is infinite. Even being finite, an argumentation framework which is not free of cycles in the attack relation can yield an infinite tree.

Let $\mathbf{W}(A)_\mathbf{P}$ the set of arguments played by **P** in a winning strategy $W(A)_\mathbf{P}$ in a game for A. It is easy to see that **P** has a winning strategy for every argument belonging to $\mathbf{W}(A)_\mathbf{P}$.

Lemma 3.4

Let A be an argument **P** has a winning strategy $W(A)_\mathbf{P}$ for. Then **P** has a winning strategy for every argument $B \in \mathbf{W}(A)_\mathbf{P}$.

Proof. Trivial. Suppose that **P** does not have a winning strategy for $B \in \mathbf{W}(A)_\mathbf{P}$. It means that once **P** plays B, she can no longer force the game towards a terminal history in which she gets 1. But then, when the game for argument A leads **P** to declare B, the game can follow a history in which **P** does not win, contradicting the assumption that $W(A)_\mathbf{P}$ is a winning strategy. ■

Lemma 3.5

The set $\mathbf{W}(A)_\mathbf{P}$ of all the arguments played by **P** in her winning strategy $W(A)_\mathbf{P}$ constitutes an admissible and strongly admissible set of arguments.

Proof. First we prove that $\mathbf{W}(A)_\mathbf{P}$ satisfies condition (1) and is conflict-free. Assume **P** has a winning strategy $W(A)_\mathbf{P}$ for A and assume by contradiction that $\mathbf{W}(A)_\mathbf{P}$ is not admissible. This implies that one of the following must be the case: (a) $\mathbf{W}(A)_\mathbf{P}$ is not conflict-free; (b) there exists some argument $B \in \mathbf{W}(A)_\mathbf{P}$ which is not acceptable w.r.t. $\mathbf{W}(A)_\mathbf{P}$. But then:

(a) The hypothesis that $B, C \in \mathbf{W}(A)_\mathbf{P}$ are such that $B \gg C$ leads to a contradiction, since it implies that if **P** plays C then **O** can win the game by playing B and following the same sequence of moves as **P** would play in $W(A)_\mathbf{P}$ if she were to defend B.

(b) It follows that there exists some $C \in \mathcal{A}$ such that $C \gg B$ and for every $D \in \mathbf{W}(A)_\mathbf{P}$ it is not the case that $D \gg C$. Then **O** can win the game by playing C when **P** plays B. This, in turn, implies that B does not belong to a winning strategy of **P**, contradicting the hypothesis.

To prove that $\mathbf{W}(A)_\mathbf{P}$ satisfies condition (2) just notice that no history built on this strategy can have cycles. Otherwise it would be infinite and the strategy would not be winning. ■

Proposition 3.6

Let $S \subseteq \mathcal{A}$ be the set of all the arguments that can be defended by **P** with a winning strategy. The set $\bigcup_{A \in S} \mathbf{W}(A)_\mathbf{P}$ is an admissible and strongly admissible set of arguments.

Proof. Given lemma 3.5 we need only prove that $\bigcup_{A \in S} \mathbf{W}(A)_\mathbf{P}$ is conflict-free. Let $A, B \in S$, and suppose by contradiction that $A' \in \mathbf{W}(A)_\mathbf{P}$ and $B' \in \mathbf{W}(B)_\mathbf{P}$ are such that $A' \gg B'$. By lemma 3.4 we have that **P** has a winning strategy $W(A')_\mathbf{P}$ for A'. But then $W(A')_\mathbf{P}$ is also a winning strategy for **O** in the game against B'. Hence, $W(B)_\mathbf{P}$ is not a winning strategy for **P**. Contradiction. ■

[4]Similarly, if $i = \mathbf{O}$, $W^*(A)_i$ is winning for **O**.

However, an argument that belongs to an admissible and strongly admissible but not conflict-free set is not ensured to be defended by **P** with a winning strategy. To see this, consider the argumentation framework $\langle\{A,B,C\},\ \{(A,B),\ (B,C),\ (C,A)\}\rangle$ in which some (strange) semantics could sanction only the single extension $\{A,B,C\}$, satisfying both admissibility and strong admissibility. But no argument in this extension can be defended by **P** (the cycle of attacks makes the game infinite). Of course, this semantics does not satisfy the conflict-free criterion, unlike all the known semantics in the literature.

3.1 Protocols capturing defensibility for finite argumentation games

Argumentation games, according to Definition 3.1, are not necessarily finite. Given the rules of the game, it is clear that a possible source of non-finiteness is the existence of cycles in the attack relation. Thus, there are basically two ways to avoid infiniteness in a framework with a finite number of arguments: (i) restricting argumentation games to argumentation frameworks in which the attack relation is acyclic; (ii) adding rules forbidding either one or both of the players to repeat arguments in some specific way. The first alternative leaves out many interesting cases. So, we (as usual in the literature) will follow the latter one.

Dung's grounded semantics is the best known strongly admissible semantics. It sanctions only one extension, the least fixed point of the characteristic function $F(S)$ (def. 2.1). No protocol needs to be added to the rules 1–6 to capture grounded semantics. This is shown, *mutatis mutandis*, in [16] and [22]. It is also known that every semantics satisfying strong admissibility is 'covered' by grounded semantics (cf. [3]). So, the game protocol needs only be extended to ensure the finiteness of the game.

PROTOCOL 1. Rules 1–6 plus

7. **P** is not allowed to play an argument that was previously played by either player.

The following two results are immediate after the considerations on grounded semantics discussed above (we make explicit the proofs anyway).

PROPOSITION 3.7
Let $A \in E$ for any extension $E \in \mathcal{E}_{\mathcal{S}}(AF)$ of a semantics $\mathcal{S}$ satisfying the conflict-free, admissibility and strong admissibility criteria. Then **P** has a winning strategy for A under protocol 1.

PROOF. If **P** could play an argument played before under her winning strategy, her strategy would lead to a non-terminating game (by engendering a cycle of attacks), contradicting the very concept of winning strategy. Hence, any winning strategy of **P** is obtained when **P** plays in accordance with Protocol 1. ■

PROPOSITION 3.8
If **P** has a winning strategy for A under protocol 1 then $A \in E$ for some set $E \subseteq \mathcal{A}$ satisfying strong admissibility (i.e. condition (2)).

PROOF. Given Lemma 3.5, we only have to prove that **P** does not need to repeat any argument played before in the same history. This is trivially true since, otherwise, there would exist some loop in her play, enabled by her strategy, contradicting the fact that it is a winning one. ■

The next protocol captures defensibility and its properties were, in essence, already found by Vreeswijk and Prakken [26].

PROTOCOL 2. Rules 1–6 plus

7. Neither player is allowed to advance an argument that was previously played by **O** (if **P** repeats **O** then she is introducing a conflict within her own strategy; if **O** repeats **O** then she is repeating an unsuccessful counter-argument).
8. Neither player is allowed to move if the last argument in the sequence was previously played by **P** (i.e. if the next move corresponds to **O** and **P** has repeated herself then **O** looses –her strategy failed; if the next move corresponds to **P** and **O** has repeated an argument that was previously played by **P** then **P** looses —**O** has made an *eo ipso* move).
9. If **O** is out of moves according to the preceding rules, she can backtrack to the last node where an argument played by **P** can lead to a different, non previously played, history.

The corresponding results arise from what is known about games for preferred semantics [26]:

PROPOSITION 3.9
Let $A \in E$ for any extension $E \in \mathcal{E}_{\mathcal{S}}(AF)$ of a semantics $\mathcal{S}$ satisfying the defensibility criterion. Then **P** has a winning strategy for A under Protocol 2.

PROOF. Note that A is acceptable with respect to E, hence **P** can win by playing only arguments in E. Since E is defensible, it is conflict-free, hence **O** cannot make an *eo ipso* move nor **P** can be forced to play an argument that was previously played by **O**. ■

PROPOSITION 3.10
If **P** has a winning strategy for A according to Protocol 2 then $A \in E$ for some set $E \subseteq \mathcal{A}$ satisfying admissibility (i.e. condition (1)).

PROOF. Assume that **P** has a winning strategy $W(A)_{\mathbf{P}}$ for A following Protocol 2 and consider any history H of W in which **O** plays optimally. Let $H_{\mathbf{P}}$ be the set of arguments used by **P** in H and assume by contradiction that $H_{\mathbf{P}}$ is not defensible. But $H_{\mathbf{P}}$ is conflict-free since, by Protocol 2, **P** can not repeat arguments used by **O**. Hence there exists some argument $B \in H_{\mathbf{P}}$ which is not acceptable w.r.t. $H_{\mathbf{P}}$, i.e. there exists $C \in \mathcal{A}$, $C \succ B$ and $\nexists D \in H_{\mathbf{P}}$, $D \succ C$. Then **O** could win the game by playing D against C. Contradiction. ■

4 Game protocols and algorithms capturing cogency-related criteria

We define specific game protocols capturing pairwise, weak and cyclic cogency.

4.1 Protocol capturing pairwise cogency

To define a protocol for pairwise cogency, notice that the only case in which a pairwise cogent set S of arguments is not admissible is when S is attacked by some self-attacking argument: Assume S is pairwise cogent but not admissible. Then there exist $A \in S$ and $B \in \mathcal{A}$ such that $B \gg A$ and $S \not\gg B$. But if $\{B\}$ is conflict-free then $\{B\} >_{cog} S$. Nevertheless, S is pairwise cogent and hence $\{B\} \not>_{cog} S$. Therefore, $\{B\}$ is not conflict-free, implying that B is self-attacking. The intuition obtained from this discussion leads naturally to the following Protocol.

PROTOCOL 3. Protocol 2 plus

10. Neither player is allowed to play a self-attacking argument.

PROPOSITION 4.1
Let $A \in E$ for any extension $E \in \mathcal{E}_{\mathcal{S}}(AF)$ of a semantics $\mathcal{S}$ satisfying the pairwise cogent criterion. Then **P** has a winning strategy for A under Protocol 3.

PROOF. Assume that **P** has no winning strategy for A. Then, there exists some history leading **O** to win, which means that in this history either (i) **O** made an *eo ipso* move, or (ii) all the arguments that attack the last move of **O** were already deployed by **O**, or (iii) all the arguments that attack the last move of **O** are self-attacking. Cases (i) and (ii) imply that there exists an odd-length cycle of attacks involving an argument played by **P**. Hence, for every set S that **P** tries to build to defend A, there exists some set S' such that $S' >_{cog} S$. Hence A cannot belong to a pairwise cogent set. For case (iii), assume B is the last argument played by **O** and for every argument C, if $C \gg B$ then $C \gg C$. Then $\{B\}$ is pairwise cogent and for every set S that **P** tries to build for A, $\{B\} \gg S$. ■

PROPOSITION 4.2
If **P** has a winning strategy for A according to protocol 3, then $A \in E$ for some set $E \subseteq \mathcal{A}$ satisfying pairwise cogency (i.e. condition (3)).

PROOF. Assume **P** has a winning strategy $W(A)_{\mathbf{P}}$ for A following Protocol 3 and consider any history H obtained with **P** playing $W(A)_{\mathbf{P}}$ in which **O** plays optimally. Let $\mathbf{H_P}$ be the set of arguments used by **P** in H and assume by contradiction that $\mathbf{H_P} \notin Cog(AF)$. This implies that there must exist some set $S >_{cog} H_{\mathbf{P}}$. But since **O** has not played an *eo ipso* move in H, $\mathbf{H_P}$ is conflict-free, hence there exist some arguments $B \in S$ and $C \in \mathbf{H_P}$ such that $B \gg C$ and for any $D \in \mathbf{H_P}$, $D \not\gg B$ (in particular for $D = C$). Then **O** could win the game by playing B against C. Contradiction. ■

Winning strategies in dialogue games provide characterizations for the extensions of various semantics and can be rather easily translated into algorithms yielding those extensions. In fact, algorithms and complexity issues for admissibility have been already widely discussed [8, 9, 11, 12, 16, 25]. Up from them, a generic procedure can be distilled, which we call *proc_Protocol_2*, implementing Protocol 2. It is intended as a basic ingredient for the algorithms that will be introduced henceforth. In fact, since it is known that more efficiency is gained by forbidding **P** to play self-conflicting moves [16], we add this restriction. Moreover, since the pairwise cogency protocol differs from the admissibility protocol only in that self-attacking arguments are excluded from the legal moves of both players, any algorithm for admissibility can be adapted easily to become an algorithm for pairwise cogency. So, *proc_Protocol_2* with this modification yields an algorithm, *proc_Protocol_3* that will be invoked by the algorithms implementing the weak and cyclic cogency protocols.

4.2 *Protocol capturing weak cogency*

The intuition underlying weak cogency is that a proponent has enough justification for her thesis if any opposing theory (i.e. set of arguments) is not better (in terms of cogency) than her own theory. We think of a meta-argumentation game seen as a series of argumentation games, which will facilitate a definition as an algorithm which invokes modularly procedure *proc_Protocol_3*. The meta-game starts with **P** aiming to defend an argument A against **O** in an argumentation game under *proc_Protocol_3*. If **P** succeeds, the meta-game ends; otherwise, **O** has still to defend with a pairwise cogent set the argument B with which she has blocked **P** in the first phase of the meta-game. Now the burden of proof is on **O**, so a new argumentation game is in order, in which **P** just tries to make **O** unable to support B with a pairwise cogent set. If **O** succeeds at this round she wins the meta-argumentation game and **P** looses; otherwise **P** wins (**O** failed in her attempt to find a better theory than **P**'s). If **P** has a winning strategy in this meta-argumentation game, be it at the first or at the second phase, her argument A belongs to a weakly cogent set.

To avoid confusion with respect to the roles that **P** and **O** play in the two rounds of the game, let us introduce some new terminology which will facilitate the definition of the algorithm. Call *exponent* the player who has the burden of proof at a round of the game, and *challenger* the other player (i.e. the one who tries to defeat the exponent). The exponent and the challenger must play under the rules governing the proponent and the opponent, respectively, in the corresponding protocol. Also call *thesis* the argument for which the exponent has the burden of proof and *challenge* the argument that allowed the challenger to block the exponent in the previous stage.

PROTOCOL 4. The game is played according to this algorithm:

proc_Protocol_4

```
BEGIN
% The game starts distributing the roles of Exponent and
% Challenger, respectively.
Exponent := P
Challenger := O
RUN proc_Protocol_3 % Exponent aims to defend her thesis argument
                    % with a pairwise cogent set.
IF Exponent succeeds
    % P has proved that her thesis belongs to a pairwise cogent set.
THEN
    Payoff(P):=1
    Payoff(O):=-1
ELSE % Round 2: O has to defend her formerly challenge argument
      % with a pairwise cogent set, while P tries to show she can't.
    Exponent := O
    Challenger := P
    RUN proc_Protocol_3
    IF Exponent succeeds
    % A better proposal than P's first one has been found by O.
    THEN
        Payoff(P):=-1
        Payoff(O):=1
    ELSE
        Payoff(P):=1
        Payoff(O):=-1
END.
```

EXAMPLE 4.3

To illustrate the procedure, let us suppose that **P** is aimed to defend A in the argumentation framework depicted in Figure 2 in the left. On the right side, we show the game tree according to Protocol 4. Assume **O** plays optimally the first round under Protocol 3, attacking A with C. As **P** is blocked (B is in conflict with her thesis) **O** succeeds. Round 2 starts by shifting the burden of proof to **O** who has to defend C. Then **O** assumes the role of the exponent **P**, who in turn becomes the challenger. Now **P** succeeds playing B and blocking **O**. Therefore, **P** wins the meta-argumentation game. (Viewed as a rational choice problem, notice that $\{A\} \geq_{cog} \{B\}$, which is a reason for **P** not to dropping $\{A\}$ in favour of $\{B\}$.)

PROPOSITION 4.4
Let $A \in E$ for any extension $E \in \mathcal{E}_{\mathcal{S}}(AF)$ of a semantics $\mathcal{S}$ satisfying the weakly cogent criterion. Then **P** has a winning strategy for A under Protocol 4.

PROOF. By way of contradiction, assume that **P** has no winning strategy for A. If A is self-attacking it is clearly excluded from any weakly cogent set. Otherwise there exists some history in which **O** could win the second round game. This means that **P** failed to build a pairwise cogent set for A at the first round game, which implies that for any set S such that $A \in S$, **O** can play an argument $B \in S'$ for some set S' and $S' >_{cog} S$. At round 2, let B be the argument that **O** defends. Since **P** has no winning strategy, no matter which set of arguments T is used by **P** in her strategy at the second round, **O** can play a strategy with arguments drawn from some set of arguments T' such that $B \in T'$ and $T' \geq_{cog} T$. Therefore, A cannot belong to any weakly cogent set of arguments. Contradiction. ∎

PROPOSITION 4.5
If **P** has a winning strategy for A under protocol 4 then $A \in E$ for some set $E \subseteq \mathcal{A}$ satisfying weak cogency (i.e. condition (4)).

PROOF. Assume **P** has a winning strategy for A in a game played according to Protocol 4. Let us analyse two cases: (i) **P** wins at Round 1; (ii) **P** wins only at Round 2.

(i) In this case, the set of arguments in the winning strategy, $\mathbf{W}(A)_{\mathbf{P}}$ is a pairwise cogent set and, hence, a weakly cogent set.
(ii) Assume $i = \mathbf{P}$ has no winning strategy for A at Round 1 but has one at Round 2. First note that if **P** cannot win the game at Round 1 she is unable to build a set for A in $Cog(AF)$, that is, for every set S such that $A \in S$ there has to exist some set S' such that $S' >_{cog} S$. Now let B be an optimal choice of **P** among the non *eo ipso* arguments played by **O** at Round 1. This choice forces $-i$ to defend B at Round 2. Since at the second round game $i = \mathbf{O}$ wins the game, it must be that $-i = \mathbf{P}$ will not be able to build a set in $Cog(AF)$ for B. Hence, for any set S' such that $B \in S'$ and $S' >_{cog} S$ for any set S such that $A \in S$, there exists some subset S'' (containing some of the arguments played by **P** at Round 2) such that $S'' >_{cog} S'$. This implies that S satisfies (4). ∎

4.3 Protocol capturing cyclic cogency

To capture cyclic cogency, we follow here a similar approach as for weak cogency: **P** tries to defend her thesis with a good theory $T \subseteq \mathcal{A}$ according to Protocol 3. If **P** succeeds then she wins: T is a pairwise cogent set, hence also a cyclically cogent set. Otherwise a sequence of argumentation games begins with both players alternating the roles of exponent and challenger, each one trying to find a better theory than the last one played by her counterpart. **P**, in particular, is aimed to show that the sequence of dialogues will yield a cycle going back to T (in particular, to the argument in T, say B, that was blocked by **O**); **O**, in time, is aimed to show that some pairwise cogent set R used against T will be found in the way. Thus, unlike Protocol 4, the sequence of dialogues goes on until some player either finds a pairwise cogent set or repeats B. The meta-game is won by **O** if, and only if, a pairwise cogent set is found: the exponent, either **P** or **O**, has succeeded at some round (note that **P** herself could be who finds a better theory than her own theory T!). Otherwise, the meta-game is won by **P** since the challenger (either **P** or **O**) has succeeded in the last round with the help of B, confirming that a theory better founded than T has not been found.

PROTOCOL 5. The game is played according to this algorithm:

proc_Protocol_5

```
BEGIN
% The game starts distributing the roles of Exponent and
% Challenger, respectively.
Exponent := P
Challenger := O
RUN proc_Protocol_3 % Exponent aims to defend her thesis.
IF Exponent  succeeds
% P has proved that her thesis belongs to a pairwise cogent set
THEN
   Payoff(P):=1
   Payoff(O):=−1
ELSE
   blocked_argument :=  last argument played by P
   REPEAT % A run against the counterpart begins.
          player_var := Exponent
          Exponent := Challenger
          Challenger := player_var
          RUN proc_Protocol_3 % Exponent defends her
                                  % formerly challenging argument.
          S^Cha := {B ∈ 𝒜: B  is an argument
                  played by Challenger}
   UNTIL blocked_argument ∈ S^Cha
   IF Exponent  succeeds
          % A better proposal than P's first one has been found.
   THEN
          Payoff(P):=−1
          Payoff(O):=1
   ELSE
          % Challenger has repeated P's blocked argument and
          % Exponent has not succeeded.
          Payoff(P):=1
          Payoff(O):=−1
END.
```

EXAMPLE 4.6

Suppose we want to know whether A can be defended by **P** with a winning strategy in the argumentation framework $\langle\{A,B,C\},\{(A,B),(B,C),(C,A)\}\rangle$ (Figure 4). The first round will end with **O** blocking A with C (**P** cannot play B since she would be playing a self-conflicting strategy). Then the protocol forces **O** to show in a second game round that her challenging argument C belongs to a pairwise cogent set. But **P** can reply with B, leaving **O** out of moves. Then in the next round **O** can attack B only with A. In this way, **O** is forced to accept A and thus **P** wins the overall game.

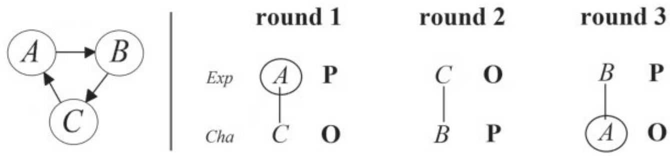

FIGURE 4. {A} is cyclically cogent. In the game, the opponent is forced to use the proponent's thesis.

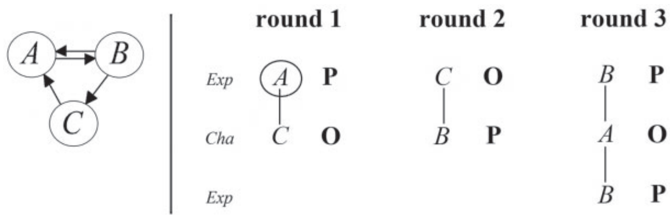

FIGURE 5. **P** wins the last round but she herself has found a better theory ({B}) than her first one ({A}), so she looses the meta-game.

EXAMPLE 4.7

Let us see now whether A belongs to some cyclically cogent set in the argumentation framework depicted in Figure 5. **O** has two options to attack A: B and C. Choosing B is not an optimal move since **P** can reply moving A again, forcing **O** to backtrack and start a new attack with C. So, assume **O** plays optimally in the first round moving C. Since the only argument attacking C is B, which is in conflict with A, **P** is blocked at A and looses the first round. Now **O** has the burden of proof on her challenging argument C: she has to assume the role of the exponent and show that C belongs to a pairwise cogent set. Once she advances C as her thesis in this second round, **P** replies with B. This leaves **O** out of moves since A, the only attacking argument of B, is in conflict with her previous move C. Then a third round starts with **P** defending the argument she used to challenge **O**, B. Then **O** challenges B playing A, her only possible move, but **P** can defend successfully B. At this point, we have the particular situation in which **P** wins this last round but looses the meta-game: **P**, by herself, has found a better theory than her initial one since $\{B\}$ is a pairwise cogent (admissible, indeed) set. (Also notice that **O** has used **P**'s blocked thesis A, but without success). In fact, $\{B\} >_{cog} \{C\} >_{cog} \{A\}$ but $\{A\} \not>_{cog} \{B\}$, hence A does not belong to a cyclically cogent set.

Next we show the results ensuring the correspondence between cyclic cogency and winning strategies under Protocol 5.

PROPOSITION 4.8

Let $A \in E$ for any extension $E \in \mathcal{E}_{\mathcal{S}}(AF)$ of a semantics $\mathcal{S}$ satisfying the cyclic cogency criterion. Then **P** has a winning strategy for A under Protocol 5.

PROOF. If A belongs to a pairwise cogent or to an admissible set, **P** is ensured to win at the first run of *proc_Protocol_3*. So let us analyse the case where A belongs to some cyclically, non-pairwise cogent set. In this case, **P** does not succeed in the first round game (she is left out of moves since **O** forces her to use a self-conflicting strategy). But since A belongs to a cyclically cogent set, it follows that either A or an argument defending (directly or indirectly) A are involved in an odd-length cycle of attacks. In either case, the second round starts with **O** trying to show that her blocking argument belongs to a pairwise cogent set. But **O** also fails being pushed by the challenger, **P**, into a self-conflicting strategy as well. Then a new round starts in which **O** challenges **P**'s challenging argument, and so on. Note now that for any odd-length k of the cycle, the player who has the role of challenger will repeat, in the $[(k+1)/2+1]$th round, the argument of **P** blocked in the first game round (i.e. $B \in \mathbf{S}_r^{Cha}$, where B is the blocked argument of **P** and $\mathbf{S}_r^{Cha}$ is obtained from the strategy played by the challenger in the round $r=(k+1)/2+1$). Hence, the game ends with a successful defence of A by **P** according to Protocol 5. ■

Proposition 4.9
If **P** has a winning strategy for A under protocol 5 then $A \in E$ for some set $E \subseteq \mathcal{A}$ satisfying cyclic cogency.

Proof. Assume **P** has a winning strategy for A under Protocol 5, and let us prove that A belongs to a cyclically cogent set. Assume **O** plays optimally when **P** plays her winning strategy. If **P** wins the first run of proc_Protocol_3 then A belongs to a pairwise cogent set and hence to a cyclic cogent set. Otherwise, **P** wins after several game rounds. Clearly, **O** won the first round. But since **P** still has a winning strategy, it follows that the strategy played by **O** does not yield a pairwise cogent set in the second round. Thus, the players play arguments involved in an odd-length cycle. Let C be the blocking argument of **O** in the first round. Then C does not belong to a pairwise cogent set since she looses in the second round, and the same is the case for all the Exponents of the following game rounds. Thus, according to the protocol, some player repeats the argument of **P** blocked in the first game round. Hence, for any set S of arguments that **O** can advance in her strategy (in particular, anyone such that $C \in S$) **P** can close a cogency cycle with a set S' such that $\{A, B\} \subseteq S'$. Therefore, A belongs to a cyclically cogent set. ■

4.4 *Sceptical games and cogency-related criteria*

Finally, an interesting question is whether justification games can be defined as to capture the arguments included in every extension that satisfies a given cogency criterion (sceptical decision problem). This is a difficult question to address in the case of preferred semantics [16] and it seems to be so also for cogency-based semantics. A strategy followed in [26] is to impose symmetric obligations on both players by asking **O** to defend her challenging argument satisfying the same requisites as those imposed on **P**. If **O** succeeds, she has proved the existence of a set that satisfies those requirements without including **P**'s thesis. But this strategy only works for preferred extensions which are also stable (i.e. extensions that attack every external argument) and not for general preferred semantics. It does not work for general cogency-based semantics neither. Nevertheless, we can show that it works for semantics sanctioning maximally pairwise, weakly or cyclic cogent sets (namely sustainable, lax or tolerant semantics).

Protocol 6. (Sceptical games)
If **P** is successful playing under protocol x, $x \in \{3, 4, 5\}$, **O** is prompted to defend her challenging argument under the same protocol x.

Proposition 4.10
P has a winning strategy for A under protocol 6 iff $A \in E$ for every set $E \subseteq \mathcal{A}$ maximally (w.r.t. $\subseteq$) satisfying the cogency criterion of protocol x.

Proof. ($\rightarrow$) Protocol 3: Let $A \in S$, where S is the pairwise cogent set supplying a winning strategy to **P** in the defence of A, and let B be any challenging argument (unsuccessfully) used by **O** (B does

not belong to any pairwise cogent set). Assume now by contradiction that S' is a maximal pairwise cogent set such that $A \notin S'$. By Lemma 2.13, we have that if $S' \cup S$ is conflict-free then it is pairwise cogent. But this contradicts the maximality of S', hence it is not conflict-free, i.e. there exist some attack among S' and S. Now, as they are pairwise cogent sets, each set can reply any attack from the other set; in particular, we will have an attack from S' to S. But then this attack can be used by **O** to build a pairwise cogent set with elements of S' not containing A. This contradicts the hypothesis.
Protocol 4: Let us take the same assumptions as in the previous proof, but assuming now that S is weakly cogent and that, by way of contradiction, S' is a maximal weakly cogent set such that $A \notin S'$. Since **O** cannot build a weakly cogent set attacking A, it follows that S' does not attack S, in particular, $S' \not\gg A$. But $S' \cup \{A\}$ cannot be conflict-free, otherwise by lemma 2.13 $S' \cup \{A\}$ would be weakly cogent, contradicting the maximality of S'. Hence $A \gg S'$, which implies that (*) $\{A\} >_{cog} S'$. But since S' is weakly cogent, it follows that there exists some set T such that $T >_{cog} \{A\}$. In turn, T cannot be weakly cogent (otherwise it would be used by **O**). Hence there exists a pairwise cogent set T' such that $T' >_{cog} T$. Note that T' cannot attack A (otherwise it would be used by **O** again), while A cannot attack T', otherwise it would not be pairwise cogent. But then $T' \cup \{A\}$ is pairwise cogent and moreover by (*) we have that $T' \cup \{A\} >_{cog} S'$. This contradicts that S' is weakly cogent.
Protocol 5: It follows immediately from the previous proof and the fact that every cyclic cogent set is also weakly cogent.
($\leftarrow$) Let $A \in E$ for every set $E \subseteq \mathcal{A}$ which maximally (w.r.t. $\subseteq$) satisfies the cogency criterion corresponding to protocol x (let us call such sets 'x-cogent'). Then for each protocol x, **P** has a winning strategy for A under protocol x. Assume by way of contradiction that **O** wins under protocol 6. Then **O** has found an x-cogent set S such that $B \in S$ where B is **O**'s challenging argument, i.e. $B \gg A$. But, by hypothesis, there exists some maximal x-cogent set T such that $\{A, B\} \subseteq T$. Then T is not conflict-free, contradicting that it is x-cogent. ■

5 Discussion

5.1 Non-admissible semantics from the point of view of cogency criteria

Extension semantics which do not satisfy admissibility have not been widely explored. Among them, sustainable semantics satisfy pairwise cogency, tolerant semantics satisfy cyclic cogency [5] and lax semantics only satisfies weak cogency [4]. This is straightforward since, in each case, the extensions are defined as to be the maximal (w.r.t. $\subseteq$) subsets of arguments satisfying the corresponding criterion. $CF2$ [2] and stage [23] semantics, on the other hand, do not satisfy neither weak nor cyclic cogency, though their behaviour is in many cases similar to that of lax and tolerant semantics regarding self-attacking arguments and odd-length cycles of attack. The argumentation framework depicted in Figure 2 is a counterexample where $\{C\}$ is a $CF2$ extension but it is neither weakly nor cyclic cogent. For stage semantics, consider the argumentation framework $\langle \{A, B, C\}, \{(A, A), (A, B), (B, C), (C, A)\} \rangle$ where the stage extension $\{C\}$ is neither weakly nor cyclic cogent too, since $\{B\} >_{cog} \{C\}$ and $\{B\}$ is pairwise cogent. Table 1 summarizes the above remarks.

5.2 Cogency criteria from a rational choice point of view

Admissibility and pairwise cogency can be interpreted as models of rational choice, with admissibility corresponding to a stringent maximalization function while pairwise cogency corresponds to a liberal maximalization function [14]. A stringent maximalization function chooses the subset of alternatives $\{S : S \succeq S'$, for any alternative $S'\}$, where '$\succeq$' is a reflexive and transitive preference relation, while the

TABLE 1. Non-admissible semantics comparison through cogency criteria

Semantics	*Pairwise cogency*	*Cyclic cogency*	*Weak cogency*
CF2	No	No	No
Stage	No	No	No
Lax	No	No	Yes
Tolerant	No	Yes	Yes
Sustainable	Yes	Yes	Yes

class of admissible subsets of $\mathcal{A}$ is $\{S : S \geq_{cog} S'$, for any subset $S' \subset \mathcal{A}\}$. On the other hand, a liberal maximalization function chooses the subset of alternatives $\{S : S' \not\succ S$, for any alternative $S'\}$, while the class of pairwise cogent subsets of $\mathcal{A}$ is $\{S : S' \not>_{cog} S$, for any subset $S' \subset \mathcal{A}\}$. Nevertheless, we cannot disregard a fundamental difference between preference relationships, on which the rational choice functions are defined, and the cogency relationship: while preferences are transitive, cogency is not. This notwithstanding, the cogency relationship can be seen as a justification for the use of cycles in argumentative disputes, being natural and not deprived of rationality.

On the other hand, while loops tend to induce problems in preference relationships, they arise with naturality in strategic decision-making, as shown in Example 2.8. So, if decisions are to be supported by arguments it seems sensible to analyse also the possible cycles in the argumentation process. In [5], cyclically cogent sets of arguments are put in correspondence with rationalizable strategies in a two-player normal-form game: every S is chosen as the best response to any S' such that $S \geq_{cog} S'$. Here, instead, we focus on extensive-form games under protocols leading to different types of dialogues. The most involved protocol from those presented in this article was designed for cyclic cogency. It characterizes the cycling argumentation process as an alternation of partial dialogues intended to find, again, a 'rationalizable theory'.

6 Conclusion

Dialogue games are usually studied as proof-theoretical counterparts of extension semantics. Our work goes one step further, relating them to general criteria of argument defence. In this sense, we have focused on a precise game-theoretical framework and have shown that particular protocols of play can be defined in such way that the arguments in the ensuing winning strategies constitute the extensions that satisfy general criteria of defence. Moreover, all the protocols presented here can be algorithmically implemented, ensuring the finiteness of the game.

In particular, we have shown this characterization for three defence criteria weaker than admissibility: pairwise, weak and cyclic cogency. These criteria exploit the intuition that acceptability depends on how argument strategies are able to protect their arguments. Unlike admissibility, which is defined in terms of defending an argument against *any* attack, cogency requires to defend an argument only against *coordinated* attacks. This yields an immediate dialectical interpretation: the opponent is urged to build a more convincing argument strategy than the proponent's. Pairwise cogency offers a rational criterion for solving undesired interferences of self-attacking arguments. In turn, weak and cyclical cogency, with different degrees of credulity, yield rationalizable choices —in the game-theoretical meaning of the term— of arguments involved in odd-length cycles of attack.

Acknowledgements

We are indebted to two anonymous referees who made helpful remarks that improved the article, especially Reviewer 2 who posed the question about the sceptical decision problem with regards to cogency criteria. This work was partially supported by Agencia Nacional de Promociones Científicas y Técnicas, Argentina, under grant PICT-2013-1489, and Universidad Nacional del Sur in Bahía Blanca, Argentina.

References

[1] L. Amgoud and C. Cayrol. A reasoning model based on the production of acceptable arguments. *Annals of Mathematics and Artificial Intelligence*, **34**, 197–215, 2002.

[2] P. Baroni, M. Giacomin, and G. Guida. SCC-recursiveness: a general schema for argumentation semantics. *Artificial Intelligence*, **168**, 162–210, 2005.

[3] P. Baroni and M. Giacomin. On principle-based evaluation of extension-based argumentation semantics. *Artificial Intelligence*, **171**, 675–700, 2007.

[4] G. Bodanza. The 'lottery paradox' paradox and other self-attacking arguments from the point of view of defeasible argumentation frameworks. In *The Many Sides of Logic*. Studies in Logic 21, W. Carnielli, M. Coniglio, and I. D'Ottaviano, eds, pp. 481–496. College Publications, 2009.

[5] G. Bodanza and F. Tohmé. Two approaches to the problems of self-attacking arguments and general odd-length cycles of attack. *Journal of Applied Logic*, **7**, 403–420, 2009.

[6] G. Bodanza, F. Tohmé, and G. R. Simari. Argumentation games for admissibility and cogency criteria. In *Computational Models of Argument—Proceedings of COMMA 2012, Austria, Frontiers in Artificial Intelligence and Applications, v. 245*, B.Verheij, S. Szeider, S. Woltran, eds, pp. 153–164. IOS Press, 2012.

[7] M. Caminada. On the issue of reinstatement in argumentation. In *Proceedings of Logics in Artificial Intelligence, 10th European Conference, JELIA*, LNCS 4160, pp. 111–123. Springer 2006.

[8] C. Cayrol, S. Doutre, and J. Mengin. On decision problems related to the preferred semantics for argumentation frameworks. *Journal of Logic and Computation*, **13**, 377–403, 2003.

[9] Y. Dimopoulos, B. Nebel, and F. Toni. On the computational complexity of assumption-based argumentation for default reasoning. *Artificial Intelligence*, **141**, 57–78, 2002.

[10] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, **77**, 321–357, 1995.

[11] P. M. Dung, P. Mancarella, and F. Toni. Computing ideal sceptical argumentation. *Artificial Intelligence*, **171**, 642–674, 2007.

[12] P. E. Dunne and T. J. M. Bench-Capon. Two party immediate response disputes: properties and eficiency. *Artificial Intelligence*, **149**, 221–250, 2003.

[13] D. Gale and F. M. Stewart. Infinite games with perfect information. *Annals of Mathematics Studies*, **28**, 245–266, 1953.

[14] H. G. Herzberger. Ordinal preference and rational choice. *Econometrica*, **41**, 187–237, 1973.

[15] H. Jakobovits and D. Vermeir. Dialectic semantics for argumentation frameworks. In *Proceedings of the 7th ICAIL*, Oslo, pp. 53–62. ACM, 1999.

[16] S. Modgil and M. Caminada. Proof theories and algorithms for abstract argumentation frameworks. In *Argumentation in AI*, I. Rahwan, G. R. Simari, eds, pp. 105–132. Springer, 2009.

[17] Y. Moschovakis, *Descriptive Set Theory*, North-Holland, 1980.

[18] M. Osborne and A. Rubinstein, *A Course in Game Theory*, MIT Press, 1994.

[19] J. Pollock. Self-defeating arguments. *Minds and Machines*, **1**, 367–392, 1991.

[20] I. Rahwan, M. I. Madakkatel, J.-F. Bonnefon, R.N. Awan, and S. Abdallah. Behavioral experiments for assessing the abstract argumentation semantics of reinstatement. *Cognitive Science*, **34**, 1483–1502, 2010.

[21] P. M. Thang, P. M. Dung, and N. D. Hung. Towards a common framework for dialectical proof procedures in abstract argumentation. *Journal of Logic and Computation*, **19**, 1071–1109, 2009.

[22] I. Viglizzo, F. Tohmé, and G. R. Simari. The foundations of DeLP: defeating relations, games and truth values. *Annals of Mathematics and Artificial Intelligence*, **57**, 181–204, 2010.

[23] B. Verheij. Two approaches to dialectical argumentation: admissible sets and argumentation stages. In *Proceedings of the Eighth Dutch Conference on Artificial Intelligence (NAIC '96)*, J.-J.Ch. Meyer, L. C. van der Gaag, eds, pp. 357–368. Utrecht University, 1996.

[24] G. A. W. Vreeswijk. Defeasible dialectics: a controversy-oriented approach towards defeasible argumentation. *Journal of Logic and Computation*, **3**, 3–27, 1993.

[25] G. A. W. Vreeswijk. An algorithm to compute minimally grounded and admissible defence sets in argument systems. In *Proceedings Computational Models of Argument (COMMA 2006)*, pp. 109–120. 2006.

[26] G. A. W. Vreeswijk and H. Prakken, Credulous and sceptical argument games for preferred semantics. In *Proceedings 7th European Workshop on Logic for Artificial Intelligence*, pp. 239–253. 2000.

Received 13 September 2013