Greedy ABA Learning for Case-Based Reasoning

Emanuele De Angelis
IASI-CNR
Rome, Italy
emanuele.deangelis@iasi.cnr.it

Maurizio Proietti
IASI-CNR
Rome, Italy
maurizio.proietti@iasi.cnr.it

Francesca Toni
Imperial
London, UK
ft@ic.ac.uk

ABSTRACT

ABA Learning is a form of logic-based learning, producing, from examples and background knowledge, symbolic representations in the form of assumption-based argumentation (ABA) frameworks that naturally encode conflicts emerging from generalising the examples as well as their resolution. ABA Learning is based on the application of transformation rules to progressively refine an initial ABA framework (the background knowledge) guided by the examples, and is typically highly nondeterministic, with the search space underpinning the choice of applied transformation rules very large. In this paper we propose a novel 'greedy' variant of ABA Learning tailored to settings where the examples and background knowledge are drawn from labelled cases as in case-based reasoning. *Greedy ABA Learning* applies the transformation rules in a fully deterministic way. We prove that, when the casebase is 'coherent' (i.e., where all cases with the same features have the same label), Greedy ABA Learning corresponds exactly with AA-CBR, another form of logic-based learning for case-based reasoning. Finally, we show that Greedy ABA Learning generalises beyond coherent casebases to deal with conflicts.

KEYWORDS

Assumption-based Argumentation; Learning from Data; Intrinsic Explainability

ACM Reference Format:
Emanuele De Angelis, Maurizio Proietti, and Francesca Toni. 2025. Greedy ABA Learning for Case-Based Reasoning. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 9 pages.

1 INTRODUCTION

ABA Learning [9, 10, 21] is a form of logic-based learning, producing symbolic representations in the form of assumption-based argumentation (ABA) frameworks [2] from an initial ABA framework (the background knowledge) and (positive and negative) examples. The learnt ABA frameworks naturally encode conflicts emerging from generalising the examples, given the background knowledge, as well as their resolution. They encode this knowledge with the use of rules, made defeasible by the presence amongst their premises of assumptions which can be "attacked" by "arguments" for their contraries [2]. Several forms of ABA Learning have been proposed in the literature, differing in how they resolve conflicts emerging in the learnt ABA frameworks, whether sceptically [9, 23], by adopting the (unique) grounded extension semantics for ABA, or credulously [10], by choosing one (stable) extension amongst the several possible.

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

ABA Learning is based on the application of transformation rules to progressively derive general rules from an initial ABA framework (the background knowledge) guided by the examples. Typically, these transformation rules may be applied in many alternative ways, and thus the search space underpinning ABA Learning is very large. For instance, the 'rote learning' transformation rule can be applied to any positive example, the 'folding' transformation rule can be applied to any two ABA rules whose premises "match", and the 'assumption introduction' transformation rule can be applied to any ABA rule to pave the way to attack arguments drawn from it. A major issue for developing effective ABA Learning strategies is to control the resulting high nondeterminism.

In this paper, we propose *Greedy ABA Learning*, a novel variant of ABA Learning tailored to settings where the examples and background knowledge are drawn from (categorical) casebases, consisting of sets of cases, each characterised by a set of (binary) features and a (binary) label. In this setting, the background knowledge corresponds to simple ABA frameworks consisting of 'facts' (rules with empty premises).

Greedy ABA Learning starts by generalising the examples exhaustively (by applying the 'rote learning' and 'folding' transformation rules) and only afterwards looks for attacks against arguments in the resulting ABA frameworks (by applying the 'assumption introduction' transformation rule exhaustively).

Greedy ABA Learning can be deployed with 'coherent' casebases, namely such that there are no two cases with the same features but different labels, as well as with 'incoherent' casebases. We prove that, when the casebase is 'coherent', Greedy ABA Learning corresponds exactly with AA-CBR [3, 8, 16], another form of logic-based learning with casebases of the same kind we consider, but using abstract argumentation [11] as the underpinning symbolic formalism rather than ABA. We also show that Greedy ABA Learning generalises beyond coherent casebases to deal (credulously) with conflicts amongst cases.

In summary, we make the following contributions:

(1) we define a novel variant of ABA Learning, using specialised versions of existing transformation rules [21] so as to limit the search space of possible solutions;
(2) we prove that this variant corresponds to an existing form of AA-CBR [3] for coherent casebases; and
(3) we show that it can naturally deal with incoherent casebases.

2 RELATED WORK

Learning argumentation frameworks from data. Several approaches to integrate/reconcile argumentation and data-driven machine learning have been proposed (e.g., see the early survey in [5] and more recently [19]). Amongst these, a few understand supervised learning with data as argumentation, notably [1] understands concept learning as abstract argumentation with preferences.

We focus on the AA-CBR approach of [3, 8, 16], which uses, as a starting point, a casebase, with each case characterised by a set of binary features and one of two outcomes, and predicts outcomes for new cases based on the calculation of grounded extensions of an abstract argumentation framework corresponding to the learning problem. However, while [3, 8] restrict attention to *coherent* casebases (i.e., such that no two cases are equipped with the same features but a different outcome), we also consider the possibility of incoherent casebases. Differently from [16], that also drops the coherence restriction, we do so by adopting a credulous semantics for abstract argumentation.

Some other approaches are not restricted to cases with binary features only. These include the approach of [18] that, similarly to [3, 8], maps outcome prediction based on cases, onto abstract argumentation, but also accommodates the tendency of features to favour one side or another. Furthermore, [4] uses AA-CBR alongside various feature engineering methods to apply AA-CBR to images and text, and [17] uses AA-CBR alongside decision trees to binarise non-binary features for prediction with any set of cases. Furthermore, [6] considers two sets of binary features per case, one of which represents dynamic information. We leave to future work accommodating these various extensions/variants of AA-CBR to provide corresponding forms of ABA learning.

Learning ABA frameworks. ABALearn [21], and its implementation in [23], is a precursor in using transformation rules for learning ABA frameworks. However, they give a nondeterministic strategy that focuses on cautious (a.k.a. sceptical) reasoning under the grounded extension semantics. Cautious reasoning under stable extension semantics is considered in [9], where the authors introduce a learning strategy, called *ASP-ABALearn*, implemented using Answer Set Programming (ASP) [13]. A recent extension of this strategy, called *ASP-ABALearn*$_B$ [10], considers brave (a.k.a. credulous) reasoning under stable extension semantics. Here we adopt a deterministic strategy for applying transformation rules, and consider reasoning under the grounded and stable extension semantics. Moreover, we also restrict the transformation strategy for its application to both coherent and incoherent casebases.

Other logic-based learning. ABA Learning is a form of logic-based learning. Other approaches to logic-based learning under stable model (i.e., stable extension) semantics include ILASP [15] and FOLD-RM [25]. However, differently from ILASP, ABA Learning features the ability of automatically synthesising brand new predicates (i.e., assumptions and their contraries), and FOLD-RM [25] can only learn stratified normal logic programs. We will present in Section 6 a variant of Greedy ABA Learning that can learn non-stratified ABA frameworks in the sense that cycles through contraries are allowed.

3 BACKGROUND

Abstract Argumentation (AA). An *AA framework (AAF)* [11] is a pair $(Args, \rightsquigarrow)$, where *Args* is a set of *arguments* and $\rightsquigarrow \subseteq Args \times Args$ is a (binary, directed) relation of *attack* between arguments. For $(\alpha, \beta) \in \rightsquigarrow$, we also write $\alpha \rightsquigarrow \beta$. Also, $\Delta \subseteq Args$ *defends* $\Gamma \subseteq Args$ iff for every $\alpha \rightsquigarrow \beta$ with $\beta \in \Gamma$, there exists $\gamma \rightsquigarrow \alpha$ with $\gamma \in \Delta$.

The semantics of an AAF is defined in terms of various notions of extensions [11], including the following. An *extension* $\Delta \subseteq Args$ is *conflict-free* iff there exists no $\alpha \rightsquigarrow \beta$ for $\alpha, \beta \in \Delta$; *admissible* iff it is conflict-free and it defends itself; *complete* iff it is admissible and for every $\alpha \in Args$, if Δ defends $\{\alpha\}$, then $\alpha \in \Delta$; *grounded* iff it is $\subseteq$-minimally complete; *stable* iff it is conflict-free and for every $\alpha \in Args \setminus \Delta$ there exists $\beta \rightsquigarrow \alpha$ for $\beta \in \Delta$.

Assumption-based Argumentation (ABA). An *ABA framework (ABAF)* [2, 7] is a tuple $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$, where $(\mathcal{L}, \mathcal{R})$ is a *deductive system*, consisting of a *language* $\mathcal{L}$ and a set $\mathcal{R}$ of rules over $\mathcal{L}$, $\mathcal{A} \subseteq \mathcal{L}$ is a non-empty set of *assumptions*, and $\overline{} : \mathcal{A} \rightarrow \mathcal{L}$ is a *contrary* mapping. Rules $r \in \mathcal{R}$ are of the form $s_0 \leftarrow s_1, \ldots, s_n$ with $n \geq 0$ and $s_i \in \mathcal{L}$ for all $i \in \{0, \ldots, n\}$: s_0 is the *head* of r, denoted by $head(r)$, and $\{s_1, \ldots, s_n\}$ is the (possibly empty) body of r, denoted by $body(r)$; if $body(r)$ is empty, then we may write r as $head(r) \leftarrow$ and call it a *fact*. We restrict attention to *flat* ABA frameworks, where $head(r) \notin \mathcal{A}$ for all $r \in \mathcal{R}$. Furthermore, as in [21], we restrict $\mathcal{L}$ to be a set of atoms, with each sentence in $\mathcal{L}$ of the form $p(t)$ for t a constant.[1]

A flat ABAF $F = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$ can be understood as an AAF $(Args^F, \rightsquigarrow^F)$ as follows. Let an *argument* with claim $s \in \mathcal{L}$, supported by $\mathsf{A} \subseteq \mathcal{A}$ and $R \subseteq \mathcal{R}$ (denoted $\mathsf{A} \vdash_R s$ or simply $\mathsf{A} \vdash s$) be a finite tree with nodes labelled by sentences in $\mathcal{L}$ or by *true*, such that the root is labelled by s, leaves are labelled by assumptions in Δ or by *true*, and for each non-leaf node n there is exactly one rule $r \in R$ such that n is labelled with $head(r)$, the number of children of n is $|body(r)|$ and every child of n is labelled with a distinct sentence in $body(r)$ or, if $body(r)$ is empty, by *true*. Then $Args^F = \{\mathsf{A} \vdash s \mid s \in \mathcal{L}\}$ and $\mathsf{A} \vdash s_1 \rightsquigarrow^F \mathsf{B} \vdash s_2$ iff $s_1 = \overline{\beta}$ for some $\beta \in \mathsf{B}$. Thus, the semantics of an ABAF F can be defined in terms of the semantics of extensions of the AAF $(Args^F, \rightsquigarrow^F)$ [24], that is, Δ is a grounded/stable extension of $F = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$ iff Δ is a grounded/stable extension of $(Args^F, \rightsquigarrow^F)$. It can be shown that every ABAF admits a unique grounded extension, denoted $\mathbb{G}(F)$. Furthermore, let an ABAF F be *stratified* iff $(Args^F, \rightsquigarrow^F)$ is acyclic. Then, a stratified ABAF F admits a unique stable extension, which coincides with $\mathbb{G}(F)$ [11].

We will use the following notion. Given $s \in \mathcal{L}$ and extension Δ of F, s is *covered in* Δ iff $(\mathsf{A} \vdash s) \in \Delta$ for some $\mathsf{A} \subseteq \mathcal{A}$. In the case where we focus on the (unique) grounded extension $\mathbb{G}(F)$ of ABAF F, we also say that s is *covered by* F, denoted $F \models s$, iff s is covered in $\mathbb{G}(F)$. Note that, if s is not covered in an extension Δ of F then, for every argument $(\mathsf{A} \vdash s)$ with $\mathsf{A} \subseteq \mathcal{A}$, $(\mathsf{A} \vdash s)$ may be attacked by Δ or not, but if F is stratified and Δ is its grounded extension $\mathbb{G}(F)$, then each such argument $(\mathsf{A} \vdash s)$ is necessarily attacked by Δ. Given argument $\alpha = \mathsf{A} \vdash_R s$ in an extension of some F, we refer to the single rule $\rho \in R$ such that the sentences in ρ's body label s' children in α as the *top rule* of the argument.

As in [21], we assume that ABAFs are given via *schemata*, using variables to represent compactly all instances over some *universe* (of constants), as in the following illustration.

[1] For simplicity, we focus on unary predicates, but, in general, our approach is applicable to predicates of any arity.

Example 1. *Let X range over universe $\{1,2\}$. Then,* $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$ *with* $\mathcal{L} = \{p(X), a(X), b(X)\}$, $\mathcal{R} = \{p(X) \leftarrow a(X)\}$,
$\mathcal{A} = \{a(X), b(X)\}$, $\overline{a(X)} = b(X)$, $\overline{b(X)} = p(X)$
represents the ABA framework
$\langle \{p(1), p(2), a(1), a(2), b(1), b(2)\},$
$\{p(1) \leftarrow a(1), \quad p(2) \leftarrow a(2)\},$
$\{a(1), a(2), b(1), b(2)\}, \overline{} \rangle$
with $\overline{a(1)} = b(1)$, $\overline{a(2)} = b(2)$, $\overline{b(1)} = p(1)$, $\overline{b(2)} = p(2)$.

Again as in [21], we also assume that $\mathcal{L}$ always contains all equalities between elements of the universe, $\mathcal{R}$ includes all *equality rules* $a = a \leftarrow$, where a belongs to the universe, and all non-equality rules in $\mathcal{R}$ are *normalised*, i.e., they are written as:

$$p_0(X_0) \leftarrow Eqs, p_1(X_1), \ldots, p_n(X_n)$$

where $p_i(X_i)$, for $0 \leq i \leq n$, is an atom (whose ground instances over the universe are) in $\mathcal{L}$ and Eqs is a (possibly empty) sequence of equalities whose variables occur in the tuples $X_0, X_1, \ldots, X_n$.[2] Further, we assume that the body of a normalised rule can be freely rewritten by using the standard axioms of equality, e.g., $Y_1 = a, Y_2 = a$ can be rewritten as $Y_1 = Y_2, Y_2 = a$. Finally, we use the notation $vars(Z)$, for Z any sequence of atoms, to refer to the set of all variables occurring in Z.

4 LEARNING PROBLEMS AND SOLUTIONS

In this paper we focus on the following learning problem. Let $\mathbb{F}$ be a set of (binary) features, and let $\mathcal{P}(\mathbb{F})$ be the powerset of $\mathbb{F}$. Let $O = \{\delta, \overline{\delta}\}$ be the set of possible outcomes, with δ the *default outcome*. Let $D \subseteq \mathcal{P}(\mathbb{F}) \times O$ be a finite *casebase* of labelled examples, each of the form (S, o_S) for $S \subseteq \mathbb{F}$ and $o_S \in O$, where D is said to be *coherent* iff for $(S, o_S), (T, o_T) \in D$, if $S = T$ then $o_S = o_T$.[3] Let $(N, ?)$, for $N \in \mathcal{P}(\mathbb{F})$, be a *new case*. Then, we strive towards predicting an outcome $o_N \in O$ for N by means of a classifier generalising the information held in D.

AA-CBR [3, 8] provides one solution for this learning problem, as reviewed next. For D coherent, let $AAF(D, \delta) = (Args, \rightsquigarrow)$ be the AAF obtained as follows:

- $Args = D \cup \{(\emptyset, \delta)\}$, where $(\emptyset, \delta)$ is the *default argument*;
- for $(S, o_S), (T, o_T) \in Args$, $(S, o_S) \rightsquigarrow (T, o_T)$ iff
  (1) $o_S \neq o_T$,
  (2) $T \subset S$, and
  (3) $\nexists (U, o_S) \in D$ such that $T \subset U \subset S$.

Given $AAF(D, \delta) = (Args, \rightsquigarrow)$, for a new case $(N, ?)$, let us define $AAF(D, \delta, N) = (Args_N, \rightsquigarrow_N)$ as follows:

- $Args_N = Args \cup \{(N, ?)\}$;
- $\rightsquigarrow_N = \rightsquigarrow \cup \{((N, ?), (T, o_T)) \mid T \nsubseteq N\}$.

Let $\mathbb{G}$ be the grounded extension of $AAF(D, \delta, N)$. Then the outcome o_N for N is

$$\text{AA-CBR}(D, \delta, N) = \begin{cases} \delta & \text{if } (\emptyset, \delta) \in \mathbb{G}, \\ \overline{\delta} & \text{otherwise.} \end{cases}$$

Note that $AAF(D, \delta, N)$ is acyclic, and, thus, the grounded extension coincides with the unique stable extension of $AAF(D, \delta, N)$.

[2] In this paper, all non-equality predicates are unary, but in general they can have any arity.
[3] Unless specified otherwise (see Section 6), we will focus on coherent casebases.

Example 2. *(Adapted from [3]) Let us consider* $\mathbb{F} = \{a, b, c, d\}$ *and* $D = \{(\{a\}, \overline{\delta}), (\{b\}, \overline{\delta}), (\{a, c\}, \delta), (\{b, d\}, \overline{\delta})\}$. *$D$ is coherent. For the new case* $(N, ?) = (\{a, d\}, ?)$, $AAF(D, \delta, N)$ *is depicted in Figure 1, its grounded extension is* $\{(\{a, d\}, ?), (\{a\}, \overline{\delta})\}$, *and thus the outcome* $AA\text{-}CBR(D, \delta, N)$ *is* $\overline{\delta}$.

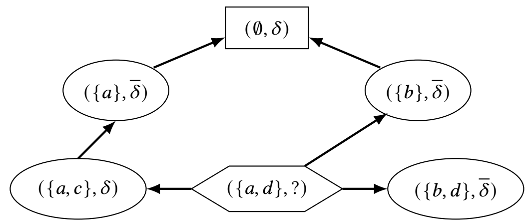

Figure 1: AAF(D, δ, N) for Example 2 as a graph (adapted from [3]), with the default argument as a rectangle, the new case as an argument shown as a hexagon, all other arguments (which are cases in D) as ovals and attacks as edges.

The outcome AA-CBR(D, δ, N) can be determined with an ABA framework, by adapting the Rule Extractor of [3] as follows.

Definition 1. *Let* $D \subseteq \mathcal{P}(\mathbb{F}) \times \{\delta, \overline{\delta}\}$ *be a coherent casebase and* $(N, ?)$ *be a new case. Let* $AAF(D, \delta) = (Args, \rightsquigarrow)$ *and* $AAF(D, \delta, N) = (Args_N, \rightsquigarrow_N)$. *Assume a naming function, name, that assigns a unique constant to every argument in Args, such that* $name((\emptyset, \delta)) = c_\delta$; *let* $C = \{name(\alpha) \mid \alpha \in Args_N\}$. *Then, the* ABA framework $F_{(D,\delta,N)}$ associated with $AAF(D, \delta, N)$ *is* $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$ *with*

- $\mathcal{L} = \{f(c) \mid f \in \mathbb{F}, c \in C\} \cup \{asm_c(X), ctr\_asm_c(X) \mid c \in C \setminus \{name((N, ?)\}, X \in C\} \cup \{default(X) \mid X \in C\}$;
- $\mathcal{R} = \mathcal{R}_\delta \cup \mathcal{R}_\chi \cup \mathcal{R}_\phi$ *where*
  - $\mathcal{R}_\delta = \{default(X) \leftarrow asm_{c_\delta}(X) \mid X \in C\}$
  - $\mathcal{R}_\chi = \{ctr\_asm_c(X) \leftarrow f_1(X), \ldots, f_n(X), asm_{c'}(X) \mid c = name(\gamma), (\beta, \gamma) \in \rightsquigarrow, c' = name(\beta), \beta = (\{f_1, \ldots, f_n\}, o)\}$;
  - $\mathcal{R}_\phi = \{f_1(c) \leftarrow, \ldots, f_n(c) \leftarrow \mid \gamma = (\{f_1, \ldots, f_n\}, o) \in Args_N, name(\gamma) = c\}$;
- $\mathcal{A} = \{asm_c(X) \mid c \in C \setminus \{c_N\}, X \in C\}$, *where* $c_N = name((N, ?))$;
- *for each* $asm_c(X) \in \mathcal{A}$, $\overline{asm_c(X)} = ctr\_asm_c(X)$.

Example 3. *Consider the learning problem in Example 2 and name such that*
$name((\{a\}, \overline{\delta})) = 1$, $name((\{b\}, \overline{\delta})) = 2$, $name((\{a, c\}, \delta)) = 3$,
$name((\{b, d\}, \overline{\delta})) = 4$, $name((\{a, d\}, ?)) = 5$.
Then, $F_{(D,\delta,N)}$ *has* $\mathcal{R} = \mathcal{R}_\delta \cup \mathcal{R}_\chi \cup \mathcal{R}_\phi$, *with* $\mathcal{R}_\phi$ *consisting of the following rules (in normalised form):*

$a(X) \leftarrow X = 1 \quad b(X) \leftarrow X = 2$
$a(X) \leftarrow X = 3 \quad c(X) \leftarrow X = 3$
$b(X) \leftarrow X = 4 \quad d(X) \leftarrow X = 4$
$a(X) \leftarrow X = 5 \quad d(X) \leftarrow X = 5$

and $\mathcal{R}_\delta \cup \mathcal{R}_\chi$ *consisting of the following rules:*[4]

[4] Note that here and everywhere, for simplicity, we omit to include assumptions in the body or rules when there is no rule with their contrary in the head. Indeed, those assumptions cannot be attacked and can be ignored. If given in full, in this example, the second and third rules in $\mathcal{R}_\chi$ would be $ctr\_asm_{c_\delta}(X) \leftarrow b(X), asm_2(X)$ and $ctr\_asm_1(X) \leftarrow a(X), c(X), asm_3(X)$.

$default(X) \leftarrow asm_{c_\delta}(X)$
$ctr\_asm_{c_\delta}(X) \leftarrow a(X), asm_1(X)$
$ctr\_asm_{c_\delta}(X) \leftarrow b(X)$
$ctr\_asm_1(X) \leftarrow a(X), c(X)$.

Given that $AAF(D, \delta, N)$ is acyclic, we get the following property.

Lemma 1. *$F_{(D,\delta,N)}$ is stratified.*

It is easy to see that the outcome obtained by AA-CBR can be equivalently determined by ascertaining whether or not the claim $default(c_N)$ is covered in $F_{(D,\delta,N)}$, by construction thereof.

Lemma 2. *AA-CBR$(D, \delta, N) = \delta$ iff $F_{(D,\delta,N)} \models default(c_N)$.*

Thus, prediction of outcomes for new cases can be achieved in ABA by covering $default(C_N)$ (in the grounded extension of $F_{(D,\delta,N)}$). For illustration, in Example 3, it is easy to see that $F_{(D,\delta,N)} \not\models default(5)$, as the argument $\{asm_1(5)\} \vdash ctr\_asm_{c_\delta}(5)$ cannot be attacked. Indeed, as we have seen in Example 2, the outcome for this new case is $\overline{\delta}$.

In addition, $F_{(D,\delta,N)}$ can be used to determine coverage of cases in D, as well as the default argument, as follows.

Lemma 3. *For any $(S, o) \in Args$, $F_{(D,\delta,N)} \models default(name((S, o))$ iff $o = \delta$.*

Thus, we can understand c_δ and cases in D with outcome δ as *positive examples* and cases in D with outcome $\overline{\delta}$ as negative examples of the *default* concept in the following ABA Learning problem, instantiating the general definition of [21]:

Definition 2. *The* ABA Learning problem associated with (D, δ, N) *is* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$, *where*

- *the* background ABA framework F_B *is* $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{\ } \rangle$ *with*
  - $\mathcal{L} = \{f(c) \mid f \in \mathbb{F}, c \in C\} \cup \{default(X) \mid X \in C\} \cup \{\alpha_0, ctr\_\alpha_0\}$;
  - $\mathcal{R} = \mathcal{R}_\phi$ *as in Definition 1;*
  - $\mathcal{A} = \{\alpha_0\}$;
  - $\overline{\alpha_0} = ctr\_\alpha_0$.
- *the* positive examples *are* $\mathcal{E}^+ = \{default(c) \mid name((S, \delta)) = c, (S, \delta) \in D\} \cup \{default(c_\delta)\}$;
- *the* negative examples *are* $\mathcal{E}^- = \{default(c) \mid name((S, \overline{\delta})) = c, (S, \overline{\delta}) \in D\}$.

Here, α_0 is a *bogus assumption*, as required in [24] to guarantee that the ABA framework includes at least one assumption, as in this simple setting where the rules amount solely to facts characterising the examples, no "real" assumption can be naturally identified.

Example 4. *Consider again the learning problem of Example 2 and the name function in Example 3. The associated ABA Learning problem is $(F_B, \mathcal{E}^+, \mathcal{E}^-)$, where F_B has $\mathcal{R} = \mathcal{R}_\phi$ as in Example 3, and the positive and negative examples are*
$\mathcal{E}^+ = \{default(c_\delta), default(3)\}$
$\mathcal{E}^- = \{default(1), default(2), default(4)\}$.

Note that, in Definition 2, $\mathcal{E}^+ \cap \mathcal{E}^- = \emptyset$ by construction. If we start from a coherent casebase, we can further pull apart $\mathcal{E}^+$ and $\mathcal{E}^-$ as follows:

Definition 3. *Let $\mathcal{R}$ be a set of rules of the form $p(X) \leftarrow X = c$ for a unary predicate p and a constant $c \in C$. Two constants $c_1, c_2 \in C$ are* discernible *in $\mathcal{R}$ if there exists a predicate p such that either $(p(X) \leftarrow X = c_1) \in \mathcal{R}$ and $(p(X) \leftarrow X = c_2) \notin \mathcal{R}$ or $(p(X) \leftarrow X = c_1) \notin \mathcal{R}$ and $(p(X) \leftarrow X = c_2) \in \mathcal{R}$.*

Proposition 1. *Let $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ be the ABA Learning problem associated with (D, δ, N), with $F_B = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{\ } \rangle$. The casebase D is coherent iff for all $c_1, c_2 \in C$, if $default(c_1) \in \mathcal{E}^+$ and $default(c_2) \in \mathcal{E}^-$, then c_1 and c_2 are discernible in $\mathcal{R}$.*

In the same way that, for the original learning problem, we strive towards predicting an outcome for the new case N by generalising the information in D (in an AAF), here we strive towards determining coverage of $default(c_N)$ from an ABA framework learnt from the examples so that all positive ones are covered and none of the negative ones are covered therein, as formally defined next:

Definition 4. *Given the ABA Learning problem $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ associated with (D, δ, N), with $F_B = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{\ } \rangle$, the* goal of ABA Learning *is to construct $F' = \langle \mathcal{L}', \mathcal{R}', \mathcal{A}', \overline{\ }' \rangle$, called a* solution, *such that $\mathcal{L} \subseteq \mathcal{L}'$, $\mathcal{R} \subseteq \mathcal{R}'$, $\mathcal{A} \subseteq \mathcal{A}'$, for all $\alpha \in \mathcal{A}$:[5] $\overline{\alpha}' = \overline{\alpha}$, and the following two conditions hold:*

- (Completeness) *for all $e \in \mathcal{E}^+$, $F' \models e$;*
- (Consistency) *for all $e \in \mathcal{E}^-$, $F' \not\models e$.*

F' is an intensional *solution when $\mathcal{R}' \setminus \mathcal{R}$ is made out of rule schemata constructed by using predicate symbols and variables only.*

Intuitively, intensionality captures a notion of generality for the learnt rules [10]. Note that $F'_{(D,\delta,N)}$ amounting to $F_{(D,\delta,N)}$ in Example 3 with additionally the bogus assumption α_0 and its contrary, is an intensional solution to the ABA Learning problem associated with the learning problem of Example 2. In the remainder, we will define a form of ABA Learning for generating solutions to the ABA Learning problem associated with (D, δ, N) in such a way that they correspond to the solutions determined by AA-CBR.

Note that in [10] solutions to the *brave ABA Learning problem* are defined by replacing the two conditions in Definition 4 with the following three conditions:

- (*Existence*) F' admits at least one stable extension Δ;
- (*Completeness*) for all $e \in \mathcal{E}^+$, e is covered in Δ;
- (*Consistency*) for all $e \in \mathcal{E}^-$, e is not covered in Δ.

In [10] F' is called a *brave* solution. Recall that for every stratified (flat) ABA framework the grounded extension is guaranteed to exist and to coincide with its (unique) stable extension. Thus, for stratified ABA frameworks, our notion of ABA Learning problem in Definition 4 is equivalent to the definition of brave ABA Learning problem in [10].

5 GREEDY ABA LEARNING FROM COHERENT CASEBASES

To achieve the goal of ABA Learning, several *strategies* and implementations thereof have been defined [9, 10, 21, 23], all combining in different ways *transformation rules*. In our novel strategy, we use the ones defined below, adapted from [21], all turning a given ABAF $\langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{\ } \rangle$ into a new ABAF $\langle \mathcal{L}', \mathcal{R}', \mathcal{A}', \overline{\ }' \rangle$[6], as follows: [7]

[5] In the specific setting of Definition 2, necessarily $\alpha = \alpha_0$.
[6] When its new components are the same as the old ones we omit to indicate them.
[7] For ABA rules: (1) H, K denote heads, (2) Eq denotes an equality, (3) B (possibly with subscripts) denotes sequences of atoms. Sequences of atoms can be freely reordered to enable the application of a transformation rule.

- *Rote Learning (RL).* For $p(t) \in \mathcal{L}$, $\mathcal{R}' = \mathcal{R} \cup \{p(X) \leftarrow X{=}t\}$.
- *Folding (Fld).* For $\rho_1, \rho_2 \in \mathcal{R}$, respectively of the form $H \leftarrow Eq, B$ and $K \leftarrow Eq$, $\mathcal{R}' = (\mathcal{R} \setminus \{\rho_1\}) \cup \{H \leftarrow Eq, K, B\}$.[8]
- *Assumption Introduction (AI).* For $\rho_1 \in \mathcal{R}$ of the form $H \leftarrow B$, let ρ_2 be $H \leftarrow B, \alpha(X)$, where X is the tuple of variables from $vars(H) \cup vars(B)$ and $\alpha(X) \notin \mathcal{L} \setminus \mathcal{A}$[9]. Then, $\mathcal{R}' = (\mathcal{R} \setminus \{\rho_1\}) \cup \{\rho_2\}$, $\mathcal{A}' = \mathcal{A} \cup \{\alpha(X)\}$, $\overline{\alpha(X)}' = ctr\_\alpha(X)$ for some $ctr\_\alpha(X) \notin \mathcal{A}'$, and $\overline{\beta}' = \overline{\beta}$ for all $\beta \in \mathcal{A}$.
- *Equality Removal (ER).* For $\rho_1 \in \mathcal{R}$ of the form $H \leftarrow Eq, B$, $\mathcal{R}' = (\mathcal{R} \setminus \{\rho_1\}) \cup \{H \leftarrow B\}$.
- *Subsumption (Su).* Suppose that $\mathcal{R}$ contains rules
  $\rho_1 : H \leftarrow B_1$ and $\rho_2 : H \leftarrow B_1, B_2$
  Then, ρ_2 is said to be *subsumed* by ρ_1 and is deleted from $\mathcal{R}$, and hence $\mathcal{R}' = \mathcal{R} \setminus \{\rho_2\}$.

Definition 5. *Given an ABAF $F = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$, an extension Δ for F, a set E of claims in $\mathcal{L}$, and a rule $(h(X) \leftarrow B) \in \mathcal{R}$, we define $Cov(\Delta, E, h(X) \leftarrow B) = \{h(t) \in E \mid$ there is an argument $\gamma \in \Delta$ with claim $h(t)$ and top rule $h(X) \leftarrow B$ such that γ is not attacked in $\Delta\}$.*

Now we present the notion of an ABA Learning derivation constructed by a sequence of applications of the transformation rules.

Definition 6 (ABA Learning from Coherent Casebases). *Given an ABA Learning problem $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ associated with the triple (D, δ, N), an* ABA Learning derivation *is a sequence*

$(F_0, \mathcal{E}_0^+, \mathcal{E}_0^-, r_0) \Rightarrow (F_1, \mathcal{E}_1^+, \mathcal{E}_1^-, r_1) \Rightarrow \ldots \Rightarrow (F_n, \mathcal{E}_n^+, \mathcal{E}_n^-, r_n) \ldots$

where, for $i \geq 0$, F_i is an ABAF and $r_i \in \{RL, Fld, ER, Su, AI\}$, such that, $\mathcal{E}_0^+ = \mathcal{E}^+$, $\mathcal{E}_0^- = \mathcal{E}^-$, r_0 is the RL label and, for $i > 0$, $(F_i, \mathcal{E}_i^+, \mathcal{E}_i^-, r_i)$ is obtained from $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, r_{i-1})$ by applying the transformation rule r_{i-1} as specified at the five points below[10]. *$\mathcal{R}_i$ and $\mathcal{A}_i$ denote, respectively, the set of rules and assumptions in F_i:*

(*RL*) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, RL) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, Fld)$, *where F_i is obtained by (repeatedly) applying RL for all $h(t) \in \mathcal{E}_{i-1}^+$;*

(*Fld*) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, Fld) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, ER)$, *where F_i is obtained by (repeatedly) applying Fld to all $h(X) \leftarrow X = c, B \in \mathcal{R}_{i-1}$ and $p(X) \leftarrow X = c \in \mathcal{R}_0$;*

(*ER*) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, ER) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, Su)$, *where F_i is obtained by (repeatedly) applying ER to all rules in $\mathcal{R}_{i-1} \setminus \mathcal{R}_0$;*

(*Su*) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, Su) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, AI)$, *where F_i is obtained by (repeatedly) applying Su and deleting from $\mathcal{R}_{i-1}$ every rule that is subsumed by another rule in $\mathcal{R}_{i-1}$;*

(*AI*) *there are three cases:*

(AI.1) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, AI) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, AI)$, *where F_i is obtained by selecting $\rho : h(X) \leftarrow B$ in $\mathcal{R}_{i-1}$ such that there exists $(k(X) \leftarrow B, \alpha(X)) \in \mathcal{R}_{i-1}$ and, by applying AI to ρ, getting $\mathcal{R}_i = (\mathcal{R}_{i-1} \setminus \{\rho\}) \cup \{h(X) \leftarrow B, \alpha(X)\}$;*

(AI.2) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, AI) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, AI)$, *where F_i is obtained by (1) selecting $\rho : h(X) \leftarrow B$ in $\mathcal{R}_{i-1}$ such that (1.a) there exists no rule in $\mathcal{R}_{i-1}$ with body $B, \alpha(X)$, with $\alpha(X) \in \mathcal{A}_{i-1}$ and (1.b) $Cov(\mathbb{G}(F_{i-1}), \mathcal{E}_{i-1}^-, \rho) \neq \emptyset$, and then (2) applying AI so that: (2.a) $\mathcal{A}_i = \mathcal{A}_{i-1} \cup \{\alpha(X)\}$, where $\alpha(X) \notin \mathcal{A}_{i-1}$ is a new assumption with contrary $\overline{\alpha(X)}$, and (2.b) $\mathcal{R}_i = (\mathcal{R}_{i-1} \setminus \{\rho\}) \cup \{h(X) \leftarrow B, \alpha(X)\}$;*

(AI.3) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, AI) \Rightarrow (F_{i-1}, \mathcal{E}_i^+, \mathcal{E}_i^-, RL)$, *where every $\rho \in \mathcal{R}_{i-1}$ with $Cov(\mathbb{G}(F_{i-1}), \mathcal{E}_{i-1}^-, \rho) \neq \emptyset$ has an assumption in its body, and $\mathcal{E}_i^+, \mathcal{E}_i^-$ are obtained as follows: (1) $\mathcal{E}_i^+ = \{\overline{\beta(t)} \mid h(t) \in Cov(\mathbb{G}(F_{i-1}), \mathcal{E}_{i-1}^-, \rho')$, for some $\rho' : h(X) \leftarrow B, \beta(X) \in \mathcal{R}_{i-1}\}$, and (2) $\mathcal{E}_i^- = \{\overline{\beta(t)} \mid h(t) \in Cov(\mathbb{G}(F_{i-1}), \mathcal{E}_{i-1}^+, \rho')$, for some $\rho' : h(X) \leftarrow B, \beta(X) \in \mathcal{R}_{i-1}\}$.*

We say that ABA Learning terminates *for the learning problem $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ and its output is an ABA framework F' if there exists a derivation $(F_B, \mathcal{E}^+, \mathcal{E}^-, RL) \Rightarrow \ldots \Rightarrow (F', \emptyset, \emptyset, RL)$. F' is denoted $ABAL(D, \delta, N)$.*

There is some nondeterminism in the construction of an ABA Learning derivation. In particular, the order of application of cases (*AI*.1) and (*AI*.2) is not fixed and, within each case, the selection of a rule ρ for the application of *AI* is arbitrary. However, it can be easily seen that the output of an ABA Learning derivation is independent of the specific sequence of these derivation steps, as stated below.

Proposition 2. *Let F' and F'' be the outputs of two ABA Learning derivations for the same learning problem $(F_B, \mathcal{E}^+, \mathcal{E}^-)$. Then, F' and F'' are equal up to the variable names, the order of atoms in bodies, and the predicate names of assumptions.*

Example 5. *Let us consider the ABA Learning problem of Example 4 and let $\mathcal{R}_0$ be the set of rules in $F_{(D,\delta,N)}$. We construct an ABA Learning derivation of the form $(RL; Fld; ER; Su; AI^*)^*$ (we only indicate the transformation rule labels) as follows, where by an* iteration *we mean a sequence $RL; Fld; ER; Su; AI^*$.*

First iteration*. Case (RL) applies, thereby getting*

$\rho_1 :\ default(X) \leftarrow X = c_\delta$

$\rho_2 :\ default(X) \leftarrow X = 3$

from $\mathcal{E}_0^+$, that is, $\mathcal{E}^+$. $\mathcal{R} := \mathcal{R}_0 \cup \{\rho_1, \rho_2\}$.

The derivation proceeds by applying case (Fld). Rule ρ_2 can be folded by using the rules in the set $\mathcal{R}_\phi$ listed in Example 3, thereby getting

$\rho_3 :\ default(X) \leftarrow X = 3, a(X), c(X)$

No rule in $\mathcal{R}_\phi$ can be used to fold ρ_1. Thus, $\mathcal{R} := (\mathcal{R} \setminus \{\rho_2\}) \cup \{\rho_3\}$, and the derivation proceeds by applying case (ER)

$\rho_4 :\ default(X) \leftarrow$

$\rho_5 :\ default(X) \leftarrow a(X), c(X)$

and we get $\mathcal{R} := (\mathcal{R} \setminus \{\rho_1, \rho_3\}) \cup \{\rho_4, \rho_5\}$.

Now, case (Su) applies and we have that ρ_5 is subsumed by ρ_4. Thus, $\mathcal{R} := \mathcal{R} \setminus \{\rho_5\}$ and the derivation proceeds by applying case (AI).

Let F and $\mathcal{E}^-$ be the ABA framework and the set of negative examples computed so far, respectively. Given that no assumption has been introduced and $Cov(\mathbb{G}(F), \mathcal{E}^-, \rho_4) = \{default(1), default(2), default(4)\}$, case (AI.2) *applies. A new assumption $\alpha_1(X)$, with contrary $ctr\_\alpha_1(X)$, is introduced in $\mathcal{A}$ and added to the body of ρ_4:*

$\rho_6 :\ default(X) \leftarrow \alpha_1(X)$

No more assumptions are required and, therefore, case (AI.3) applies: $\mathcal{R} := (\mathcal{R} \setminus \{\rho_4\}) \cup \{\rho_6\}$, $\mathcal{E}^+ := \{ctr\_\alpha_1(1), ctr\_\alpha_1(2), ctr\_\alpha_1(4)\}$, and $\mathcal{E}^- := \{ctr\_\alpha_1(c_\delta), ctr\_\alpha_1(3)\}$.

Second iteration*. By (RL), we get:*

$\rho_7 : ctr\_\alpha_1(X) \leftarrow X = 1$

$\rho_8 : ctr\_\alpha_1(X) \leftarrow X = 2$

[8] The folding transformation rule in [21] is defined to be applicable to rules ρ_1, ρ_2 with more general bodies.

[9] The assumption $\alpha(X)$ may belong to $\mathcal{A}$ or be a new one, as specified in the ABA Learning derivations defined in this section (Definition 6) in the next section (Definition 7).

[10] When multiple transformation rules of the same type are applied in sequence, we assume that each application takes the output of the previous as its input.

$\rho_9 : ctr\_\alpha_1(X) \leftarrow X = 4$
and $\mathcal{R} := \mathcal{R} \cup \{\rho_7, \rho_8, \rho_9\}$. *By (Fld) and (ER), we derive:*
$\rho_{10} : ctr\_\alpha_1(X) \leftarrow a(X)$
$\rho_{11} : ctr\_\alpha_1(X) \leftarrow b(X)$
$\rho_{12} : ctr\_\alpha_1(X) \leftarrow b(X), d(X)$
$\mathcal{R} := (\mathcal{R} \setminus \{\rho_7, \rho_8, \rho_9\}) \cup \{\rho_{10}, \rho_{11}, \rho_{12}\}$. *By (Su) we get that* ρ_{12} *is subsumed by* ρ_{11}. *Hence, it is deleted,* $\mathcal{R} := \mathcal{R} \setminus \{\rho_{12}\}$, *and the derivation moves on with (AI). Now, case (AI.2) applies and a new assumption* $\alpha_2(X)$, *with contrary* $ctr\_\alpha_2(X)$, *is introduced in* $\mathcal{A}$ *and added to the body of* ρ_{10}:
$\rho_{13} : ctr\_\alpha_1(X) \leftarrow a(X), \alpha_2(X)$
while no assumption is added to the body of rule ρ_{11}. *Hence, we get* $\mathcal{R} := (\mathcal{R} \setminus \{\rho_{10}\}) \cup \{\rho_{13}\}$. *Then, by (AI.3), we get* $\mathcal{E}^+ := \{ctr\_\alpha_2(3)\}$ *and* $\mathcal{E}^- := \{ctr\_\alpha_2(1)\}$.
Third iteration. *By (RL), we get*
$\rho_{14} : ctr\_\alpha_2(X) \leftarrow X = 3$
and then, by (Fld) and (ER), we get
$\rho_{15} : ctr\_\alpha_2(X) \leftarrow a(X), c(X)$
Given that no rule covers negative examples, by case (AI.3) we get $\mathcal{E}^+ = \emptyset$ *and* $\mathcal{E}^- = \emptyset$. *Thus, the ABA Learning derivation terminates and returns an ABAF* F' *whose set of rules is* $\mathcal{R}' = \mathcal{R}_0 \cup \{\rho_6, \rho_{11}, \rho_{13}, \rho_{15}\}$. $\mathcal{R}'$ *is equal (modulo the predicate names of the assumptions) to the set of rules obtained in Example 3 by applying Definition 1.*

Any ABA Learning derivation can be viewed as an adaptation of the *ABALearn* algorithm presented in [10] specialised to the case where the input ABA Learning problem is constructed from a coherent casebase, by using the *greedy folding* strategy as defined at step (*Fld*) and the assumption introduction strategy defined by (*AI*). Indeed, ABALearn is a nondeterministic algorithm and it is parametric with respect to the specific implementation of the Folding and Assumption Introduction transformation rules.

In our more specific context we strengthen some results about ABALearn. Indeed, we get the following two lemmas.

Lemma 4 (Soundness of ABA Learning for Coherent Casebases). *If ABA Learning terminates for the learning problem* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ *associated with* (D, δ, N), *where* D *is a coherent casebase, then its output* $ABAL(D, \delta, N)$ *is an intensional solution of the problem.*

Proof. (Sketch) Let us consider the sequence, for $i > 0$:
$(F_{i-1}, \mathcal{E}^+_{i-1}, \mathcal{E}^-_{i-1}, RL) \Rightarrow (F_i, \mathcal{E}^+_{i-1}, \mathcal{E}^-_{i-1}, Fld)$
$\Rightarrow (F_{i+1}, \mathcal{E}^+_{i-1}, \mathcal{E}^-_{i-1}, ER)$
$\Rightarrow^k (F_{i+k+1}, \mathcal{E}^+_{i-1}, \mathcal{E}^-_{i-1}, AI)$
$\Rightarrow (F_{i+k+2}, \mathcal{E}^+_{i}, \mathcal{E}^-_{i}, RL)$
The core of the proof consists in showing that if F' is a solution of $(F_{i+k+2}, \mathcal{E}^+_i, \mathcal{E}^-_i)$, then F' is a solution of $(F_{i-1}, \mathcal{E}^+_{i-1}, \mathcal{E}^-_{i-1})$. To see that this invariant holds, note that (1) $\mathcal{E}^+_i$ consists of the contraries to the assumptions introduced by (*AI*) that must be covered in $\mathbb{G}(F')$ to avoid the coverage of negative examples in $\mathcal{E}^-_{i-1}$, and (2) $\mathcal{E}^-_i$ consists of the contraries to the assumptions introduced by (*AI*) that must *not* be covered in $\mathbb{G}(F')$ to preserve the coverage of the positive examples in $\mathcal{E}^+_{i-1}$. Thus, if we get to $(F', \emptyset, \emptyset, RL)$, F' is a solution of $(F_0, \mathcal{E}^+_0, \mathcal{E}^-_0)$. Moreover, F' is an intensional solution, because, due to a previous (*ER*), no rule derived by (*AI*) contains occurrences of constants. □

Lemma 5 (Termination of ABA Learning for Coherent Casebases). *Let* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ *be the ABAF associated with* (D, δ, N), *where* D *is a coherent casebase. Then, ABA Learning terminates for the input* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ *and its output* $ABAL(D, \delta, N)$ *is stratified.*

Proof. (Sketch) Let us consider the set $\mathcal{R}^{AI}_n$ of rules obtained by an ABA Learning derivation of the form:
$(F_B, \mathcal{E}^+, \mathcal{E}^-, RL) \Rightarrow \ldots$
$\Rightarrow (F_{n-1}, \mathcal{E}^+_{n-1}, \mathcal{E}^-_{n-1}, AI) \Rightarrow (F_n, \mathcal{E}^+_n, \mathcal{E}^-_n, RL)$
for $n > 0$. $\mathcal{R}^{AI}_n$ is of the form $\mathcal{R}_\phi \cup \mathcal{R}^\lambda_n$, where $\mathcal{R}_\phi$ are the rules in F_B and $\mathcal{R}^\lambda_n$ are the learnt rules. Each rule in $\mathcal{R}^\lambda_n$ will be of one of the following forms:
$default(X) \leftarrow \alpha(X)$ (ρ_0)
$ctr\_\gamma(X) \leftarrow f_1(X), \ldots, f_k(X), \beta(X)$
$ctr\_\eta(X) \leftarrow g_1(X), \ldots, g_m(X)$
where $f_1, \ldots, f_k, g_1, \ldots, g_m$ are predicates in $\mathbb{F}$, $\alpha(X), \beta(X)$ are assumptions and $ctr\_\gamma(X), ctr\_\eta(X)$ are contraries. Let us consider the AAF $(Args^\lambda, \rightsquigarrow^\lambda)$, where $Args^\lambda = \bigcup_{n>0} \mathcal{R}^\lambda_n$ and $\rightsquigarrow^\lambda$ is defined as follows: for two rules $\rho_1, \rho_2 \in Args^\lambda$, $\rho_1 \rightsquigarrow^\lambda \rho_2$ iff the head of ρ_1 is of the form $ctr\_\alpha_1(X)$ and the assumption $\alpha_1(X)$ occurs in the body of ρ_2.

Let $\rho_0\rho_1 \ldots$ be any sequence of rules in $Args^\lambda$ where, for $i \geq 0$, $\rho_{i+1} \rightsquigarrow^\lambda \rho_i$. By construction, ρ_i is of the form $ctr\_\gamma(X) \leftarrow f_1(X), \ldots, f_k(X), \beta(X)$ and ρ_{i+1} is of the form $ctr\_\beta(X) \leftarrow g_1(X), \ldots, g_m(X), \eta(X)$, where $\{f_1(X), \ldots, f_k(X)\} \subset \{g_1(X), \ldots, g_m(X)\}$ and $\eta(X)$ may be absent. Since the set $\mathbb{F}$ of predicates is finite, and we can assume that no duplicates occur in the body of a rule, there is a maximum length of such sequences. Thus, ABA Learning constructs a finite, acyclic, directed graph $(Args^\lambda, \rightsquigarrow^\lambda)$, that is, ABA Learning terminates for any input $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ and the output $ABAL(D, \delta, N)$ is a stratified ABAF. □

Our soundness and termination results are stronger than the ones for the ABALearn algorithm [10], as the latter may terminate with failure, that is, without producing an intensional solution. Moreover, the result of ABALearn is not necessarily a stratified ABAF, which in our case guarantees soundness with respect to ground extensions (and not stable extensions as in [10]). Furthermore, the next result enforces that an ABA Learning derivation always gets an output that is isomorphic to the one of Definition 1, thus generalising the outcome of Example 5.

Theorem 6. *Let* D *be a coherent casebase,* δ *be the default outcome, and* N *be a new case. Let* $F_{(D,\delta,N)}$ *be the ABAF associated with* $AAF(D, \delta, N)$ *as shown in Definition 1. Let* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ *be the ABA Learning problem associated with* (D, δ, N), *constructed as shown in Definition 2. Then,*

1. *ABA Learning terminates for the input ABA Learning problem* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$, *and returns an ABAF* $ABAL(D, \delta, N)$;
2. $ABAL(D, \delta, N)$ *is a stratified ABAF;*
3. $ABAL(D, \delta, N)$ *is an intensional solution of* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$;
4. $F_{(D,\delta,N)} = ABAL(D, \delta, N)$ *(modulo variable names, predicate names of assumptions, order of atoms in bodies, and presence of the bogus assumption and its contrary in* $ABAL(D, \delta, N)$*).*

Proof. (Sketch) Points (1) and (2) follow directly from Lemma 5. Point (3) follows from Lemma 4.

The main step of the proof for Point (4) consists in showing that the abstract argumentation framework $AAF(D, \delta) = (Args, \rightsquigarrow)$ is isomorphic to the abstract argumentation framework $(Args^\lambda, \rightsquigarrow^\lambda)$.

To see this, consider the mapping $\Phi : Args^{\lambda} \to D \cup \{(\emptyset, \delta)\}$, where $\Phi(\rho) = (X, o_X)$ iff X is the set of features occurring as predicates in the body of ρ (e.g., $\Phi(default(X) \leftarrow \alpha_1(X)) = (\emptyset, \delta)$). To show that Φ is a bijection, we also use the fact that in $Args^{\lambda}$ there is no pair of rules of the form ρ_1 and ρ_2 such that: (i) $head(\rho_1) = head(\rho_2)$, and (ii) the set of non-assumption atoms occurring in $body(\rho_1)$ is a subset of the set of atoms occurring in $body(\rho_2)$. Indeed, ρ_2 would not be derived, due to the application of rule *Su*. We can also see that $\rho_1 \rightsquigarrow^{\lambda} \rho_2$ iff $\Phi(\rho_1) \rightsquigarrow \Phi(\rho_2)$, where $\rightsquigarrow$ is the attack relation in $AAF(D, \delta)$. Then, Point 4 follows from Definition 1, as in particular the following holds: (i) the rules of $F_{(D,\delta,N)}$ are $\mathcal{R}_\delta \cup \mathcal{R}_\chi \cup \mathcal{R}_\phi$, (ii) the rules of $ABAL(D, \delta, N)$ are $\mathcal{R}_\phi \cup Args^{\lambda}$, and (iii) $\mathcal{R}_\delta \cup \mathcal{R}_\chi = \Phi^{-1}(D \cup \{(\emptyset, \delta)\})$, modulo variable names, predicate names of assumptions, and order of atoms in bodies.

Finally, note that the language of the background ABA framework F_B includes a bogus assumption α_0 and its contrary ctr_α_0 (see Definition 2). Thus, α_0 and ctr_α_0 also appear in the language of $ABAL(D, \delta, N)$, while they do not appear in $F_{(D,\delta,N)}$. However, neither α_0 nor ctr_α_0 appear in the rules of $ABAL(D, \delta, N)$. □

We get the following straightforward consequence of Lemma 2 and Theorem 6(4).

Corollary 7. *For a coherent casebase D, $AA\text{-}CBR(D, \delta, N) = \delta$ iff $ABAL(D, \delta, N) \models default(c_N)$.*

6 GREEDY ABA LEARNING FROM INCOHERENT CASEBASES

In the previous sections we have dealt with coherent casebases. When we drop the coherence assumption and we admit an incoherent casebase D, we can still use Definition 2 to construct an ABA learning problem $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ associated with (D, δ, N). However, we cannot guarantee that a solution of $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ (in the sense of Definition 4) exists and can be computed by an ABA Learning derivation (see Definition 6), as shown by the following simple example (a variant of the well known Nixon diamond problem [22]).

Example 6 (Nixon diamond). *Let us consider the ABA framework $F_B = \langle \mathcal{L}, \mathcal{R}, \mathcal{A}, \overline{} \rangle$, where $\mathcal{R}$ is the following set of rules:*

$quaker(X) \leftarrow X = jerry \quad republican(X) \leftarrow X = jerry$
$quaker(X) \leftarrow X = richard \quad republican(X) \leftarrow X = richard$
$quaker(X) \leftarrow X = margery$
$republican(X) \leftarrow X = jennifer$

We also assume that in the universe there is a constant c_δ representing the default case and a rule

$republican(X) \leftarrow X = george$

which represents the feature characterising the new case. The positive and negative examples are the following sets:

$\mathcal{E}_0^+ = \{pacifist(c_\delta), pacifist(jerry), pacifist(margery)\}$
$\mathcal{E}_0^- = \{pacifist(richard), pacifist(jennifer)\}$.

We do not present the source triple (D, δ, N), as the mapping is straightforward.

F_B is incoherent, because jerry and richard, who are classified as pacifist and nonpacifist, respectively, are characterised by the same features, that is, for both the predicates quaker and republican hold. In terms of Definition 3, jerry and richard are not discernible, and by Proposition 1, F_B corresponds to an incoherent casebase.

An ABA Learning derivation is constructed by iterations analogous to Example 5. By the first three iterations we obtain the rules:

$\rho_1 : pacifist(X) \leftarrow \alpha_1(X)$
$\rho_2 : ctr\_\alpha_1(X) \leftarrow republican(X), \alpha_2(X)$.
$\rho_3 : ctr\_\alpha_2(X) \leftarrow republican(X), quaker(X), \alpha_3(X)$.

At the fourth iteration, by (RL) we get:

$\rho_4 : ctr\_\alpha_3(X) \leftarrow X = richard$

and, by (Fld) and (ER):

$\rho_5 : ctr\_\alpha_3(X) \leftarrow republican(X), quaker(X)$

Thus, by (AI.1), we use the previously introduced assumption $\alpha_3(X)$, and we get:

$\rho_6 : ctr\_\alpha_3(X) \leftarrow republican(X), quaker(X), \alpha_3(X)$

The learnt ABA framework F', with rules $\mathcal{R} \cup \{\rho_1, \rho_2, \rho_3, \rho_6\}$, is not a solution according to Definition 4. Indeed, $\mathbb{G}(F') = \mathbb{G}(F_B)$, and hence F' does not cover any positive example. This is due to the self-attacking rule ρ_6, which does not allow the construction of a grounded extension that includes an argument for a claim of the form $pacifist(t)$.

By formalising the above example as a casebase (D, δ, N), we can also see that the construction of Definition 1 is not applicable either, as the AAF $(Args_N, \rightsquigarrow_N)$ is a directed graph with cycles. To overcome this difficulty, we generalise the ABA Learning problem presented in Definition 4 by considering *brave* ABA Learning problems and modifying the definition of an ABA Learning derivation based on stable extensions, instead of grounded extensions, and handling cycles through contraries in a sound way. The modified ABA Learning derivation always guarantees the computation of a *brave* solution.

Definition 7 (Brave ABA Learning). *Given an ABA Learning problem $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ associated with (D, δ, N), a* brave ABA Learning derivation *is a sequence*

$(F_0, \mathcal{E}_0^+, \mathcal{E}_0^-, \Delta_0, r_0) \Rightarrow (F_1, \mathcal{E}_1^+, \mathcal{E}_1^-, \Delta_1, r_1) \Rightarrow \ldots$
$\Rightarrow (F_n, \mathcal{E}_n^+, \mathcal{E}_n^-, \Delta_n, r_n) \ldots$

constructed as in Definition 6, with the following modifications:
(1) for $i \geq 1$, Δ_{i-1} is a stable extension of F_{i-1} such that: (1.a) $\Delta_0 = \mathbb{G}(F_0)$, and (1.b) for $r_{i-1} = RL$, with $i > 1$, $\Delta_i = \Delta_{i-1} \cup \{\emptyset \vdash e \mid e \in \mathcal{E}_{i-1}^+\}$; (1.c) for $r_{i-1} \in \{Fld, ER, Su, AI\}$, with $i > 1$, every $e \in \mathcal{E}_{i-1}^+$ is covered in Δ_{i-1} and the set $\{e \in \mathcal{E}_{i-1}^- \mid e$ is covered in $\Delta_{i-1}\}$ is minimal;
(2) the three cases (AI.1)–(AI.3) are replaced by the following ones:
(AI.1B) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, \Delta_{i-1}, AI) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, \Delta_{i-1}, AI)$, where F_i is obtained by selecting $\rho : h(X) \leftarrow B$ in $\mathcal{R}_{i-1}$ such that there exists $(k(X) \leftarrow B, \alpha(X)) \in \mathcal{R}_{i-1}$, and applying AI to ρ, thereby getting $\mathcal{R}_i = (\mathcal{R}_{i-1} \setminus \{\rho\}) \cup \{h(X) \leftarrow B, \beta(X)\}$, with $\overline{\beta(X)} = k(X)$ if $h(X) = \overline{\alpha(X)}$, and $\beta(X) = \alpha(X)$, otherwise;
(AI.2B) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, \Delta_{i-1}, AI) \Rightarrow (F_i, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, \Delta_{i-1}, AI)$, where F_i is obtained by (1) selecting $\rho : h(X) \leftarrow B$ in $\mathcal{R}_{i-1}$ such that (1a) there exists no rule in $\mathcal{R}_{i-1}$ with body $B, \alpha(X)$, with $\alpha(X) \in \mathcal{A}_{i-1}$ and (1b) $Cov(\Delta_{i-1}, \mathcal{E}_{i-1}^-, \rho) \neq \emptyset$, and (2) applying AI so that: (2a) $\mathcal{A}_i = \mathcal{A}_{i-1} \cup \{\alpha(X)\}$, where $\alpha(X) \notin \mathcal{A}_{i-1}$ is a new assumption with contrary $\overline{\alpha(X)}$, and (2b) $\mathcal{R}_i = (\mathcal{R}_{i-1} \setminus \{\rho\}) \cup \{h(X) \leftarrow B, \alpha(X)\}$;
(AI.3B) $(F_{i-1}, \mathcal{E}_{i-1}^+, \mathcal{E}_{i-1}^-, \Delta_{i-1}, AI) \Rightarrow (F_{i-1}, \mathcal{E}_i^+, \mathcal{E}_i^-, \Delta_{i-1}, RL)$, where every $\rho \in \mathcal{R}_{i-1}$ with $Cov(\Delta_{i-1}, \mathcal{E}_{i-1}^-, \rho) \neq \emptyset$ has an assumption in its body, and $\mathcal{E}_i^+, \mathcal{E}_i^-$ are obtained as follows: (1) $\mathcal{E}_i^+ = \{\overline{\alpha(t)} \mid h(t) \in Cov(\Delta_{i-1}, \mathcal{E}_{i-1}^-, \rho')$, for some $\rho' \in \mathcal{R}_{i-1}\}$, and (2) $\mathcal{E}_i^- = \{\overline{\alpha(t)} \mid h(t) \in Cov(\Delta_{i-1}, \mathcal{E}_{i-1}^+, \rho')$, for some $\rho' \in \mathcal{R}_{i-1}\}$.

In (AI.1B) we handle the case where we derive a rule $ctr\_\alpha(X) \leftarrow B$, and a rule $ctr\_\gamma(X) \leftarrow B, \alpha(X)$ had already been learnt in a previous step. In this case, (AI.1B) avoids the introduction of the self-attacking rule $ctr\_\alpha(X) \leftarrow B, \alpha(X)$ and, instead, learns the rule $ctr\_\alpha(X) \leftarrow B, \gamma(X)$. This learning step may generate an ABA framework with more than one stable extension.

Note that in (AI.2B), (AI.3B) we consider a suitable stable extension Δ_{i-i} of F_{i-1} that covers all positive examples, instead of the grounded extension $\mathbb{G}(F_{i-1})$. The following lemma ensures that, for all $i > 1$, such a stable extension Δ_{i-1} exists, and hence the notion of a brave ABA Learning derivation is well-defined.

Lemma 8. *Let* $(F_0, \mathcal{E}_0^+, \mathcal{E}_0^-, \Delta_0, r_0) \Rightarrow (F_1, \mathcal{E}_1^+, \mathcal{E}_1^-, \Delta_1, r_1) \Rightarrow \ldots \Rightarrow (F_n, \mathcal{E}_n^+, \mathcal{E}_n^-, \Delta_n, r_n) \ldots$ *be a brave ABA Learning derivation. Then, for* $i \geq 1$, *(1) for* $r_{i-1} = RL$, *every positive example* $e \in \mathcal{E}_{i-1}^+$ *is covered in* Δ_i, *and (2) for* $r_{i-1} \in \{Fld, ER, Su, AI\}$, *every* $e \in \mathcal{E}_{i-1}^+$ *is covered in* Δ_{i-1}.

Brave ABA Learning enjoys properties similar to ABA Learning, when we refer to stable extensions instead of grounded extensions.

Lemma 9 (Soundness of Brave ABA Learning). *If brave ABA Learning terminates for the ABA Learning problem* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ *associated with* (D, δ, N), *where D is any (coherent or not) casebase, then its output* $ABAL(D, \delta, N)$ *is an intensional solution.*

Proof. (Sketch) The proof generalises the one for Lemma 4, by considering a stable extension Δ of the learnt ABAF $ABAL(D, \delta, N)$, instead of $\mathbb{G}(ABAL(D, \delta, N))$. □

Lemma 10 (Termination of Brave ABA Learning). *Let* F_B *be the ABAF associated with* (D, δ, N), *where D is any (coherent or not) casebase. Then, brave ABA Learning terminates for the ABA Learning problem* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$.

Proof. (Sketch) The proof is similar to the one for Lemma 5. However, in the case where coherence is not assumed, the AAF $(Args^\lambda, \rightsquigarrow^\lambda)$ satisfies the following weaker property: let $\rho_0, \rho_1 \ldots$ be any sequence of rules in $Args^\lambda$ where, for $i \geq 0$, $\rho_{i+1} \rightsquigarrow^\lambda \rho_i$. By construction, ρ_i is of the form $ctr\_\gamma(X) \leftarrow f_1(X), \ldots, f_k(X), \beta(X)$ and ρ_{i+1} is of the form $ctr\_\beta(X) \leftarrow g_1(X), \ldots, g_m(X), \eta(X)$, where $\{f_1(X), \ldots, f_k(X)\} \subseteq \{g_1(X), \ldots, g_m(X)\}$ and $\eta(X)$ may be absent. In the case where $\{f_1(X), \ldots, f_k(X)\} = \{g_1(X), \ldots, g_m(X)\}$, by *(AI.1B)*, we must have $\eta(X) = \gamma(X)$ and no new assumption is introduced. Thus, no infinite sequence can be constructed by brave ABA Learning. □

For a coherent casebase, a brave ABA Learning derivation coincides with an ABA Learning derivation as presented in Definition 1. Indeed, the case where we derive mutually attacking rules at step *(AI.1B)* will never occur. Thus, the learnt ABAF F' will be stratified and, as already mentioned, the grounded extension of F' coincides with its (unique) stable extension. The following theorem summarizes the results for brave ABA Learning.

Theorem 11. *Let D be a casebase,* δ *be the default outcome, and N be a new case. Then,*

(1) *Brave ABA Learning terminates for the input ABA Learning problem* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ *associated with* (D, δ, N), *and returns an ABAF* $ABAL(D, \delta, N)$ *which is an intensional brave solution of* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$;

(2) *If D is coherent, then* $ABAL(D, \delta, N)$ *coincides with the output of an ABA Learning derivation as in Definition 6.*

The output of ABA Learning can be used to predict the outcomes of new cases using the following notion (with coherent or incoherent casebases alike).

Definition 8. *Let* F' *be the output of an ABA Learning derivation for the learning problem* $(F_B, \mathcal{E}^+, \mathcal{E}^-)$ *associated with* (D, δ, N). *We say that the outcome of* F' *for the new case N is* δ *iff there exists a stable extension* Δ *of* F' *such that (i) for all* $e \in \mathcal{E}^+ \cup \{default(c_N)\}$, *e is covered in* Δ, *and (ii) for all* $e \in \mathcal{E}^-$, *e is not covered in* Δ.

We now revisit the Nixon Diamond example, which illustrates a case of incoherence for which the output of brave ABA Learning is a non-stratified ABAF admitting several stable extensions.

Example 7 (Nixon Diamond (continued)). *From rule* ρ_5, *by the modified step (AI.1B) of brave ABA Learning, we get, instead of* ρ_6:

$\rho_7 : ctr\_\alpha_3(X) \leftarrow republican(X), quaker(X), \alpha_2(X)$

The learnt ABA framework F'', *with rules* $\mathcal{R} \cup \{\rho_1, \rho_2, \rho_3, \rho_7\}$, *admits several stable extensions. Among these, there is a unique stable extension* Δ *in which all the positive examples are covered and no negative is covered, and hence* F'' *is a brave solution of the learning problem considered in Example 6. In* Δ *the new case* $\underline{pacifist}(george)$ *is not covered, and hence the predicted outcome is* $\overline{\delta}$, *i.e., george is predicted to be non-pacifist.*

7 CONCLUSION

We have proposed Greedy ABA Learning, a novel logic-based learning method from casebases, driven by the goal of reducing the high nondeterminism inherent in ABA and limit the search space of possible solutions for ABA Learning. We have proven that Greedy ABA Learning generalises AA-CBR, another logic-based learning method, beyond coherent casebases.

This paper opens several avenues for future work. We plan to experiment with (implementations of) Greedy ABA Learning on tabular datasets, as well as compare it experimentally with other forms of and systems for ABA Learning, notably [10, 23]. It would also be interesting to see whether Greedy ABA Learning could be fruitfully applied to other data modality, e.g., in combination with feature engineering as in [4]. We also intend to integrate Greedy ABA Learning within neuro-symbolic pipelines, in the spirit of [20]. Finally, it would be interesting to explore the addition of preferences to Greedy ABA Learning, e.g., to match [14], and to see whether our method could find natural applicability in legal settings, where AA-CBR has been shown to provide useful insights [12].

ACKNOWLEDGMENTS

We thank support from the Royal Society, UK (IEC\R2\222045). Toni was partially funded by the ERC (grant agreement No. 101020934) and by J.P. Morgan and the RAEng, UK, under the Research Chairs Fellowships scheme (RCSRF2021\11\45). De Angelis and Proietti were supported by the MUR PRIN 2022 Project DOMAIN funded by the EU – NextGenerationEU (2022TSYYKJ, CUP B53D23013220006, PNRR, M4.C2.1.1) and by the PNRR MUR project PE0000013-FAIR (CUP B53C22003630006). De Angelis and Proietti are members of the INdAM-GNCS research group. Finally, we would like to thank Adam Gould for helpful suggestions.

REFERENCES

[1] Leila Amgoud and Mathieu Serrurier. 2008. Agents that argue and explain classifications. *Auton. Agents Multi Agent Syst.* 16, 2 (2008), 187–209. https://doi.org/10.1007/S10458-007-9025-6

[2] Andrei Bondarenko, Phan Minh Dung, Robert A. Kowalski, and Francesca Toni. 1997. An Abstract, Argumentation-Theoretic Approach to Default Reasoning. *Artif. Intell.* 93 (1997), 63–101. https://doi.org/10.1016/S0004-3702(97)00015-5

[3] Oana Cocarascu, Kristijonas Cyras, and Francesca Toni. 2018. Explanatory Predictions with Artificial Neural Networks and Argumentation. In *IJCAI/ECAI Workshop on Explainable Articial Intelligence (XAI-18).*

[4] Oana Cocarascu, Andria Stylianou, Kristijonas Cyras, and Francesca Toni. 2020. Data-Empowered Argumentation for Dialectically Explainable Predictions. In *ECAI 2020 - Proceedings of the 24th European Conference on Artificial Intelligence, 29 August-8 September 2020, Santiago de Compostela, Spain (Frontiers in Artificial Intelligence and Applications, Vol. 325)*, Giuseppe De Giacomo, Alejandro Catalá, Bistra Dilkina, Michela Milano, Senén Barro, Alberto Bugarín, and Jérôme Lang (Eds.). IOS Press, 2449–2456. https://doi.org/10.3233/FAIA200377

[5] Oana Cocarascu and Francesca Toni. 2016. Argumentation for Machine Learning: A Survey. In *Computational Models of Argument - Proceedings of COMMA 2016, Potsdam, Germany, 12-16 September, 2016 (Frontiers in Artificial Intelligence and Applications, Vol. 287)*, Pietro Baroni, Thomas F. Gordon, Tatjana Scheffler, and Manfred Stede (Eds.). IOS Press, 219–230. https://doi.org/10.3233/978-1-61499-686-6-219

[6] Kristijonas Cyras, David Birch, Yike Guo, Francesca Toni, Rajvinder Dulay, Sally Turvey, Daniel Greenberg, and Tharindi Hapuarachchi. 2019. Explanations by arbitrated argumentative dispute. *Expert Syst. Appl.* 127 (2019), 141–156. https://doi.org/10.1016/J.ESWA.2019.03.012

[7] Kristijonas Cyras, Xiuyi Fan, Claudia Schulz, and Francesca Toni. 2017. Assumption-based Argumentation: Disputes, Explanations, Preferences. *FLAP* 4, 8 (2017). http://www.collegepublications.co.uk/downloads/ifcolog00017.pdf

[8] Kristijonas Cyras, Ken Satoh, and Francesca Toni. 2016. Abstract Argumentation for Case-Based Reasoning. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference, KR 2016, Cape Town, South Africa, April 25-29, 2016*, Chitta Baral, James P. Delgrande, and Frank Wolter (Eds.). AAAI Press, 549–552. http://www.aaai.org/ocs/index.php/KR/KR16/paper/view/12879

[9] Emanuele De Angelis, Maurizio Proietti, and Francesca Toni. 2023. ABA Learning via ASP. In *Proceedings 39th International Conference on Logic Programming, ICLP 2023, Imperial College London, UK, 9th July 2023 - 15th July 2023 (EPTCS, Vol. 385)*, Enrico Pontelli, Stefania Costantini, Carmine Dodaro, Sarah Alice Gaggl, Roberta Calegari, Artur S. d'Avila Garcez, Francesco Fabiano, Alessandra Mileo, Alessandra Russo, and Francesca Toni (Eds.). 1–8. https://doi.org/10.4204/EPTCS.385.1

[10] Emanuele De Angelis, Maurizio Proietti, and Francesca Toni. 2024. Learning Brave Assumption-Based Argumentation Frameworks via ASP. In *ECAI 2024 - Proceedings of the 27th European Conference on Artificial Intelligence, 19-24 October 2024, Santiago de Compostela, Spain (Frontiers in Artificial Intelligence and Applications, Vol. 392)*, Ulle Endriss, Francisco S. Melo, Kerstin Bach, Alberto José Bugarín Diz, Jose Maria Alonso-Moral, Senén Barro, and Fredrik Heintz (Eds.). IOS Press, 3445–3452. https://doi.org/10.3233/FAIA240896

[11] Phan Minh Dung. 1995. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. *Artif. Intell.* 77, 2 (1995), 321–358. https://doi.org/10.1016/0004-3702(94)00041-X

[12] Wachara Fungwacharakorn, Ken Satoh, and Bart Verheij. 2023. Constructing and Explaining Case Models: A Case-Based Argumentation Perspective. In *New Frontiers in Artificial Intelligence - JSAI-isAI 2023 International Workshops, JURISIN, SCIDOCA, EmSemi and AI-Biz, Kumamoto, Japan, June 4-6, 2023, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 14644)*, Mayumi Bono, Yasufumi Takama, Ken Satoh, Le-Minh Nguyen, and Setsuya Kurahashi (Eds.). Springer, 100–114. https://doi.org/10.1007/978-3-031-60511-6_7

[13] Michael Gelfond and Vladimir Lifschitz. 1991. Classical Negation in Logic Programs and Disjunctive Databases. *New Gener. Comput.* 9, 3/4 (1991), 365–386. https://doi.org/10.1007/BF03037169

[14] Adam Gould, Guilherme Paulino-Passos, Seema Dadhania, Matthew Williams, and Francesca Toni. 2024. Preference-Based Abstract Argumentation for Case-Based Reasoning. In *Proceedings of the 21st International Conference on Principles of Knowledge Representation and Reasoning, KR 2024, Hanoi, Vietnam. November 2-8, 2024*, Pierre Marquis, Magdalena Ortiz, and Maurice Pagnucco (Eds.). https://doi.org/10.24963/KR.2024/37

[15] Mark Law, Alessandra Russo, and Krysia Broda. 2014. Inductive Learning of Answer Set Programs. In *Logics in Artificial Intelligence - Proceedings of the 14th European Conference, JELIA 2014, Funchal, Madeira, Portugal, September 24-26, 2014. (Lecture Notes in Computer Science, Vol. 8761)*, Eduardo Fermé and João Leite (Eds.). Springer, 311–325. https://doi.org/10.1007/978-3-319-11558-0_22

[16] Guilherme Paulino-Passos and Francesca Toni. 2021. Monotonicity and Noise-Tolerance in Case-Based Reasoning with Abstract Argumentation. In *Proceedings of the 18th International Conference on Principles of Knowledge Representation and Reasoning, KR 2021, Online event, November 3-12, 2021.* 508–518. https://doi.org/10.24963/KR.2021/48

[17] Guilherme Paulino-Passos and Francesca Toni. 2023. Learning Case Relevance in Case-Based Reasoning with Abstract Argumentation. In *Legal Knowledge and Information Systems - JURIX 2023: The Thirty-sixth Annual Conference, Maastricht, The Netherlands, 18-20 December 2023 (Frontiers in Artificial Intelligence and Applications, Vol. 379)*, Giovanni Sileno, Jerry Spanakis, and Gijs van Dijck (Eds.). IOS Press, 95–100. https://doi.org/10.3233/FAIA230950

[18] Henry Prakken and Rosa Ratsma. 2022. A top-level model of case-based argumentation for explanation: Formalisation and experiments. *Argument Comput.* 13, 2 (2022), 159–194. https://doi.org/10.3233/AAC-210009

[19] Nicoletta Prentzas, Constantinos S. Pattichis, and Antonis C. Kakas. 2023. Explainable Machine Learning via Argumentation. In *Explainable Artificial Intelligence - Proceedings of the First World Conference, xAI 2023, Lisbon, Portugal, July 26-28, 2023, Part III (Communications in Computer and Information Science, Vol. 1903)*, Luca Longo (Ed.). Springer, 371–398. https://doi.org/10.1007/978-3-031-44070-0_19

[20] Maurizio Proietti and Francesca Toni. 2023. A Roadmap for Neuro-argumentative Learning. In *Proceedings of the 17th International Workshop on Neural-Symbolic Learning and Reasoning, La Certosa di Pontignano, Siena, Italy, July 3-5, 2023 (CEUR Workshop Proceedings, Vol. 3432)*, Artur S. d'Avila Garcez, Tarek R. Besold, Marco Gori, and Ernesto Jiménez-Ruiz (Eds.). CEUR-WS.org, 1–8. https://ceur-ws.org/Vol-3432/paper1.pdf

[21] Maurizio Proietti and Francesca Toni. 2024. Learning Assumption-Based Argumentation Frameworks. In *Inductive Logic Programming - Proceedings of the 31st International Conference, ILP 2022, Windsor Great Park, UK, September 28-30, 2022 (Lecture Notes in Computer Science, Vol. 13779)*, Stephen H. Muggleton and Alireza Tamaddoni-Nezhad (Eds.). Springer, 100–116. https://doi.org/10.1007/978-3-031-55630-2_8

[22] Raymond Reiter and Giovanni Criscuolo. 1981. On Interacting Defaults. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence, IJCAI '81, Vancouver, BC, Canada, August 24-28, 1981*, Patrick J. Hayes (Ed.). William Kaufmann, 270–276. http://ijcai.org/Proceedings/81-1/Papers/054.pdf

[23] Cristina Tirsi, Maurizio Proietti, and Francesca Toni. 2023. ABALearn: An Automated Logic-Based Learning System for ABA Frameworks. In *AIxIA 2023 - Advances in Artificial Intelligence - Proceedings of the XXIInd International Conference of the Italian Association for Artificial Intelligence, AIxIA 2023, Rome, Italy, November 6-9, 2023 (Lecture Notes in Computer Science, Vol. 14318)*, Roberto Basili, Domenico Lembo, Carla Limongelli, and Andrea Orlandini (Eds.). Springer, 3–16. https://doi.org/10.1007/978-3-031-47546-7_1

[24] Francesca Toni. 2014. A tutorial on Assumption-based Argumentation. *Argument & Computation* 5, 1 (2014), 89–117. https://doi.org/10.1080/19462166.2013.869878

[25] Huaduo Wang, Farhad Shakerin, and Gopal Gupta. 2022. FOLD-RM: A Scalable, Efficient, and Explainable Inductive Learning Algorithm for Multi-Category Classification of Mixed Data. *Theory and Practice of Logic Programming* 22, 5 (2022), 658–677. https://doi.org/10.1017/S1471068422000205