

Explanation for Case-Based Reasoning via Abstract Argumentation

Kristijonas ČYRAS^{a,1}, Ken SATOH^b and Francesca TONI^a

^aImperial College London, UK

^bNational Institute of Informatics, Tokyo, Japan

Abstract. Case-based reasoning (CBR) is extensively used in AI in support of several applications, to assess a new situation (or case) by recollecting past situations (or cases) and employing the ones most similar to the new situation to give the assessment. In this paper we study properties of a recently proposed method for CBR, based on instantiated Abstract Argumentation and referred to as AA-CBR, for problems where cases are represented by abstract factors and (positive or negative) outcomes, and an outcome for a new case, represented by abstract factors, needs to be established. In addition, we study properties of explanations in AA-CBR and define a new notion of lean explanations that utilize solely relevant cases. Both forms of explanations can be seen as dialogical processes between a proponent and an opponent, with the burden of proof falling on the proponent.

Keywords. Case-Based Reasoning, Abstract Argumentation, Explanation

1. Introduction

Case-based reasoning (CBR), as overviewed in [28], is extensively used in various applications of AI (see e.g. [23,28]). At a high-level, in CBR a reasoner in need to assess a new situation, or *new case*, recollects past situations, or *past cases*, and employs the ones most similar to the new situation to give the assessment. Several approaches to CBR use (forms of) argumentation, e.g. [1,27] and, more recently, the AA-CBR approach of [11].

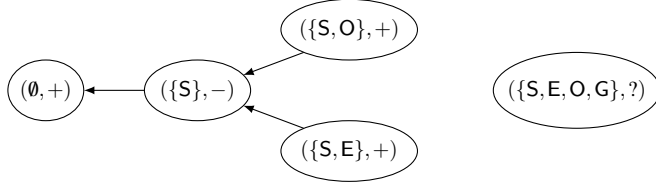
AA-CBR instantiates Abstract Argumentation (AA) [12] to resolve conflicts amongst most similar past cases with diverging outcomes. It provides: 1) a method for computing outcomes for new cases, given past cases and a *default* outcome; and 2) *explanations* for computed outcomes, as dialogical exchanges between a proponent, in favour of the default outcome for the new case, and an opponent, against the default outcome.

As common in the literature (see e.g. [5,25,28]), in AA-CBR past cases are represented as sets of *factors* (also known as features or attribute-value pairs, cf. [30]) together with an outcome, which may be positive (+) or negative (−). AA-CBR then relies upon the *grounded extension* [12] of an AA framework with, as arguments, a default case (with an empty set of factors and the default outcome), past cases (with their outcomes) and a new case (with unknown outcome). A past case attacks another past case or the default

¹Corresponding Author: Kristijonas Čyras, Department of Computing, Imperial College London, United Kingdom; E-mail: k.cyras13@imperial.ac.uk.

case if they have a different outcomes, the former is more specific than the latter and at least as concise as any other similarly more specific, conflicting past case. The following example, used or adapted throughout the paper, illustrates AA-CBR.

Example 1. Suppose Bob wishes to rent his spare room to get between £800 and £900 per month, and decides to use an online AA-CBR system to determine whether this amount is reasonable and why, based on similar lodgings being rented. Let N , the new case, represent the set of features of Bob's room, e.g. $N = \{S, E, O, G\}$ (the room is Small, with an En-suite bathroom in an Open-plan flat with a Gym in the building). Here the default outcome is $+$, indicating Bob's bias for the price range £800–£900. The past cases are either of the form $(X, +)$, for lodgings in the desired price range, or $(Y, -)$, for lodgings in different (lower or higher) price ranges, with X, Y the feature sets of these lodgings. For example, suppose the past cases are $(\{S\}, -)$ (Small rooms go for lower prices), $(\{S, E\}, +)$ (En-suite compensates for Small room), $(\{S, O\}, +)$ (Open-plan flat compensates for Small). Then, the corresponding (instantiated) AA framework [11] is depicted below (with attacks represented by arrows, $(\emptyset, +)$ the argument for the default case, and $(\{S, E, O, G\}, ?)$ the argument for the new case):



$\mathbb{G} = \{(\{S, E, O, G\}, ?), (\{S, E\}, +), (\{S, O\}, +), (\emptyset, +)\}$ is the grounded extension of this AA framework. Since $(\emptyset, +) \in \mathbb{G}$, the outcome for the new case determined by AA-CBR is $+$, with two possible explanations \mathcal{T}_P and \mathcal{T}'_P depicted below (with P standing for proponent and O standing for opponent):



Thus, for example, \mathcal{T}_P explains the recommendation $+$ dialectically as follows: the default outcome $+$ needs to be defended against the objection posed by past case $(\{S\}, -)$, and this can be achieved by using past case $(\{S, E\}, +)$, that cannot be objected against.

In this paper we propose a novel form of explanations, called *lean explanations*, and study properties of both forms of explanations in AA-CBR. Explanations can naturally be seen as dialogical exchanges between a proponent and an opponent, the former having the burden of proof for explaining as well as establishing the outcome of the new case.

The paper is organized as follows. We first recall, in Section 2, the necessary background. In Section 3 we prove some properties of AA-CBR, in the context of some related work, and in Section 4 we investigate properties of explanations in AA-CBR. We then allot Section 5 to relate AA-CBR with proof standards and burden of proof. We conclude with a discussion on related and future work in Section 6.

2. Background

AA-CBR [11] assumes a fixed but otherwise arbitrary (possibly infinite) set \mathbb{F} of *factors*, and a set $\{+, -\}$ of *outcomes*, one of which is singled out as the *default outcome* d . The *complement* of d is indicated as \bar{d} , and is: $+$ if $d = -$; and $-$ if $d = +$. A *case* is a pair (X, o) with $X \subseteq \mathbb{F}$ and $o \in \{+, -\}$; a *case base* is a finite set $CB \subseteq \mathcal{P}(\mathbb{F}) \times \{+, -\}$ such that for $(X, o_X), (Y, o_Y) \in CB$, if $X = Y$, then $o_X = o_Y$; a *new case* is a set $N \subseteq \mathbb{F}$.

AA-CBR maps the problem of determining the outcome for a new case into a membership problem within the grounded extension of an AA framework [12] obtained from the case base CB , the new case N and the default outcome d . In general, following [12], an *AA framework* is a pair $(\text{Args}, \rightsquigarrow)$, where Args is a set (of *arguments*) and \rightsquigarrow is a binary relation on Args (where, for $\mathbf{a}, \mathbf{b} \in \text{Args}$, if $\mathbf{a} \rightsquigarrow \mathbf{b}$, then we say that \mathbf{a} *attacks* \mathbf{b}). For a set of arguments $E \subseteq \text{Args}$ and an argument $\mathbf{a} \in \text{Args}$, E *defends* \mathbf{a} if for all $\mathbf{b} \rightsquigarrow \mathbf{a}$ there exists $\mathbf{c} \in E$ such that $\mathbf{c} \rightsquigarrow \mathbf{b}$. Then, the *grounded extension* of $(\text{Args}, \rightsquigarrow)$ can be constructed as $\mathbb{G} = \bigcup_{i \geq 0} G_i$, where G_0 is the set of all unattacked arguments, and $\forall i \geq 0$, G_{i+1} is the set of arguments that G_i defends. For any $(\text{Args}, \rightsquigarrow)$, the grounded extension \mathbb{G} always exists and is unique, and, if $(\text{Args}, \rightsquigarrow)$ is well-founded [12], extensions under other semantics are equal to \mathbb{G} . AA-CBR uses the following instance of AA [11]:

Definition 2. The AA framework corresponding to a case base CB , a default outcome $d \in \{+, -\}$ and a new case N is $(\text{Args}, \rightsquigarrow)$ satisfying the following conditions:

- $\text{Args} = CB \cup \{(\emptyset, d)\} \cup \{(N, ?)\}$;
- for $(X, o_X), (Y, o_Y) \in CB \cup \{(\emptyset, d)\}$, it holds that $(X, o_X) \rightsquigarrow (Y, o_Y)$ iff
 - * $o_X \neq o_Y$, and (different outcomes)
 - * $Y \subsetneq X$, and (specificity)
 - * $\nexists (Z, o_Z) \in CB$ with $Y \subsetneq Z \subsetneq X$; (concision)
- for $(Y, o_Y) \in CB$, $(N, ?) \rightsquigarrow (Y, o_Y)$ holds iff $Y \not\subseteq N$.

$(N, ?)$ is referred to as the *new case argument* and (\emptyset, d) as the *default case*.

In what follows, $(\text{Args}, \rightsquigarrow)$ is the AA framework corresponding to a given, generic CB , d and N , and \mathbb{G} is its grounded extension. Note the following: $(\text{Args}, \rightsquigarrow)$ is finite (as case bases are); $\mathbb{G} \neq \emptyset$ (as $(N, ?)$ is unattacked); $(\text{Args}, \rightsquigarrow)$ is well-founded (due to the specificity requirement in Definition 2), so that \mathbb{G} is a unique extension under other semantics. AA-CBR decides the outcome for the new case as follows [11]:

Definition 3. The AA outcome of the new case N is:

- the default outcome d , if $(\emptyset, d) \in \mathbb{G}$;
- \bar{d} , otherwise, if $(\emptyset, d) \notin \mathbb{G}$.

In AA-CBR, explanations for AA outcomes are defined in terms of *dispute trees* [11,13,14], where a *dispute tree* for $\mathbf{a} \in \text{Args}$ is a tree \mathcal{T} such that:

1. every node of \mathcal{T} is of the form $[L:\mathbf{x}]$, with $L \in \{P, O\}$, $\mathbf{x} \in \text{Args}$: the node is *labelled* by argument \mathbf{x} and assigned the *status* of either *proponent* (P) or *opponent* (O);
2. the root of \mathcal{T} is a P node labelled by \mathbf{a} ;
3. for every P node n , labelled by some $\mathbf{b} \in \text{Args}$, and for every $\mathbf{c} \in \text{Args}$ such that $\mathbf{c} \rightsquigarrow \mathbf{b}$, there exists a child of n , which is an O node labelled by \mathbf{c} ;
4. for every O node n , labelled by some $\mathbf{b} \in \text{Args}$, there exists at most one child of n which is a P node labelled by some $\mathbf{c} \in \text{Args}$ such that $\mathbf{c} \rightsquigarrow \mathbf{b}$;
5. there are no other nodes in \mathcal{T} except those given by 1–4.

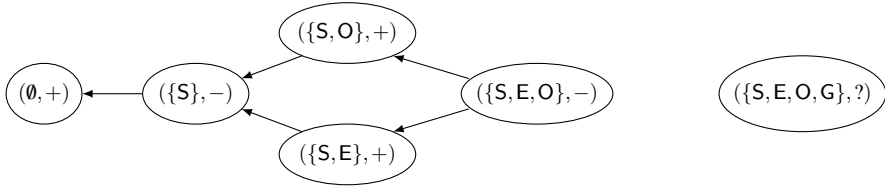
The *defence set* of \mathcal{T} , denoted by $\mathcal{D}(\mathcal{T})$, is the set of all arguments labelling P nodes in \mathcal{T} . A dispute tree \mathcal{T} is an *admissible dispute tree* iff (i) every 0 node in \mathcal{T} has a child, and (ii) no argument in \mathcal{T} labels both P and 0 nodes. A dispute tree \mathcal{T} is a *maximal dispute tree* [11] iff for all opponent nodes $[0:x]$ which are leaves in \mathcal{T} there is no $y \in \text{Args}$ such that $y \rightsquigarrow x$. Note that an admissible dispute tree \mathcal{T} for some $a \in \text{Args}$ is also a maximal dispute tree for $a \in \text{Args}$ [11, Lemma 4]. Indeed, in an admissible dispute tree, each 0 node has exactly one child (a P node); thus, no 0 node is a leaf, and so the dispute tree is maximal.

Explanations in AA-CBR are defined as follows [11]:

Definition 4. If the AA outcome of N is d , then an *explanation for why the AA outcome of N is d* is any admissible dispute tree for (\emptyset, d) . If the AA outcome of N is \bar{d} , then an *explanation for why the AA outcome of N is \bar{d}* is any maximal dispute tree for (\emptyset, d) .

Example 1 illustrates the notion of explanation for why the outcome is d . The following example illustrates the notion of explanation for why the outcome is \bar{d} .

Example 5 (Example 1 ctd.). Suppose there is an additional case $(\{S, E, O\}, -)$ in CB . Then the corresponding AA framework is depicted below:



Here, $\mathbb{G} = \{(\{S, E, O, G\}, ?), (\{S, E, O\}, -), (\{S\}, -)\}$, so the AA outcome of $\{S, E, O, G\}$ is $-$, for which the dispute trees (in linear notation) $\mathcal{T}_0 : [P : (\{\emptyset\}, +)] - [O : (\{S\}, -)] - [P : (\{S, E\}, +)] - [O : (\{S, E, O\}, -)]$ and $\mathcal{T}'_0 : [P : (\{\emptyset\}, +)] - [O : (\{S\}, -)] - [P : (\{S, O\}, +)] - [O : (\{S, E, O\}, -)]$ are explanations.

3. Properties of AA outcomes

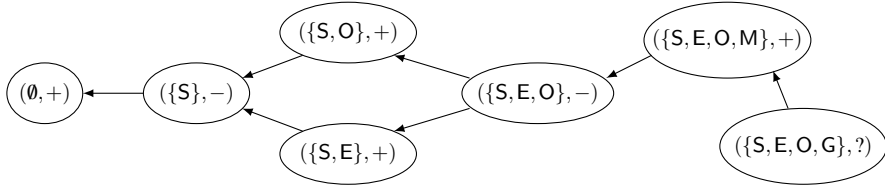
In this section, we prove several properties of AA outcomes, focusing on aspects that have been considered in some related work on CBR. Where indicated, these properties were stated in [11] already, but their proofs omitted there.

In the context of, particularly, legal CBR, as well as CBR in general, two properties are identified as important [3], namely that cases employed in determining the outcome of a new case N should be *most on point* and *untrumped*, where (X, o) is:

most on point iff no other case with the same outcome shares a more inclusive set of factors with the new case, i.e. $X \cap N$ is \subseteq -maximal for $X' \cap N$ with $(X', o) \in CB$;
untrumped iff no counterexample is more on point, i.e. there is no $(Y, o_Y) \in CB$ satisfying both $o_Y \neq o$ and $X \cap N \subsetneq Y \cap N$.

These two constraints together can be summarized into a single condition of $X \cap N$ being \subseteq -maximal among all $Y \cap N$ with $(Y, o') \in CB$. These properties—of being most on point and untrumped—allow for ‘deviating’ factors (i.e. factors not present in N) amongst past cases to be used to determine the outcome of the new case, whereas AA outcome does not. For an illustration, consider the following modification of our running example.

Example 6. Suppose there is an additional past case $(\{S, E, O, M\}, +)$ in Example 5: a Motorway next to the building is disadvantageous and the price of a Small En-suite room in an Open plan flat next to a Motorway falls into Bob's price range:



Here, both $(\{S, E, O\}, -)$ and $(\{S, E, O, M\}, +)$ are most on point and untrumped, yet $(\{S, E, O, M\}, +)$ is attacked by the new case and so effectively discarded from influencing the AA outcome, which is $-$, as in Example 5.

The notion of AA outcome fulfils a variant of the properties of being most on point and untrumped, that disregards ‘deviating’ factors and focuses instead on *nearest* cases [11], where (X, o_X) is

nearest to N iff $X \subseteq N$, and $\nexists (Y, o_Y) \in CB$ with $Y \subseteq N$ and $X \subsetneq Y$.

In other words, (X, o_X) is nearest to N iff $X \subseteq N$ is \subseteq -maximal in the case base. In Example 6, $(\{S, E, O\}, -)$ is nearest, but $(\{S, E, O, M\}, +)$ is not, because $M \notin N$.

Like elsewhere in the literature, e.g. [22,23,28], in AA-CBR nearest cases are very important. In particular, when CB contains a single nearest case (X, o) to N , the AA outcome of N is fully determined by (X, o) , independently of what d is, as follows:

Proposition 1 ([11, Proposition 2]). *If there is a unique nearest case (X, o) to N , then, for any $d \in \{+, -\}$, the AA outcome of N is o .*

Proof. Let $(X, o) \in CB$ be the unique nearest case to N . Consider a chain of attacks $(Y, o_Y) \rightsquigarrow \dots \rightsquigarrow (\emptyset, d)$, with $n \geq 1$ arguments and (Y, o_Y) unattacked in $(Args, \rightsquigarrow)$. First, we know that $(Y, o_Y) \in \mathbb{G}$. Assuming $o_Y \neq o$, we find $Y \subsetneq X$ (as (X, o) is unique nearest to N), whence $(X, o) \rightsquigarrow (Y, o_Y)$ gives a contradiction. So $o_Y = o$. Thus, if $o = d$, then n is odd, and so \mathbb{G} defends (\emptyset, d) , so that $(\emptyset, d) \in \mathbb{G}$. Else, if $o = \bar{d}$, then n is even, so that \mathbb{G} attacks (\emptyset, d) , and so $(\emptyset, d) \notin \mathbb{G}$. In any case, the AA outcome of N is o . \square

In Example 6, the AA outcome of $\{S, E, O, G\}$ is $-$, the outcome of the unique nearest case $(\{S, E, O\}, -)$ (and the complement of the default outcome $+$). If instead the default outcome was $-$, the structure of the AA framework would change (in particular, the attack relation would be different), but the AA outcome would remain unchanged.

In AA-CBR, nearest cases are important as they belong to the grounded extension:

Proposition 2 ([11, Lemma 1]). \mathbb{G} contains all the nearest past cases to N .

Proof. Let $(X, o_X) \in CB$ be nearest to N . Then $X \subseteq N$, so $(N, ?) \not\rightsquigarrow (X, o_X)$. Now assume that $(Y, o_Y) \rightsquigarrow (X, o_X)$, for some $(Y, o_Y) \in CB$. Then $Y \not\subseteq N$, whence $(N, ?) \rightsquigarrow (Y, o_Y)$. Since the new case argument $(N, ?)$ is unattacked in $(Args, \rightsquigarrow)$, we have $(N, ?) \in \mathbb{G}$. As (Y, o_Y) was arbitrary, we know that \mathbb{G} defends (X, o_X) , so that $(X, o_X) \in \mathbb{G}$. \square

This result shows that AA-CBR takes into account all the most similar past cases when determining the outcome of a new case. This is in contrast with some forms of the

conventional k -nearest neighbour approaches to CBR, where some of the nearest cases may be ignored in order to decide the new case [28] (see also Section 6 for a discussion).

Note that \mathbb{G} contains not only the nearest cases (as well as the new case N), but also some other past cases: in Examples 5, 6, \mathbb{G} includes $(\{S\}, -)$, which is not nearest to N , but still ‘relevant’ to the AA outcome. Overall, past cases, as arguments, can be classified into those deemed *relevant* and *irrelevant* for deciding the new case, as follows:

Definition 7. An argument $(X, o) \in \text{Args} \setminus \{(\emptyset, d)\}$ is said to be:

- **relevant** if $X \cap N \neq \emptyset$;
- **irrelevant** otherwise, if $X \cap N = \emptyset$.

By convention, the default case (\emptyset, d) is also deemed relevant.

Since arguments in AA-CBR are cases, with an abuse of notation we sometimes talk about cases being relevant and irrelevant.

In Example 6, all cases (including the default case) are relevant. If there was, say, a case $(\{H\}, -)$ in CB , it would be irrelevant, as $\{H\} \cap \{S, E, O, G\} = \emptyset$.

The relevance criteria defined above will play a role in characterizing explanations of AA outcomes, which we will investigate in the next section.

4. Explanations of AA Outcomes

The notion of AA outcome allows to determine algorithmically whether a new case N should be assigned the default outcome (d) or not (\bar{d}), by determining whether the default case (\emptyset, d) belongs or not (respectively) to the grounded extension \mathbb{G} of the AA framework $(\text{Args}, \rightsquigarrow)$ corresponding to the given case base CB , d and N . Explanations of AA outcomes (Section 2) exploit the argumentative re-interpretation afforded by AA-CBR utilizing not only the nearest cases, but also dialectical exchanges of relevant arguments (cf. e.g. [26]). In this section, we prove several properties of explanations in AA-CBR. Where indicated, these properties were stated in [11], but their proofs omitted there.

The following result will help us to characterize explanations of AA outcomes.

Theorem 3. $(\emptyset, d) \in \mathbb{G}$ iff there exists an admissible dispute tree \mathcal{T} for (\emptyset, d) .

Proof. By Theorem 3.2 in [14], there is an admissible dispute tree \mathcal{T} for (\emptyset, d) iff (\emptyset, d) is in some admissible extension. Every admissible extension is contained in some preferred extension [12] and, as $(\text{Args}, \rightsquigarrow)$ is well founded, \mathbb{G} is the only preferred extension. Thus, (\emptyset, d) is in some admissible extension iff $(\emptyset, d) \in \mathbb{G}$, and so the claim follows. \square

Thus, an explanation for the default outcome always exists:

Proposition 4 ([11, Proposition 3]). *If the AA outcome of the new case N is the default outcome d , then there is an explanation \mathcal{T} for why the AA outcome of N is d , which is moreover such that the defence set $\mathcal{D}(\mathcal{T})$ is admissible and $\mathcal{D}(\mathcal{T}) \subseteq \mathbb{G}$.*

Proof. Existence of explanations follows from Theorem 3. Further, Theorem 3.2, part (ii), in [14], says that if $\mathbf{a} \in \text{Args}$ belongs to an admissible set $A \subseteq \text{Args}$ of arguments, then there exists an admissible dispute tree \mathcal{T} for \mathbf{a} such that $\mathcal{D}(\mathcal{T}) \subseteq A$ and $\mathcal{D}(\mathcal{T})$ is admissible. Since \mathbb{G} is admissible and $(\emptyset, d) \in \mathbb{G}$, there is an admissible dispute tree \mathcal{T} for (\emptyset, d) with $\mathcal{D}(\mathcal{T})$ admissible and $\mathcal{D}(\mathcal{T}) \subseteq \mathbb{G}$. \square

An analogous result holds regarding explanations for the non-default outcome:

Proposition 5 ([11, Proposition 5]). *If the AA outcome of the new case N is \bar{d} , then there is an explanation \mathcal{T} for why the AA outcome of N is \bar{d} , and moreover $\mathcal{D}(\mathcal{T}) \not\subseteq \mathbb{G}$.*

Proof. Theorem 3.1 in [14] states that a dispute tree \mathcal{T} such that every 0 node in \mathcal{T} has a child, is necessarily admissible if it is *finite*. Since dispute trees in our setting are guaranteed to be finite, any dispute tree with all leaves labelled P would be admissible, yielding $(\emptyset, d) \in \mathbb{G}$ (by Theorem 3), contradicting the AA outcome of N being \bar{d} . Thus, some dispute tree \mathcal{T} for (\emptyset, d) will have all 0 unattacked in $(\text{Args}, \rightsquigarrow)$, and so be maximal, as required. Further, if $\mathcal{D}(\mathcal{T}) \subseteq \mathbb{G}$ then, by definition of dispute trees and grounded extensions, $(\emptyset, d) \in \mathbb{G}$, which is again a contradiction. Hence, $\mathcal{D}(\mathcal{T}) \not\subseteq \mathbb{G}$. \square

In Example 1, dispute trees $\mathcal{T}_p : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, E\}, +)]$ and $\mathcal{T}_p' : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, O\}, +)]$ are explanations for why the AA outcome of $\{S, E, O, G\}$ is $+$, with $\mathcal{D}(\mathcal{T}_p) = \{(\{S, E\}, +), (\emptyset, +)\} \subseteq \mathbb{G}$ and $\mathcal{D}(\mathcal{T}_p') = \{(\{S, O\}, +), (\emptyset, +)\} \subseteq \mathbb{G}$, both admissible. Each explanation serves Bob to legitimize why he is justified in asking the price he has in mind. Similarly, in Example 5 (where CB from Example 1 is augmented with $(\{S, E, O\}, -)$), the trees $\mathcal{T}_0 : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, E\}, +)] - [0 : (\{S, E, O\}, -)]$ and $\mathcal{T}_0' : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, O\}, +)] - [0 : (\{S, E, O\}, -)]$ are explanations for why the AA outcome of $\{S, E, O, G\}$ is $-$, with the same defence sets as \mathcal{T}_p and \mathcal{T}_p' (respectively), yet no longer contained in \mathbb{G} . Each explanation indicates that Bob should reconsider his price tag.

The next result says that every case that should be considered in explaining as well as determining the AA outcome is indeed considered.

Proposition 6. *For every nearest case (X, o) , there is an explanation \mathcal{T} (for why the AA outcome of N is either d or \bar{d}) s.t. for some (X', o) with $X' \subseteq X$ we find $(X', o) \in \mathcal{D}(\mathcal{T})$.*

Proof. If a nearest case (X, o) does not itself appear in any explanation, then some (X', o) with $X' \subseteq X$ must appear in some explanation \mathcal{T} . In any event, if either $o = d$ or $o = \bar{d}$, we find $(X', o) \in \mathcal{D}(\mathcal{T})$ by construction of \mathbb{G} and \mathcal{T} . \square

In Example 5, the unique nearest case $(\{S, E, O\}, -)$ to $N = \{S, E, O, G\}$ labels a node in both explanations \mathcal{T}_0 and \mathcal{T}_0' (as above) for why the AA outcome of N is $-$. Observe that $(\{S, E, O\}, -)$ is unattacked. This need not always happen: in Example 6, $(\{S, E, O\}, -)$ is still a unique nearest case to N , but this time attacked by $(\{S, E, O, M\}, +)$, which is in turn attacked by $(N, ?)$. In any event, $(\{S, E, O\}, -)$ labels a node in both possible explanations, namely $\mathcal{T}_1 : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, E\}, +)] - [0 : (\{S, E, O\}, -)] - [P : (\{S, E, O, M\}, +)] - [0 : (\{S, E, O, G\}, ?)]$ and $\mathcal{T}_2 : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, O\}, +)] - [0 : (\{S, E, O\}, -)] - [P : (\{S, E, O, M\}, +)] - [0 : (\{S, E, O, G\}, ?)]$, for why the AA outcome of N is $-$.

A nearest case need not itself appear in any explanation; instead, some of its ‘proper subsets’ will. For instance, suppose that in Example 5, instead of $(\{S, E, O\}, -)$, we have $(\{S, E, O\}, +)$, which is then a unique nearest case to N . The AA outcome of N is then $+$, for which $\mathcal{T}_p : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, E\}, +)]$ and $\mathcal{T}_p' : [P : (\emptyset, +)] - [0 : (\{S\}, -)] - [P : (\{S, O\}, +)]$ are explanations, with $\mathcal{D}(\mathcal{T}_p) = \{(\{S, E\}, +), (\emptyset, +)\}$ and $\mathcal{D}(\mathcal{T}_p') = \{(\{S, O\}, +), (\emptyset, +)\}$. Thus, $(\{S, E, O\}, +)$ does not label any node in either \mathcal{T}_p or \mathcal{T}_p' , but $(\{S, E\}, +)$ and $(\{S, O\}, +)$ do.

In general, every argument in an explanation has a reason to appear there:

Proposition 7. Every argument labelling a node in an explanation \mathcal{T} (for why the AA outcome of N is either d or \bar{d}), is either relevant or attacked by $(N, ?)$.

Proof. By definition of relevance, if $(X, o) \neq (\emptyset, d)$ labels a node in \mathcal{T} and is irrelevant, then $X \cap N = \emptyset$, so that $(N, ?) \rightsquigarrow (X, o)$, by definition of attack. \square

In Example 5, every argument labelling a node in any of the explanations \mathcal{T}_0 and \mathcal{T}'_0 is relevant. To see that irrelevant arguments can also appear in explanations, consider a single past case $(\{A\}, +)$, default outcome $-$ and a new case $\{B\}$. In the corresponding $(Args, \rightsquigarrow)$, we find $(\{B\}, ?) \rightsquigarrow (\{A\}, +) \rightsquigarrow (\emptyset, -)$ and $\mathbb{G} = \{(\{B\}, ?), (\emptyset, -)\}$, so that the AA outcome of $\{B\}$ is $-$, for which $\mathcal{T} : [P : (\emptyset, -)] - [O : (\{A\}, +)] - [P : (\{B\}, ?)]$ is an explanation. Here, $(\{A\}, +)$ is irrelevant. Observe further that there could be many more similar irrelevant cases $(\{A_1\}, +)$, $(\{A_2\}, +)$, \dots , whence there would be as many explanations, all of them containing an irrelevant case.

To avoid overpopulation of explanations with irrelevant arguments, we next propose a leaner version of explanations that contain only relevant arguments.

Definition 8. Let a **relevant dispute tree** be a dispute tree in construction of which only relevant arguments can label nodes. A **lean explanation** for why the AA outcome of N is d (resp., \bar{d}) is an admissible (resp., maximal) relevant dispute tree for (\emptyset, d) .

As for (standard) explanations, the *defence set* of a lean explanation \mathcal{T}^L , denoted by $\mathcal{D}(\mathcal{T}^L)$, is the set of all arguments labelling P nodes in \mathcal{T}^L .

The explanations discussed in Examples 1, 5 and 6 are lean, whereas the explanations discussed in the example before Definition 8 are not: the only lean explanation there is simply $\mathcal{T}^L : [P : (\emptyset, -)]$.

Note that a lean explanation for why the AA outcome of N is d (resp., \bar{d}) is a maximal subtree of an explanation for why the AA outcome of N is d (resp., \bar{d}) such that no parent node is labelled by an irrelevant argument. Plainly, lean explanations can be obtained from (standard) explanations by removing irrelevant nodes, as well as their children.

From Proposition 7 and Definition 8 it trivially follows that nothing is irrelevant in lean explanations:

Corollary 8. Every argument labelling a node in a lean explanation \mathcal{T}^L (for why the AA outcome of N is either d or \bar{d}) is relevant.

Simultaneously, lean explanations keep desirable properties in the following sense.

Corollary 9. If the AA outcome of N is d , then there is a lean explanation \mathcal{T}^L for why the AA outcome of N is d , such that $\mathcal{D}(\mathcal{T}^L) \cup \{(N, ?)\}$ is admissible and $\mathcal{D}(\mathcal{T}^L) \subseteq \mathbb{G}$.

If the AA outcome of N is \bar{d} , then there is a lean explanation \mathcal{T}^L for why the AA outcome of N is \bar{d} , such that $\mathcal{D}(\mathcal{T}^L) \not\subseteq \mathbb{G}$.

Proof. Follows from Propositions 4 and 5, in the first instance noticing that $(N, ?)$ need not have a relevant parent in an explanation (and hence $(N, ?) \notin \mathcal{D}(\mathcal{T}^L)$). \square

Utilizing the following definition, we see that lean explanations also impose a certain structure to the otherwise unstructured collection of relevant cases.

Definition 9. Let $(X, o_X), (Y, o_Y) \in Args$ be relevant. We say that (X, o_X) is **more relevant** than (Y, o_Y) if either $Y \subsetneq X$ or $(X, o_X) = (N, ?)$.

Proposition 10. Every argument labelling a node in a lean explanation (for why the AA outcome of N is either d or \bar{d}) is more relevant than the argument labelling its parent.

Proof. Follows from Definitions 2 (attack), 8 (lean explanations) and 9. \square

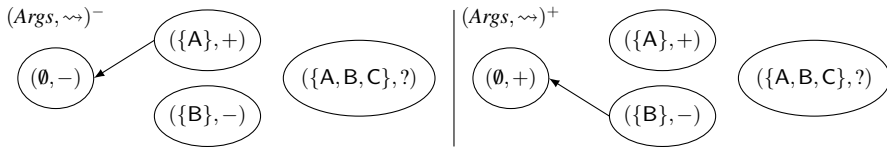
Apart from structuring past cases and providing dialogical justifications for why a particular outcome is assigned to a new case, explanations also yield hints on modifying the situation so as to achieve the desired outcome. For example, imagine that Alice wants to let a Small room with En-suite ($N = \{S, E\}$) for over £900, and past cases are $(\{E\}, -)$ (En-suite falls short) and $(\{S, E, H\}, +)$ (High-speed internet brings the rent over £900). The AA outcome of N is $-$ (as $(\{E\}, -)$ is a unique nearest case to N), with an explanation $[P: (\emptyset, +)] - [0: (\{E\}, -)] - [P: (\{S, E, H\}, +)] - [0: (\{S, E\}, ?)]$, from which Alice sees that installing High-speed internet would allow her to ask the price she wants.

This last illustration, together with the previously considered examples, hint at a feature of AA-CBR, namely that AA outcomes and (lean) explanations exhibit a certain asymmetry between the proponent and the opponent. This is in line with the asymmetry observed in the context of CBR, e.g. in [26].

5. Proof Standards for AA Outcome

In this section we show that the asymmetry described above is a manifestation of the *burden of proof* [18] falling onto the proponent, by introducing, for AA-CBR frameworks, a variant of a well-known *proof standard*. Consider the following example.

Example 10. Let $N = \{A, B, C\}$ and $CB = \{(\{A\}, +), (\{B\}, -)\}$. Consider the two default outcomes $-$ and $+$ in turn. Below are depicted the AA frameworks $(Args, \rightsquigarrow)^d$ corresponding to CB , $d \in \{-, +\}$ and N :



In $(Args, \rightsquigarrow)^-$, the AA outcome of N is $+$ (non-default), with a unique (lean) explanation $\mathcal{T}^- : [P: (\emptyset, -)] - [0: (\{A\}, +)]$. Likewise, In $(Args, \rightsquigarrow)^+$, the AA outcome is $-$, with a unique (lean) explanation $\mathcal{T}^+ : [P: (\emptyset, +)] - [0: (\{B\}, -)]$. Thus, no matter what the default outcome is, the burden of establishing as well as explaining the AA outcome is on the proponent's side.

In what follows, we formalize this feature of AA-CBR in terms of proof standards.

In our setting, following [20], a proof standard can be seen as a function taking a statement and an AA framework and returning an element of $\{\text{TRUE}, \text{FALSE}\}$. Then, a statement s is satisfied by a proof standard STD in $(Args, \rightsquigarrow)$ iff $\text{STD}(s, (Args, \rightsquigarrow)) = \text{TRUE}$. In the context of AA-CBR, the following statement is of interest:

s_d : “given a case base CB and a default outcome d , the outcome of the new case N is d ”.

We identify a proof standard that meets this statement:

Definition 11. The **Scintilla of Evidence** proof standard SE is defined as follows:

$SE(s_d, (Args, \rightsquigarrow)) = \text{TRUE}$ iff there exists an admissible dispute tree \mathcal{T} for (\emptyset, d) .

Intuitively, the SE proof standard amounts to the proponent P having a good line of defence (a tree of arguments and attacks) in a dialectical exchange of arguments for and against the default outcome. [7,18,20], to name a few, give proof standards with the same name. Our variant is in the same spirit in that a statement by a proponent P meets the standard “if it is supported by at least one *defensible* P argument”, where, in our variant, we interpret support as as a dispute tree, and defensible as admissible.

Directly from Theorem 3, we get that AA outcome meets the SE proof standard, in that accepting the default case equates with the satisfaction of the SE proof standard:

Theorem 11. $(\emptyset, d) \in \mathbb{G}$ iff $SE(s_d, (Args, \rightsquigarrow)) = \text{TRUE}$.

This result also indicates that no matter what the default outcome d is, the burden of proof to establish that the AA outcome of the new case N is d falls upon the proponent, in that the proponent needs to construct an admissible dispute tree that dialectically justifies the outcome. This is witnessed in examples we considered, particularly Example 10.

6. Related and Future Work

Argumentation has perhaps been most prominently applied to legal CBR. For example, [6] use AA frameworks to reason with particular types of animal cases by representing legal natural language arguments involved in cases as formal arguments, at the same time taking into account preference information over values promoted by arguments. [1] show how to represent the well known legal CBR systems HYPO [4], CATO [2] and IBP [8] in Abstract Dialectical Frameworks [7]. Another strand of research concerns argumentation schemes—patterns to create and/or classify arguments in order to decide how precedent cases determine the new case (e.g. [20,21,27]). There, proof standards can be employed to evaluate arguments based on argumentation schemes, as in e.g. [20].

In contrast, we are not focused on legal CBR, but rather on general CBR, as overviewed in [28]. In that setting, our work stands out in its aim to provide explanations as to why a particular outcome was obtained in solving CBR problems. To this end we exploit the dialectical aspect that AA supplies by way of dispute trees. Explaining AA outcome can be seen through a dialogical exchange of arguments (namely, past cases) between a proponent in favour of the default outcome, and an opponent against the default outcome. Explanations in AA-CBR relate to the notion of burden of proof from legal CBR, see e.g. [18,26]. However, legal CBR exhibits characteristics not necessarily applicable to CBR in general, or at least not to the type of problems we consider. For instance, in legal CBR, there is usually more granularity to factors [22]; certain hypothetical reasoning and/or background knowledge is involved [3]. Whether and how our approach can be applied to, for instance, legal CBR, is a line of future work: it would be interesting to look at other proof standards as well as burdens of production and persuasion [18]; relating AA-CBR to argumentation based on the discovery of association

rules, as in e.g. [31], and to argument and theory construction from legal cases, as in e.g. [10], would be interesting too.

In terms of general CBR, determining *why* an outcome is computed is deemed crucial, but is inherently hard to define formally [30]. A common form of explanation in CBR amounts to displaying the most similar past cases. In particular, *transparency*, in not trying to “hide conflicting evidence” [30, p. 134], is identified as desirable. This is fulfilled in AA-CBR, as the grounded extension contains all past cases nearest to the new case, be they of agreeing or diverging outcomes (cf. Proposition 2). However, merely displaying the nearest cases (especially with contrasting outcomes) is not always sufficient to explain the proposed outcome. To address this issue, *k*-nearest neighbour approaches produce only the most similar among the nearest neighbours. But then, “the transparency goal is no longer fulfilled [...] if $k > 1$ ” [30, p. 136]. In contrast, (lean) explanations in AA-CBR amount to (relevant) dispute trees, where not only the nearest cases, but also cases relevant to the AA outcome play a role.

In our setting, relevance of past cases is defined via their commonalities with the new case, in terms of factors shared. By contrast, [24] proposed *supporting/opposition* criteria based on counting the ratio (or probability) of how often a factor appears in a case with the outcome d/\bar{d} . We provide explanations without quantifying the appearance of factors, but we plan to investigate such a possibility in the future.

Several works define methods for determining explanations for the (non-)acceptability of arguments in argumentation, see e.g. [16,17,19,29]. These works use trees as the underlying mechanism for computing explanations, but not in a CBR setting. Study of formal relationships with these works is left for future work. Other work in argumentation, e.g. [9], investigates the usefulness of explanation in argumentation with users. Similar explorations for our approach are also left for the future.

Last but not least, computational complexity is an important aspect of explanations in CBR. The construction of the grounded extension \mathbb{G} of a given $(Args, \rightsquigarrow)$ is P-complete [15], so we conjecture that extracting explanations from \mathbb{G} results in a *low construction overhead* [30], as follows: if f is the (fixed) number $|\mathbb{F}|$ of factors, letting n to be the number $|Args|$ of arguments, to construct a (maximal or admissible) dispute tree for (\emptyset, d) we need to traverse the constructed graph of \mathbb{G} from (\emptyset, d) in depth at most f , in every layer exploring at most n^f arguments, so the process is polynomial in n with $O(n^{f^2})$. Precise analysis of this conjecture, as well as the complexity of construction of AA frameworks corresponding to case bases, is left for future work.

References

- [1] Al-Abdulkarim, L., Atkinson, K., Bench-Capon, T.J.M.: Abstract Dialectical Frameworks for Legal Reasoning. In: Hoekstra, R. (ed.) *Leg. Knowl. Inf. Syst.*, pp. 61–70. IOS Press (2014)
- [2] Aleven, V.: Teaching Case-based Argumentation Through a Model and Examples. Ph.D. thesis, University of Pittsburgh (1997)
- [3] Aleven, V.: Using Background Knowledge in Case-Based Legal Reasoning: A Computational Model and an Intelligent Learning Environment. *Artif. Intell.* 150, 183–237 (2003)
- [4] Ashley, K.D.: *Modeling Legal Argument - Reasoning With Cases and Hypotheticals*. MIT Press (1990)
- [5] Athakravi, D., Satoh, K., Law, M., Broda, K., Russo, A.: Automated Inference of Rules with Exception from Past Legal Cases Using ASP. In: Calimeri, F., Ianni, G., Truszczyński, M. (eds.) *LPNMR. Lecture Notes in Computer Science*, vol. 9345, pp. 83–96. Springer, Lexington (2015)

- [6] Bench-Capon, T.J.M., Modgil, S.: Case Law in Extended Argumentation Frameworks. In: ICAIL. pp. 118–127. ACM, Barcelona (2009)
- [7] Brewka, G., Woltran, S.: Abstract Dialectical Frameworks. In: Lin, F., Sattler, U., Truszczyński, M. (eds.) KR. AAAI Press, Toronto (2010)
- [8] Brüninghaus, S., Ashley, K.D.: Predicting Outcomes of Case-Based Legal Arguments. In: Zelezniok, J., Sartor, G. (eds.) ICAIL. pp. 233–242. ACM, Edinburgh (2003)
- [9] Cerutti, F., Tintarev, N., Oren, N.: Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation. In: Schaub, T., Friedrich, G., O’Sullivan, B. (eds.) ECAI. Frontiers in Artificial Intelligence and Applications, vol. 263, pp. 207–212. IOS Press, Prague (2014)
- [10] Chorley, A., Bench-Capon, T.J.M.: AGATHA: Using Heuristic Search to Automate the Construction of Case Law Theories. *Artif. Intell. Law* 13(1), 9–51 (2005)
- [11] Čyras, K., Satoh, K., Toni, F.: Abstract Argumentation for Case-Based Reasoning. In: Baral, C., Delgrande, J.P., Wolter, F. (eds.) KR. pp. 549–552. AAAI Press, Cape Town (2016)
- [12] Dung, P.M.: On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-person Games. *Artif. Intell.* 77, 321–357 (1995)
- [13] Dung, P.M., Kowalski, R., Toni, F.: Dialectic Proof Procedures for Assumption-Based, Admissible Argumentation. *Artif. Intell.* 170(2), 114–159 (2006)
- [14] Dung, P.M., Mancarella, P., Toni, F.: Computing Ideal Sceptical Argumentation. *Artif. Intell.* 171(10–15), 642–674 (2007)
- [15] Dunne, P.E., Hunter, A., McBurney, P., Parsons, S., Wooldridge, M.: Weighted Argument Systems: Basic Definitions, Algorithms, and Complexity Results. *Artif. Intell.* 175(2), 457–486 (2011)
- [16] Fan, X., Toni, F.: On Computing Explanations in Abstract Argumentation. In: Bonet, B., Koenig, S. (eds.) AAAI. pp. 1496–1502. AAAI Press, Austin, Texas (2015)
- [17] Fan, X., Toni, F.: On Explanations for Non-Acceptable Arguments. In: Black, E., Modgil, S., Oren, N. (eds.) TAFA. Lecture Notes in Computer Science, vol. 9524, pp. 112–127. Springer, Buenos Aires (2015)
- [18] Farley, A.M., Freeman, K.: Burden of Proof in Legal Argumentation. In: McCarty, T. (ed.) ICAIL. pp. 156–164. ACM, College Park (1995)
- [19] García, A.J., Chesñevar, C., Rotstein, N., Simari, G.R.: Formalizing Dialectical Explanation Support for Argument-Based Reasoning in Knowledge-Based Systems. *Expert Syst. Appl.* 40, 3233–3247 (2013)
- [20] Gordon, T., Prakken, H., Walton, D.: The Carneades Model of Argument and Burden of Proof. *Artif. Intell.* 171(10–15), 875–896 (2007)
- [21] Gordon, T., Walton, D.: Legal Reasoning with Argumentation Schemes. In: ICAIL. pp. 137–146. ACM, Barcelona (2009)
- [22] Horty, J., Bench-Capon, T.J.M.: A Factor-Based Definition of Precedential Constraint. *Artif. Intell. Law* 20(2), 181–214 (2012)
- [23] de Mántaras, R.L.: Case-Based Reasoning. In: Paliouras, G., Karkaletsis, V., Spyropoulos, C.D. (eds.) Mach. Learn. Its Appl. Adv. Lect., Lecture Notes in Computer Science, vol. 2049, pp. 127–145. Springer (2001)
- [24] McSherry, D.: Explaining the Pros and Cons of Conclusions in CBR. In: Funk, P., González-Calero, P.A. (eds.) ECCBR. Lecture Notes in Computer Science, vol. 3155, pp. 317–330. Springer, Madrid (2004)
- [25] Ontañón, S., Plaza, E.: An Argumentation-Based Framework for Deliberation in Multi-Agent Systems. In: Rahwan, I., Parsons, S., Reed, C. (eds.) ArgMAS. pp. 178–196. Springer, Honolulu (2007)
- [26] Prakken, H., Sartor, G.: Modelling Reasoning with Precedents in a Formal Dialogue Game. *Artif. Intell. Law* 6(2–4), 231–287 (1998)
- [27] Prakken, H., Wyner, A., Atkinson, K., Bench-Capon, T.J.M.: A Formalization of Argumentation Schemes for Legal Case-Based Reasoning in ASPIC+. *J. Log. Comput.* 25(5), 1141–1166 (2015)
- [28] Richter, M., Weber, R.: Case-Based Reasoning. Springer (2013)
- [29] Schulz, C., Toni, F.: Justifying Answer Sets Using Argumentation. *Theory Pract. Log. Program.* 16(1), 59–110 (2016)
- [30] Sørmo, F., Cassens, J., Aamodt, A.: Explanation in Case-Based Reasoning—Perspectives and Goals. *Artif. Intell. Rev.* 24(2), 109–143 (2005)
- [31] Wardeh, M., Bench-Capon, T.J.M., Coenen, F.: PADUA: A Protocol for Argumentation Dialogue Using Association Rules. *Artif. Intell. Law* 17(3), 183–215 (2009)