Argumentation-based Reasoning with Preferences

Kristijonas Čyras

Imperial College London, UK

Abstract. One of the main objectives of AI is modelling human reasoning. Since preference information is an indispensable component of common-sense reasoning, the two should be studied in tandem. Argumentation is an established branch of AI dedicated to this task. In this paper, we study how argumentation with preferences models human intuition behind a particular decision making scenario concerning reasoning with rules and preferences. To this end, we present an example of a common-sense reasoning problem complemented with a survey of decisions made by human respondents. The survey reveals an answer that contrasts with solutions offered by various argumentation formalisms. We argue that our results call for advancements of approaches to argumentation with preferences as well as for examination of the type of problems of reasoning with preferences put forward in this paper. Our work contributes to the line of research on preference handling in argumentation, and it also enriches the discussions on the increasingly important topic of preference treatment in AI at large.

1 Introduction

Preferences are quotidian phenomena. They appear in everyday reasoning problems, within utterances such as "I prefer Brownie over Almond cake". Since AI is inherently concerned with human type-of reasoning, preferences are of natural interest to the AI community. Lately, dealing with preference information has increasingly become of great interest in areas of AI that involve reasoning under uncertainty (see e.g. [13, 18, 26, 34] for discussions).

Argumentation (as overviewed in [34]) is by now an acknowledged branch of AI that is particularly concerned with reasoning with incomplete, uncertain and conflicting information. It is widely used in areas such as multi-agent systems, common-sense reasoning, decision making (see e.g. [2, 15, 26, 29, 34]). Preferences in argumentation, as well as in AI on the whole, are used to, for example, qualify the uncertainty of, or discriminate over, information. A principal issue, however, is the lack of consensus on how preferences should be accounted for (see e.g. [26] for a relatively recent overview). Hence the variety of formalisms of argumentation with preferences: e.g. [1, 3, 5, 6, 9, 12, 21, 22, 24, 27, 30, 32, 33, 36, 37].

In this paper we propose a simple common-sense reasoning scenario concerning decision making in the presence of preference information. The scenario is accompanied by a social experiment, i.e. a survey, that indicates the intuitive

decision that people make when faced with the problem in question. By investigating how current approaches to argumentation with preferences handle the scenario, we find they yield different outcomes than human intuition dictates.

The scenario is described as follows.

Example 1 (Cakes). There are three pieces of cakes on a table: a piece of Almond cake, a Brownie, and a piece of Cheesecake. You want to get as many cakes as possible, and the following are the rules of the game.

a) You can take cakes from the table in two 'rounds':
   1. In the first round you can take at most two cakes;
   2. In the second round you can take at most one cake.

b) If you take Almond cake and Cheesecake in the first round, Brownie will not be available in the second round. (Nothing is known about other possible combinations.)

Finally, very importantly, suppose that you *prefer* Brownie over Almond cake. (No other preferences.)

Which pair(s) of cakes would you choose in the first round?

This example falls into the family of reasoning (or, decision making) problems that involve reasoning with *rules* and *preferences*. In this case, there is essentially a single rule, namely that taking Almond cake and Cheesecake removes Brownie.[1] There is also only one preference, namely that Brownie is preferred over Almond cake. Given this information, what could be the reasoning outcome?

Obviously, {Almond cake, Cheesecake} is not a good option, because it prevents one from getting the Brownie. Meanwhile, among pairs {Almond cake, Brownie} and {Brownie, Cheesecake}, both result into the desired outcome. In addition, neither of the two is in conflict with the preference information: on the one hand, since there are no preferences involving Cheesecake, {Almond cake, Brownie} does not violate the preference of Brownie over Almond cake; on the other hand, {Brownie, Cheesecake} satisfies the preference too, as along Cheesecake, the more preferred item is chosen instead of the less preferred one.

So both choices {Almond cake, Brownie} and {Brownie, Cheesecake} seem to be equally good, or 'rational', in the sense that both lead to the desired outcome of obtaining all the cakes and both satisfy the preference relation. As a decision, one could thus randomly choose between the two pairs.

However, upon the inception of the problem, it felt as if the {Brownie, Cheesecake} choice is somehow more intuitive, or preferred. We have therefore posed the question, as formulated in Example 1, to our colleagues and friends, and, perhaps somewhat surprisingly though not unexpectedly, the answer "Brownie and Cheesecake" dominated. With the hypothesis in mind that {Brownie, Cheesecake} is more often the preferred choice, we have then conducted an anonymous survey with precisely the formulation as in Example 1. There were four possible answers:

[1] Since all three cakes are wanted, condition a) simply indicates that the first round decision is pivotal and involves basically only one rule expressed in condition b).

- {Almond cake, Brownie}
- {Brownie, Cheesecake}
- Indifferent between {Almond cake, Brownie}, and {Brownie, Cheesecake}. (I.e. randomly choose one of the two pairs.)
- Other

The three concrete answer choices were randomized for each respondent, while the 'Other' choice, which allowed for a specification, was always the last one.

In the survey, out of total 84 participants, 41 (i.e. 48.81 %) chose "Brownie and Cheesecake", and only 12 (11.90 %) chose "Almond cake and Brownie", while among the rest, 32 (38.10 %) participants said they were indifferent between the two pairs, and 1 (1.19 %) person suggested taking Almond cake and Cheesecake.[2] So, not only that {Brownie, Cheesecake} dominated more than four times {Almond cake, Brownie}; it also got 10 % more responses than the option of being indifferent between the two pairs. This suggests that Brownie and Cheesecake is the more 'intuitive' choice in this problem.

There may presumably be multiple explanations for the mismatch between the results of the survey and the 'rational' solution delineated above. We are, however, not going to speculate on this issue. Rather, we will attempt to present formalizations of the problem in question in various approaches to argumentation with preferences. The analysis will reveal that current formalisms opt mostly for the normative 'rational' solution (i.e. either of the two pairs), rather than the observed 'intuitive' one (Brownie and Cheesecake).

2 Cakes Example in Argumentation

In argumentation information is represented via *arguments* and *attacks* among them. For instance, abstract argumentation (AA) [20] frameworks are tuples $(\mathit{Args}, \rightsquigarrow)$ with a set Args of arguments and a binary attack relation $\rightsquigarrow$ on Args: for $\mathsf{A}, \mathsf{B} \in \mathit{Args}$, we say that A attacks B iff $\mathsf{A} \rightsquigarrow \mathsf{B}$. Whereas in AA the internal structure of arguments is unknown, structured argumentation (see e.g. [7] for a recent overview) allows for more granularity in defining arguments and attacks. Usually, a formal language for representing knowledge is assumed, and arguments are constructed as deductions from premises to conclusions, with attacks constructively defined.

For both abstract and structured argumentation, *semantics* is an essential procedure for designating sets $E \subseteq \mathit{Args}$ of arguments, called *extensions*, which can be deemed collectively acceptable. For example, stable semantics [20] requires that for no arguments $\mathsf{A}, \mathsf{B} \in E$ it holds that $\mathsf{A} \rightsquigarrow \mathsf{B}$, and that for each $\mathsf{A} \notin E$ there is $\mathsf{B} \in E$ with $\mathsf{B} \rightsquigarrow \mathsf{A}$; e.g. if $\mathit{Args} = \{\mathsf{A}, \mathsf{B}\}$ and $\mathsf{B} \rightsquigarrow \mathsf{A}$ is the only

[2] The survey results can be found at https://www.surveymonkey.com/results/SMGLNNBZ8Q/ . We have surveyed PhD students at the Department of Computing, Imperial College London. Invitations to take the survey were distributed by email via the Department's PhD students' mailing list. 79 responses were obtained in a single day. We do not claim any statistically significant findings.

attack, then $\{\mathsf{B}\}$ is a unique stable extension. Many different argumentation semantics have been investigated (see e.g. [4, 34] for overviews). Broadly speaking, they represent different modes of reasoning with respect to certainty of information, i.e. more sceptical, or more credulous. For the purposes of this paper, we will focus on stable semantics, as other semantics do not yield significantly different outcomes with respect to our Cakes example.

A commonality of majority of argumentation formalisms that deal with preference information is to use preferences to modify the attack relation: attacks from less preferred arguments are discarded, i.e. if B is less preferred than A (in symbols, $\mathsf{B} \lhd \mathsf{A}$), then $\mathsf{B} \rightsquigarrow \mathsf{A}$ fails; see e.g. [1, 6, 9, 12, 21, 22, 24, 27, 33]. Some formalisms (e.g. [36, 37]) employ preferences to select among extensions the most 'preferred' ones, or even combine the two uses (e.g. [3]). In what follows we select several representatives of the broad family of approaches, and show that in modelling the Cakes problem, they do not yield the 'intuitive' answer to choose Brownie and Cheesecake. This, we believe, illustrates that the task of capturing human reasoning in argumentation calls for novel approaches to preference handling. Complete specifications of different formalisms are omitted due to space constraints, but simplified expositions sufficient for our purposes are given.

2.1 Deductive Argumentation

We begin with a well known formalism called Deductive Argumentation [8, 9] in which classical logic is commonly used as a basis. Deductive Argumentation can be seen as a representative of those formalisms that employ forms of propositional, first-order, conditional or temporal logics, e.g. [1, 8, 9, 22, 25].

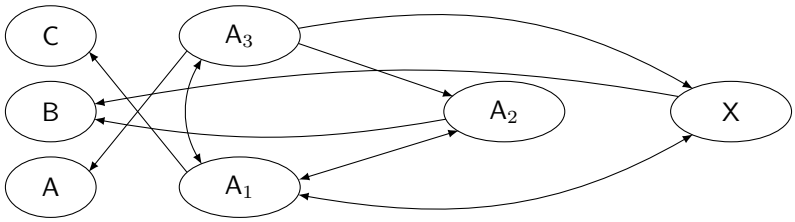
Assume classical propositional logic (PL) with atoms $\mathsf{a}, \mathsf{b}, \mathsf{c}, \ldots$, connectives $\neg, \wedge, \rightarrow$ (including falsum $\bot$) and the classical consequence relation $\vdash$. A knowledge base is a set Δ of (PL) formulas. Given Δ, for $A \subseteq \Delta$ and a formula α, $\langle A, \alpha\rangle$ is an *argument* if $A \vdash \alpha$, $A \nvdash \bot$, and there is no $A' \subsetneq A$ with $A' \vdash \alpha$. We say $\langle A, \alpha\rangle$ *attacks* $\langle B, \beta\rangle$, written $\langle A, \alpha\rangle \rightsquigarrow \langle B, \beta\rangle$, if $\alpha \vdash \neg\mathsf{b}$ for some $\mathsf{b} \in B$. Given a preference relation (strict partial order) $\lhd$ over arguments, we say $\langle A, \alpha\rangle$ *defeats* $\langle B, \beta\rangle$, written $\langle A, \alpha\rangle \hookrightarrow \langle B, \beta\rangle$, if $\langle A, \alpha\rangle \rightsquigarrow \langle B, \beta\rangle$ and $\langle A, \alpha\rangle \not\lhd \langle B, \beta\rangle$.

Deductive Argumentation is completely modular with respect to preference relations $\lhd$ over arguments and is not concerned how they are obtained. Thus, given a preference relation $<$ over, say, atoms, one has to lift it up to the argument level. This is known as preference aggregation. For instance, we can employ versions of the well known Elitist and Democratic ordering principles (see e.g. [32]), respectively called Disjoint Elitist and Disjoint Democratic orders, denoted by $\lhd_{DE}$ and $\lhd_{DD}$, and defined as follows:

$\langle S, \sigma\rangle \lhd_{DE} \langle S', \sigma'\rangle$ iff $\exists \mathsf{s} \in S \setminus S'$ such that $\forall \mathsf{s}' \in S' \setminus S$ we have $\mathsf{s} < \mathsf{s}'$;
$\langle S, \sigma\rangle \lhd_{DD} \langle S', \sigma'\rangle$ iff $\forall \mathsf{s} \in S \setminus S'$ it holds that $\exists \mathsf{s}' \in S' \setminus S$ with $\mathsf{s} < \mathsf{s}'$.

Let us see how Deductive Argumentation equipped with these preference relations copes with our Cakes example.

Example 2 (Cakes in Deductive Argumentation). Suppose that atoms $\mathsf{a}, \mathsf{b}, \mathsf{c}$ stand for the different pieces of cake, namely Almond cake, Brownie, Cheesecake, respectively. The rule that taking Almond cake and Cheesecake in the first round kicks out Brownie in the second one, can be represented as $\mathsf{a} \wedge \mathsf{c} \rightarrow \neg\mathsf{b}$. Finally, the preference of Brownie over Almond cake is specified as $\mathsf{a} < \mathsf{b}$. We thus have the knowledge base $\Delta = \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{a} \wedge \mathsf{c} \rightarrow \neg\mathsf{b}\}$ with preference $\mathsf{a} < \mathsf{b}$. The following arguments can then be constructed:[3] $\mathsf{A} = \langle\{\mathsf{a}\}, \mathsf{a}\rangle$, $\mathsf{B} = \langle\{\mathsf{b}\}, \mathsf{b}\rangle$, $\mathsf{C} = \langle\{\mathsf{c}\}, \mathsf{c}\rangle$, $\mathsf{A}_1 = \langle\{\mathsf{a}, \mathsf{b}\}, \mathsf{a} \wedge \mathsf{b}\rangle$, $\mathsf{A}_2 = \langle\{\mathsf{a}, \mathsf{c}\}, \mathsf{a} \wedge \mathsf{c}\rangle$, $\mathsf{A}_3 = \langle\{\mathsf{b}, \mathsf{c}\}, \mathsf{b} \wedge \mathsf{c}\rangle$, $\mathsf{X} = \langle\{\mathsf{a}, \mathsf{c}\}, \neg\mathsf{b}\rangle$. Comparing relevant arguments, we have: $\mathsf{A}_2 \lhd_{DE} \mathsf{A}_3$; $\mathsf{X} \lhd_{DE} \mathsf{A}_3$; $\mathsf{A}_2 \lhd_{DD} \mathsf{A}_3$; $\mathsf{X} \lhd_{DD} \mathsf{A}_3$. Consequently, regarding both defeat relations $\hookrightarrow_{DE}$ and $\hookrightarrow_{DD}$ (with respect to the orderings $\lhd_{DE}$ and $\lhd_{DD}$), the same argument framework $(\mathit{Args}, \hookrightarrow)$ (where $\hookrightarrow = \hookrightarrow_{DE} = \hookrightarrow_{DD}$) results. It can be represented graphically as follows (in this and further illustrations of argument frameworks, nodes hold arguments and directed edges indicate defeats).



The sets $E = \{\mathsf{B}, \mathsf{C}, \mathsf{A}_3\}$ and $E' = \{\mathsf{A}, \mathsf{B}, \mathsf{A}_1\}$ are the only stable extensions of $(\mathit{Args}, \hookrightarrow)$, representing that one can choose either Brownie and Cheesecake (E), or Almond cake and Brownie (E'). This is the 'rational' solution.

2.2 ASPIC$^+$

We turn to a well established structured argumentation formalism ASPIC$^+$ [14, 32, 33]. It is a widely used expressive approach representative of the use of *defeasible* rules[4] and preference aggregation in argumentation (see e.g. [21, 31–33]).

Lengthy details on ASPIC$^+$ frameworks must be omitted, but the following suffices for our purposes (see [32] for details). Assume a logical language $\mathcal{L}$ closed under negation $\neg$. Specify a set $\mathcal{R}_s$ of (strict) rules and a set $\mathcal{K}_p$ of premises. To satisfy certain rationality criteria [14], impose contraposition (also called transposition) on rules: whenever $\varphi_1, \ldots, \varphi_n \rightarrow \psi \in \mathcal{R}_s$, we have $\varphi_1, \ldots, \varphi_{i-1}, -\psi, \varphi_{i+1}, \ldots, \varphi_n \rightarrow -\varphi_i \in \mathcal{R}_s$ too, for all i.[5]

Arguments are then constructed as deductions from premises using rules; attacks ($\rightsquigarrow$) are obtained by deducing complements of premises; an attack $\mathsf{A} \rightsquigarrow \mathsf{B}$

[3] By definition, there are infinitely many arguments, but it suffices to consider only a finite number of them, as they represent the essential information; see [9] for details.

[4] Two types of rules are commonly used in argumentation: *strict* rules, whose consequent necessarily follows from the antecedent; and *defeasible* rules, whose consequent *normally* (e.g. unless there are exceptions to the rule) follows from the antecedent.

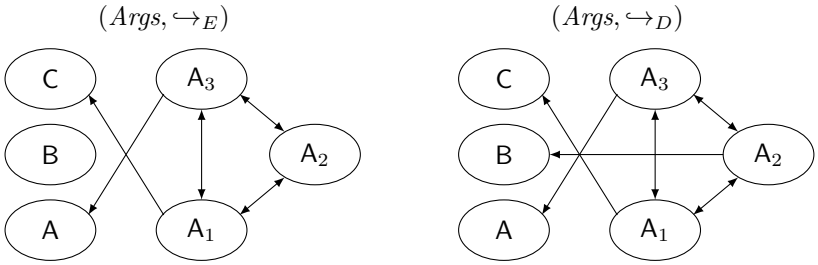
[5] For $\varphi \in \mathcal{L}$, its *complement* $-\varphi$ is: $\neg\psi$ if $\varphi = \psi$; and ψ if $\varphi = \neg\psi$.

succeeds as a defeat only if $\mathsf{A} \not\lhd \mathsf{B}$. ASPIC$^+$ offers two particular argument ordering principles, namely Elitist and Democratic, denoted by $\lhd_E$ and $\lhd_D$ respectively. For our purposes, given a preference relation $<$ over premises, they can be defined as follows: for $S, S' \subseteq \mathcal{K}_p$,

$S \lhd_E S'$ iff $\exists \mathsf{s} \in S$ such that $\forall \mathsf{s}' \in S'$ it holds that $\mathsf{s} < \mathsf{s}'$;
$S \lhd_D S'$ iff $\forall \mathsf{s} \in S$ it holds that $\exists \mathsf{s}' \in S'$ such that $\mathsf{s} < \mathsf{s}'$.

We next formalize the Cakes example in ASPIC$^+$.

Example 3 (Cakes in ASPIC$^+$). Assuming language $\mathcal{L} = \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \neg\mathsf{a}, \neg\mathsf{b}, \neg\mathsf{c}\}$, we have the (strict) rules $\mathcal{R}_s = \{\mathsf{a}, \mathsf{c} \rightarrow \neg\mathsf{b}, \quad \mathsf{a}, \mathsf{b} \rightarrow \neg\mathsf{c}, \quad \mathsf{b}, \mathsf{c} \rightarrow \neg\mathsf{a}\}$ and premises $\mathcal{K}_p = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ (which stand for cakes), together with preference $\mathsf{a} < \mathsf{b}$. We can construct arguments $\mathsf{A} = [\mathsf{a}]$, $\mathsf{B} = [\mathsf{b}]$, $\mathsf{C} = [\mathsf{c}]$, $\mathsf{A}_1 = [\mathsf{A}, \mathsf{B} \rightarrow \neg\mathsf{c}]$, $\mathsf{A}_2 = [\mathsf{A}, \mathsf{C} \rightarrow \neg\mathsf{b}]$, and $\mathsf{A}_3 = [\mathsf{B}, \mathsf{C} \rightarrow \neg\mathsf{a}]$, and obtain the following attacks: $\mathsf{A}_1 \rightsquigarrow \mathsf{C}, \mathsf{A}_2, \mathsf{A}_3$; $\mathsf{A}_2 \rightsquigarrow \mathsf{B}, \mathsf{A}_1, \mathsf{A}_3$; $\mathsf{A}_3 \rightsquigarrow \mathsf{A}, \mathsf{A}_1, \mathsf{A}_2$. Comparing argument premises yields (among other relationships) $\mathsf{A}_2 \lhd_E \mathsf{B}$, but $\mathsf{A}_2 \not\lhd_D \mathsf{B}$. Hence, $\mathsf{A}_2 \not\rightarrow_E \mathsf{B}$, but $\mathsf{A}_2 \hookrightarrow_D \mathsf{B}$, while other attacks succeed as defeats. Therefore, under the Elitist comparison, $\{\mathsf{B}, \mathsf{C}, \mathsf{A}_3\}$, $\{\mathsf{A}, \mathsf{B}, \mathsf{A}_1\}$ and $\{\mathsf{A}, \mathsf{B}, \mathsf{C}, \mathsf{A}_2\}$ are stable extensions (representing the 'rational' choice, and the impossible choice of all three cakes). whereas under the Democratic comparison, $\{\mathsf{B}, \mathsf{C}, \mathsf{A}_3\}$, $\{\mathsf{A}, \mathsf{B}, \mathsf{A}_1\}$ and $\{\mathsf{A}, \mathsf{C}, \mathsf{A}_2\}$ are stable extensions ('rational' choice, and an additional arguably unintuitive choice). The two argument frameworks are drawn below.



2.3 p_ABA

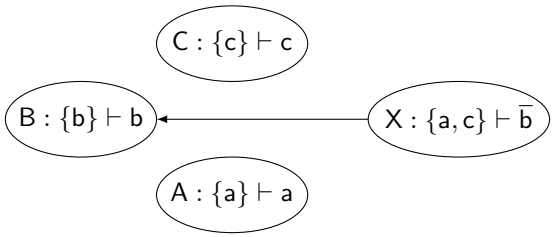
We next investigate *Assumption-Based Argumentation Equipped with Preferences* (p_ABA) [37], which is an extension of the well studied structured argumentation formalism Assumption-Based Argumentation (ABA) [10, 35]. While ABA itself does not possess a mechanism to deal with explicit preferences, to the best of our knowledge p_ABA is the only structured argumentation formalism to use preferences on the extension level to discriminate among extensions.

ABA frameworks are tuples $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$ with $\mathcal{L}$ a language, $\mathcal{R}$ a set of rules, $\mathcal{A} \subseteq \mathcal{L}$ a distinguished set of *assumptions*, and a total mapping $^- : \mathcal{A} \rightarrow \mathcal{L}$ giving *contraries* of assumptions. Similarly to ASPIC$^+$, arguments are constructed as deductions from sets of assumptions using rules, and the attacks are directed at assumptions by deducing their contraries (consult e.g. [35]).

p_ABA frameworks $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, <)$ generalize ABA frameworks by adding a preference ordering $<$ over $\mathcal{L}$. This ordering then induces an ordering over arguments, which in turn induces an ordering $\sqsubseteq$ over extensions of the underlying ABA framework. For two extensions E and E' of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$, $E \sqsubseteq E'$ essentially requires that there is a consequence φ of E' but not E, for which among consequences of E but not E' there is ψ but not ψ' with $\psi < \varphi < \psi'$.[6] p_ABA then selects those extensions E that are not strictly less preferred than any other extension E' (i.e. $E \not\sqsubset E'$ for any E'), subsequently called $\mathcal{P}$-extensions.

Our Cakes example can be cast in p_ABA as follows.

Example 4 (Cakes in p_ABA). We have a p_ABA framework $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, <)$ with language $\mathcal{L} = \{\mathsf{a}, \mathsf{b}, \mathsf{c}, \overline{\mathsf{a}}, \overline{\mathsf{b}}, \overline{\mathsf{c}}\}$, rules $\mathcal{R} = \{\overline{\mathsf{b}} \leftarrow \mathsf{a}, \mathsf{c}\}$, assumptions $\mathcal{A} = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$ (standing for cakes) and preference $\mathsf{a} < \mathsf{b}$. We construct arguments $\mathsf{A} : \{\mathsf{a}\} \vdash \mathsf{a}$, $\mathsf{B} : \{\mathsf{b}\} \vdash \mathsf{b}$, $\mathsf{C} : \{\mathsf{c}\} \vdash \mathsf{c}$, and $\mathsf{X} : \{\mathsf{a}, \mathsf{c}\} \vdash \overline{\mathsf{b}}$. The only attack is $\mathsf{X} \rightsquigarrow \mathsf{B}$, so the underlying ABA framework $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-)$, as depicted below, has a unique stable extension $\{\mathsf{A}, \mathsf{C}, \mathsf{X}\}$ (choice of Almond cake and Cheesecake).



Since the ABA extension is unique, in p_ABA preferences do not play a role and $\{\mathsf{A}, \mathsf{C}, \mathsf{X}\}$ is a unique stable $\mathcal{P}$-extension of $(\mathcal{L}, \mathcal{R}, \mathcal{A}, ^-, <)$.

2.4 Abstract Argumentation

We finally consider AA with preferences. While originally [20] it did not incorporate preference handling mechanism, there are many proposals on how to extend AA with preference information, e.g. [1, 3, 5, 6, 12, 22, 27, 30]. Those approaches follow essentially the same guideline that an attack succeeds only if the attacker is not less preferred than the attackee. We therefore employ, following e.g. [27], a generic representative approach where given an AA framework $(\mathit{Args}, \rightsquigarrow)$ and a preference relation $\lhd$ over arguments, the defeat relation $\hookrightarrow$ is obtained from the condition $\mathsf{A} \hookrightarrow \mathsf{B}$ iff $\mathsf{A} \rightsquigarrow \mathsf{B}$ and $\mathsf{A} \not\lhd \mathsf{B}$.

Consider our attempt to render the Cakes problem in AA with preferences.

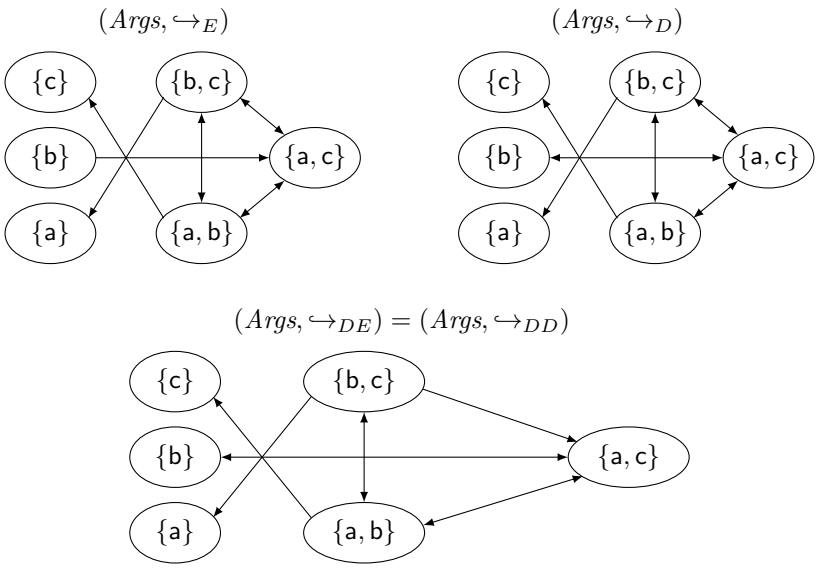
Example 5 (Cakes in AA). Let a, b and c stand for taking a piece of Almond cake, the Brownie and a piece of Cheesecake, respectively. Every combination of a, b and c thus represents a choice of cakes, e.g. $\{\mathsf{b}, \mathsf{c}\}$ typifies the choice of Brownie and Cheesecake. For simplicity, since taking all three pieces of cakes at

[6] For simplicity, we omit the precise definitions of the consequence operator as well as the relation $\sqsubseteq$; see [35, 37] for details.

once is not permitted, we can disregard the choice $\{\mathsf{a},\mathsf{b},\mathsf{c}\}$; similarly, not taking anything, i.e. $\emptyset$, can be discarded.[7] So let $\mathit{Args} = \wp(\{\mathsf{a},\mathsf{b},\mathsf{c}\}) \setminus \{\emptyset, \{\mathsf{a},\mathsf{b},\mathsf{c}\}\} = \{\{\mathsf{a}\},\{\mathsf{b}\},\{\mathsf{c}\},\{\mathsf{a},\mathsf{b}\},\{\mathsf{a},\mathsf{c}\},\{\mathsf{b},\mathsf{c}\}\}$ be the set of arguments.

We can then stipulate the attack relation as follows: for $S, S' \in \mathit{Args}$, let $S \rightsquigarrow S'$ stand for 'if you want the cakes from S, do not choose S''. For example, if Brownie and Cheesecake is wanted, then do not take Almond cake and Cheesecake, and vice versa, i.e. $\{\mathsf{b},\mathsf{c}\} \rightsquigarrow \{\mathsf{a},\mathsf{c}\}$ and $\{\mathsf{a},\mathsf{c}\} \rightsquigarrow \{\mathsf{b},\mathsf{c}\}$. Similarly, if you want Brownie and Cheesecake, then do not choose Almond cake, i.e. $\{\mathsf{b},\mathsf{c}\} \rightsquigarrow \{\mathsf{a}\}$; yet if Almond cake is desired, then it is not a problem to take Brownie and Cheesecake, i.e. $\{\mathsf{a}\} \not\rightsquigarrow \{\mathsf{b},\mathsf{c}\}$. This way, due to the rule that taking Almond cake and Cheesecake prevents getting Brownie, we have (among others) an attack $\{\mathsf{b}\} \rightsquigarrow \{\mathsf{a},\mathsf{c}\}$.

Preference information tells us that $\mathsf{a} < \mathsf{b}$. To account for this we can employ the Elitist and Democratic ordering principles, as well as their Disjoint counterparts, defined in previous sections. Then $S \hookrightarrow S'$ iff $S \rightsquigarrow S'$ and $S \not\lhd S'$, where $\lhd \in \{\lhd_E, \lhd_D, \lhd_{DE}, \lhd_{DD}\}$. Relevant to the attacks, we find: $\{\mathsf{a},\mathsf{c}\} \lhd_E \{\mathsf{b}\}$; $\{\mathsf{a},\mathsf{c}\} \lhd_{DE} \{\mathsf{b},\mathsf{c}\}$; $\{\mathsf{a},\mathsf{c}\} \lhd_{DD} \{\mathsf{b},\mathsf{c}\}$. So we obtain that $\{\mathsf{a},\mathsf{c}\} \not\hookrightarrow_E \{\mathsf{b}\}$, $\{\mathsf{a},\mathsf{c}\} \not\hookrightarrow_{DE} \{\mathsf{b},\mathsf{c}\}$ and $\{\mathsf{a},\mathsf{c}\} \not\hookrightarrow_{DD} \{\mathsf{b},\mathsf{c}\}$. The following graphically depicts the argument frameworks in question.



In $(\mathit{Args}, \hookrightarrow_E)$, $(\mathit{Args}, \hookrightarrow_{DE})$ and $(\mathit{Args}, \hookrightarrow_{DD})$, the sets $E = \{\{\mathsf{b}\},\{\mathsf{c}\},\{\mathsf{b},\mathsf{c}\}\}$ and $E' = \{\{\mathsf{a}\},\{\mathsf{b}\},\{\mathsf{a},\mathsf{b}\}\}$ are the only stable extensions (corresponding to the 'rational' choice), whereas $(\mathit{Args}, \hookrightarrow_D)$, in addition to E and E', admits $E'' = \{\{\mathsf{a}\},\{\mathsf{c}\},\{\mathsf{a},\mathsf{c}\}\}$ as a stable extension.

[7] With the attack relation specified next, this simplification is meant purely to make the argument framework easier to read.

We note that in *Preference-based Argumentation Frameworks* (PAFs) [3], the preference relation $\lhd$ over arguments is used to *reverse*, rather than discard attacks. However, in Example 5 such attack reversal would result into already existing attacks, so the outcome would be the same.

3 Related Work and Discussion

We briefly discuss a few distinguished formalisms, namely Extended Argumentation Frameworks (EAFs) [30, 31], Argumentation Frameworks with Recursive Attacks (AFRA) [5] and Defeasible Logic Programming (DeLP) [23, 24]. Both EAFs and AFRA are generalizations of AA that allow for attacks to be directed at attacks, as well as at arguments. The intuition is that an attack on an attack $\mathsf{A} \rightsquigarrow \mathsf{B}$ expresses a preference for B over A. However, if a preference relation among arguments is not explicitly given, as in the Cakes example, then it is not trivial to figure out what additional attacks on attacks should be introduced.

DeLP, on the other hand, is another structured argumentation formalism that relies on defeasible rules (see Sect. 2.2), and in addition involves the concept of negation as failure [16] from Logic Programming. To represent any non-factual information DeLP employs defeasible rules, and if negation as failure is used, then attacks succeed irrespective of the preferences, like in ABA.

In terms of modelling the Cakes problem from Example 1 in the formalisms considered, we concede that different representations are possible. We chose a natural way to represent a single (strict) rule and a single preference involving two alternatives, and believe that other representations would come at the expense of perspicuity and generality.

We have seen that among the selected formalisms of argumentation (both abstract and structured) with preferences none opt exclusively for the 'intuitive' answer "Brownie and Cheesecake" when modelling the Cakes problem. In Deductive Argumentation (Sect. 2.1) the normative 'rational' solution of choosing indifferently between {Almond cake, Brownie} and {Brownie, Cheesecake} was obtained. ASPIC$^+$ (Sect. 2.2) yielded other choices (such as {Almond cake, Cheesecake}) in addition to the 'rational' one. Preference information in p_ABA (Sect. 2.3) did not really play a role, and just like ABA, the formalism generated a unique choice {Almond cake, Cheesecake}. Finally, AA with preferences (Sect. 2.4) produced either the 'rational choice', or additionally the one afforded by p_ABA too.

The fact that current approaches do not opt exclusively for the 'intuitive' solution in the Cakes example does not mean those argumentation formalisms are ill-suited to handle preferences. Each formalism has been shown to satisfy some desirable properties and adhere to certain intuitions. Some approaches may also aim for being more normative systems, in the sense of complying with certain rationality standards, rather than mimicking human intuition at the expense of rationality. Indeed, common-sense reasoning and human decision making are not necessarily rational endeavours (see e.g. [28]). However, as argumentation is a technique to model common-sense reasoning, it is natural to demand of

formalisms that aspire to this goal to be able to accordingly model common-sense reasoning scenarios such as the one considered in this paper.

4 Conclusion and Future Work

In this paper we investigated whether and how existing approaches to argumentation with preferences model human intuition behind common-sense reasoning scenarios involving rules and preference information. To this end, we proposed an example of a decision making problem that assumes essentially a single strict rule and a single preference. The example was supplemented with both an informal solution to the problem, referred to as the 'rational' choice, and a survey, results of which indicated a solution, referred to as the 'intuitive' choice, that human respondents have opted for most often. We rendered the problem in several representative formalisms of argumentation with preferences, and saw that they produce either the 'rational' or some other choices, but not exclusively the 'intuitive' one. As argumentation is used in areas such as common-sense reasoning and decision making among others, the results, we argue, call for a deeper analysis of how to adequately represent in argumentation the type of problems as the one considered here. Still further, our investigation opens up space for discussions on how to deal with such reasoning problems involving preferences in argumentation-based approaches to, for instance, multi-agent systems, as well as in AI more generally.

Recently, we have proposed a novel formalism, ABA$^+$ [17], to account for preferences in argumentation. ABA$^+$ extends the established structured argumentation formalism Assumption-Based Argumentation [10, 35] with an explicit preference relation over assumptions and integrates an attack reversal mechanism (as proposed for abstract argumentation in Preference-based Argumentation Frameworks [3]) based on preference information. In particular, preferences in ABA$^+$ are incorporated into the attack relation, dispensing with preference aggregation techniques on either argument or extension level. Preliminary analysis indicates that such a method yields the 'intuitive' solution to the Cakes problem discussed in this paper, and we will investigate further.

Additionally, we plan to study scenarios as in the Cakes Example themselves and their treatment in AI. In particular, it would be interesting to see what existing reasoning techniques (e.g. constraint logic programming, Event Calculus) would capture human intuitions in such scenarios. This research direction also includes enquiry into the meaning of preferences and their relation to, for example, rules. In addition, it involves exploring what formal properties regarding preferences can be conceived, as in e.g. [11]. It furthermore means designing and running more surveys to get a better grasp behind human intuitions. Such research is essential to understanding preferences and approaching common-sense reasoning problems from the argumentation perspective in particular, and from the point of view of AI in general.

In AI at large there is much debate on how to manage preferences (see e.g. [19, 26]). Our work contributes to this line of research by presenting a simple example

of reasoning with preferences, opening up space for discussions on how to cope with such common-sense reasoning scenarios in argumentation, and challenging the existing and potential argumentation formalisms to model the human intuition behind them.

References

1. Amgoud, L., Cayrol, C.: A Reasoning Model Based on the Production of Acceptable Arguments. Annals of Mathematics and Artificial Intelligence 34(1-3), 197–215 (2002)
2. Amgoud, L., Prade, H.: Using Arguments for Making and Explaining Decisions. Artificial Intelligence 173(3-4), 413–436 (2009)
3. Amgoud, L., Vesic, S.: Rich Preference-Based Argumentation Frameworks. International Journal of Approximate Reasoning 55(2), 585–606 (2014)
4. Baroni, P., Caminada, M., Giacomin, M.: An Introduction to Argumentation Semantics. The Knowledge Engineering Review 26(04), 365–410 (2011)
5. Baroni, P., Cerutti, F., Giacomin, M., Guida, G.: AFRA: Argumentation Framework with Recursive Attacks. International Journal of Approximate Reasoning 52(1), 19–37 (2011)
6. Bench-Capon, T.: Persuasion in Practical Argument Using Value Based Argumentation Frameworks. Journal of Logic and Computation 13(3), 429–448 (2003)
7. Besnard, P., García, A., Hunter, A., Modgil, S., Prakken, H., Simari, G., Toni, F.: Introduction to Structured Argumentation. Argument & Computation 5(1), 1–4 (2014)
8. Besnard, P., Hunter, A.: A Logic-based Theory of Deductive Arguments. Artificial Intelligence 128(1-2), 203–235 (2001)
9. Besnard, P., Hunter, A.: Constructing Argument Graphs with Deductive Arguments: A Tutorial. Argument & Computation 5(1), 5–30 (2014)
10. Bondarenko, A., Dung, P.M., Kowalski, R., Toni, F.: An Abstract, Argumentation-Theoretic Approach to Default Reasoning. Artificial Intelligence 93(97), 63–101 (1997)
11. Brewka, G., Eiter, T.: Preferred Answer Sets for Extended Logic Programs. Artificial Intelligence 109(1-2), 297–356 (1999)
12. Brewka, G., Ellmauthaler, S., Strass, H., Wallner, J., Woltran, S.: Abstract Dialectical Frameworks Revisited. In: Rossi, F. (ed.) Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI). pp. 803–809. IJCAI/AAAI, Beijing (2013)
13. Brewka, G., Truszczynski, M., Niemelä, I.: Preferences and Nonmonotonic Reasoning. AI Magazine 29(4), 69–78 (2008)
14. Caminada, M., Amgoud, L.: On the Evaluation of Argumentation Formalisms. Artificial Intelligence 171(5-6), 286–310 (2007)
15. Carrera, Á., Iglesias, C.: A Systematic Review of Argumentation Techniques for Multi-Agent Systems Research. Artificial Intelligence Review 44(4), 509–535 (2015)
16. Clark, K.: Negation as Failure. In: Gallaire, H., Minker, J. (eds.) Logic and Data Bases, pp. 293–322. Springer (1978)
17. Čyras, K., Toni, F.: ABA+: Assumption-Based Argumentation with Preferences. In: Principles of Knowledge Representation and Reasoning: Proceedings of the Fifteenth International Conference (KR), to appear. Cape Town (2016)

18. Delgrande, J., Schaub, T., Tompits, H., Wang, K.: A Classification and Survey of Preference Handling Approaches in Nonmonotonic Reasoning. Computational Intelligence 20(2), 308–334 (2004)
19. Domshlak, C., Hüllermeier, E., Kaci, S., Prade, H.: Preferences in AI: An Overview. Artificial Intelligence 175(7-8), 1037–1052 (2011)
20. Dung, P.M.: On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-person Games. Artificial Intelligence 77, 321–357 (1995)
21. Dung, P.M.: An axiomatic Analysis of Structured Argumentation with Priorities. Artificial Intelligence 231, 107–150 (2016)
22. Dunne, P., Hunter, A., McBurney, P., Parsons, S., Wooldridge, M.: Weighted Argument Systems: Basic Definitions, Algorithms, and Complexity Results. Artificial Intelligence 175(2), 457–486 (2011)
23. García, A., Simari, G.: Defeasible Logic Programming: An Argumentative Approach. Theory and Practice of Logic Programming 4(2), 95–138 (2004)
24. García, A., Simari, G.: Defeasible Logic Programming: DeLP-servers, Contextual Queries, and Explanations for Answers. Argument & Computation 5(1), 63–88 (2014)
25. Gorogiannis, N., Hunter, A.: Instantiating Abstract Argumentation with Classical Logic Arguments: Postulates and Properties. Artificial Intelligence 175(9-10), 1479–1497 (2011)
26. Kaci, S.: Working with Preferences. Less is More. Springer (2011)
27. Kaci, S., van der Torre, L.: Preference-Based Argumentation: Arguments Supporting Multiple Values. International Journal of Approximate Reasoning 48(3), 730–751 (2008)
28. Kahneman, D., Tversky, A.: Prospect Theory: An Analysis of Decision under Risk. Econometrica 47(2), 263–291 (1979)
29. Kakas, A., Moraitis, P.: Argumentation Based Decision Making for Autonomous Agents. In: The Second International Joint Conference on Autonomous Agents & Multiagent Systems (AAMAS). pp. 883–890. ACM Press, Melbourne (2003)
30. Modgil, S.: Reasoning About Preferences in Argumentation Frameworks. Artificial Intelligence 173(9-10), 901–934 (2009)
31. Modgil, S., Prakken, H.: Reasoning About Preferences in Structured Extended Argumentation Frameworks. Frontiers in Artificial Intelligence and Applications 216(9-10), 347–358 (2010)
32. Modgil, S., Prakken, H.: A General Account of Argumentation with Preferences. Artificial Intelligence 195, 361–397 (2013)
33. Modgil, S., Prakken, H.: The ASPIC+ Framework for Structured Argumentation: A Tutorial. Argument & Computation 5(1), 31–62 (2014)
34. Rahwan, I., Simari, G.: Argumentation in Artificial Intelligence. Springer (2009)
35. Toni, F.: A Tutorial on Assumption-Based Argumentation. Argument & Computation 5(1), 89–117 (2014)
36. Wakaki, T.: Preference-Based Argumentation Built from Prioritized Logic Programming. Journal of Logic and Computation 25(2) (2013)
37. Wakaki, T.: Assumption-Based Argumentation Equipped with Preferences. In: Dam, H.K., Pitt, J.V., Xu, Y., Governatori, G., Ito, T. (eds.) PRIMA: Principles and Practice of Multi-Agent Systems - 17th International Conference. Lecture Notes in Computer Science, vol. 8861, pp. 116–132. Gold Coast (2014)