Toward a Logical Tool for Generating New Arguments in an Argumentation Based Framework

Geoffroy Aubry and Vincent Risch LSIS - UMR CNRS 6168 Domaine Universitaire de Saint-Jérôme 13397 Marseille Cedex 20, France {geoffroy.aubry, vincent.risch}@lsis.org

Abstract

Following the framework proposed by Besnard and Hunter for argumentation, this paper aims to propose a logical tool for the generation of new arguments when two formal agents have to face their respective knowledge. The following notions are addressed: the behaviour of an agent facing an argument, the answer of an agent in front of a set of formulas, and relations among arguments. X-logics, a nonmonotonic extension of classical propositional logic proposed by Siegel and Forget, is used as the background formalism for representing the reasoning of the agents on arguments.

1 Introduction

The problem of how to represent argumentation with logical tools is an old question, tackled by many authors in many different ways (see [4]). A common view is that an argument is composed with a set of reasons, a conclusion, and a method of inference by wich the conclusion is meant to follow from the reasons. One question is to find a way for representing and processing the arguments exchanged by two agents: obviously, this question is related with nonmonotonic reasoning¹ (e.g. [7], [5], [1] among others). Another important question concerns the way of combining sentences for or against a given conclusion, which is related to the notion of *acceptability* of an argument [5]. The association of linking sentences with an argumentation tree has been investigated by [2]. It is noteworthy that this last work avoids any reference to a dialectic representation of distinct agents. In this paper, we consider a logical framework composed of two formal agents facing their respective knowledge and debating about them: one agent supports a

certain conclusion, and the other has to find an answer regarding his state of knowledge. Our approach attempts to give a formal description of how the second agent can generate a relevant answer in front of the argument claimed by the first agent. The necessity to cope with a possible revision of the respective knowledge of the agents leads us to use X-logics, a nonmonotonic extension of classical propositional logics, for representing the reasoning of the agents on arguments. Our paper is organized as follows: section 2 below briefly describes X-logics, section 3 introduces the notion of agent, section 4 the notion of attitude of an agent in front of a set of formulas, section 5 has to do with answers and generation of new arguments, and section 6 with an application to the framework of [2]. Formaly, our langage is classical propositional logics where \vdash denotes the classical consequence relation, \top and \perp the usual truth values, and \neg , \lor , \land , \Rightarrow the usual connectors. Formulas are denoted by lower case letters whereas sets of formulas are denoted by shift case letters. A finite set A of formulas is logically interpreted by the conjunction of its elements, that is a sentence. We abuse the notation $\neg A$ as a shorthand for the negation of the conjunction of the formulas in A. A finite consistent set of formulas is called a *knowledge base*². Finaly, given E a set of sets, $\min(E)$ is the set of minimal sets of E regarding inclusion.

2 X-logics

X-logics were defined in [9] as an attempt for defining a proof theory for nonmonotonic logics from any classical logic with a given set X of formulas. Whereas classically $K \vdash f$ iff $\overline{K \cup \{f\}} = \overline{K}$ (where the line over the sets denotes the theorems associated with these sets), X-logics can be considered as a generalization (hence a weakening) of \vdash , namely \vdash_X , defined such as $K \vdash_X f$ iff $\overline{K \cup \{f\}} \cap$ $X = \overline{K} \cap X$ *i.e.* \vdash_X is monotonic only on X. When X

¹A way of reasoning that assumes that truth is no given for eternity, contrary to classical logics.

²Note that in [2], consistency is not required for knowledge bases

equals the language, \vdash_X is equivalent to \vdash . If $X = \{\bot\}$ then $K \vdash_X f$ is equivalent to $K \nvDash \neg f$ which describes the consistency relation between K and f (" $K \land f$ is satisfiable" holds), provided K is consistent by itself. If $X = \emptyset$, all the formulas can be entailed. Actually, the following result holds:

Theorem 2.1. $K \vdash_X f$ iff $(\forall x \in X \setminus \overline{K})(K \land f \nvDash x)$

Note that $K \vdash_X f$ if every theorem (regarding \vdash) of $K \cup \{f\}$ which is in X is a theorem of K (by adding f to K the set of classical theorems which are in X does not grow). Intuitively, X can be considered as the set of "pertinent" informations relevant to our mode of reasoning: a set K of informations entails f if the addition of f to K does not produce more pertinent formulas than with K alone. Although this was already proved independently, also note this theorem shows that X-logics are supraclassical.

Let us make use of the following terminology: if $K \vdash_X f$ we say that f is *admissible* by K, and *non admissible* otherwise. The following properties obviously hold:

Property 2.2. 1) (metacoherence) A formula cannot be both admissible and non admissible. 2) (paraconsistency) Both a formula and its negation can be admissible.

As shown in [3], X-logics coincide with permissive inference relations which are completely characterized by Reflexivity, Left Logical Equivalence, Right Weakening, Conjunctive Cautious Monotony, Cut and Or.

3 Agents

In the litterature, some argumentation theories consider the notion of *proponant-opponent* [8, 11] whereas other describe argumentation systems in which arguments made from a unique set of formulas are linked together, in a kind of abstract game among arguments [7, 5, 1, 2]. Following [10], we make use of the following notion of *agent*:

Definition 3.1. An agent is a couple [K, X] where K is a knowledge base, and $X \supseteq \{\bot\}$, a set of formulas. The set of agents, a subset of $2^{\mathcal{L}} \times 2^{\mathcal{L}}$, is denoted by \mathcal{A} .

Whereas K is used as a representation of the knowledge of an agent, X can be considered as a set of contraints on the reasoning of this agent. The obligation made for X to contain at least the contradiction is motivated by the requirement for an agent to reason consistently. In other words, in the context of an agent, the notion of admissibility covers at least consistency (K is consistent and $\perp \in X$).

Having in mind to further define the construction of new arguments by an agent, we are more especially interested by the different possible cases regarding the admissibility of a formula or its negation, that is the attitudes that this agent can have in front of a given formula. Theses attitudes depend both on the knowledge and the constraints; as already mentionned, we may perfectly have that both $K\vdash_X f$ and $K\vdash_X \neg f$, that is f as well as $\neg f$ are admissible by the agent (property 2.2.2). Actually, regarding the admissibility/non-admissibility of a formula or its negation, the following elementary sentences can be enumerated: $K \vdash_X f$, or $K \nvDash_X f$, or $K \vdash_X \neg f$, or $K \nvDash_X \neg f$. Let us label respectively (a), (b), (c), (d) these elementary cases.

4 Attitudes

Theorem 4.1. Exactly eight distinct cases can be considered regarding the admissibility of a formula or its negation: (a), (b), (c), (d), (ac), (ad), (bc), (bd).

Let us call *attitude* any one among these eight available cases. The table 1 gives a synthetic view of how the attitudes organize themselves regarding the language: each line is a partition among the formulas, and each row describes inclusions among attitudes ((bc) is included in (b) and (c)).

а		b	
ad	ac	bc	bd
d	С		d

Table 1. Relations among attitudes

Definition 4.2. Consider an agent [K, X] and a formula f. [K, X] is for f iff $K \vdash_X f$ and $K \nvDash_X \neg f$, neutral about f iff $K \vdash_X f$ and $K \vdash_X \neg f$, against f iff $K \nvDash_X f$ and $K \vdash_X \neg f$, puzzled by f iff $K \nvDash_X f$ and $K \nvDash_X \neg f$.

By extension, an agent is for (resp. neutral about, against, puzzled by) a set of formulas iff it is for (resp. neutral about, against, puzzled by) the conjunction of the formulas of this set.

Replacing the labels by their corresponding sentences in table 1 yields the graph 1: every node is one of the eight avaible cases, and the edges link pair of nodes that have a non-empty intersection; the three lines/partitions of the previous table are the three horizontal plans that intersect with the vertex of the octahedron, whereas the non-horizontal edges detail the embedding among attitudes.

Remark. There is an obvious interpretation of the middle plan of the octahedron in the framework of Belnap's logic \mathcal{FOUR} : by associating respectively the four attitudes for, neutral, against, and puzzled with the truth values *true*, \top , *false* and \bot of this logic, the complete lattice first proposed by Fitting [6] is reconstructed via the X-inference. However, whereas Belnap's logic has only one negation, an obvious distinction has to be made by differentiating between negation of a formula and negation of admissibility



Figure 1. Octahedron of attitudes

(*i.e.* non-admissibility). Moreover, an agent does not commit to an attitude in the absolute, but relatively to a set of formulas, which makes more complex the interpretation of attitudes as pure truth values.

Property 4.3. Consider an agent Φ and a set of formulas $A: \Phi$ is for A iff it is against $\neg A, \Phi$ is neutral about A iff it is neutral about $\neg A, \Phi$ is puzzled by A iff it is puzzled by $\neg A, \Phi$ is for the tautologies and against the contradictions.

How does the behaviour of an agent evolve when a new formula is added to the initial set of formulas about which he is for, neutral, against, or puzzled? This question is especially relevant when one consider the possibility for an agent to construct a strategy.

Property 4.4. Giving two sets A and B of formulas, the table 2 lists the attitudes an agent may adopt about $A \cup B$, regarding respectively his attitude about A and his attitude about B.



Table 2. Evolution of the attitudes of an agent regarding a growing set of formulas

When a set grows, the agent facing this set tends to be against it. This may also indicate that a cautious strategy for avoiding conflicts would be to construct only small arguments.

The attitudes an agent can have in front of a set of formulas leads to a partition of this set. Confrontation operators allow to reach one or the other element of the partition. We also define a "technical" operator for non-admissibility.

Definition 4.5. The operator $|_+$ (resp. $|_0$, $|_-$ and $|_p$) maps an agent [K, X] and a set E of formulas with the subsets of E such as this agent is for (resp. neutral, against or puzzled) these subsets:

$$\begin{aligned} |_{+}: & \mathcal{A} \times 2^{\mathcal{L}} & \longrightarrow & 2^{2^{\mathcal{L}}} \\ & [K,X] \mid_{+} E & \longmapsto & \{P \subseteq E \mid K \vdash_{X} P \text{ and } K \nvDash_{X} \neg P\} \\ |_{0}: & \mathcal{A} \times 2^{\mathcal{L}} & \longrightarrow & 2^{2^{\mathcal{L}}} \\ & [K,X] \mid_{0} E & \longmapsto & \{P \subseteq E \mid K \vdash_{X} P \text{ and } K \vdash_{X} \neg P\} \\ |_{-}: & \mathcal{A} \times 2^{\mathcal{L}} & \longrightarrow & 2^{2^{\mathcal{L}}} \\ & [K,X] \mid_{-} E & \longmapsto & \{P \subseteq E \mid K \nvDash_{X} P \text{ and } K \vdash_{X} \neg P\} \\ |_{p}: & \mathcal{A} \times 2^{\mathcal{L}} & \longrightarrow & 2^{2^{\mathcal{L}}} \\ & [K,X] \mid_{p} E & \longmapsto & \{P \subseteq E \mid K \nvDash_{X} P \text{ and } K \nvDash_{X} \neg P\} \end{aligned}$$

Definition 4.6. The non-admissibility operator $|_{t}$ maps an agent [K, X] and a set E of formulas with the subsets of E such as each of these subsets is non-admissible by the agent [K, X]:

Note that the non-admissibility operator is more interesting than a potentential dual admissibility operator, since the first allows to identify precisely conflicting formulas (due to the existential quantifier in $(\exists x \in X \setminus \overline{K})(K \cup A \vdash x))$.

However, note also that using the non-admissibility operator is merely a technical choice, which does not limit our approach only to the generation of disputing arguments. This is studied in detail in the next section.

5 Generating arguments

Definition 5.1 (cf. [10, 2, 1]). Let Δ be a finite set of formulas. An argument of Δ is the couple $\langle S, c \rangle$ such as: 1) $S \nvDash \bot$; 2) S is a minimal subset of Δ such as $S \vdash c$.

S is called the support of the argument, and *c* its conclusion. The sets of arguments is a subset of $2^{\mathcal{L}} \times \mathcal{L}$, written *Arg.* supp (α) denotes the support of the argument α , and concl (α) denotes its conclusion.

Essentialy, an argument can be considered as a set of formulas (the support) used for classically proving a formula (the conclusion). The minimality criterion allows to only consider relevant formulas for proving the conclusion. Besnard and Hunter [2] show that the following relation of attack can handle various existing relations in the litterature: **Definition 5.2.** [2] An argument β attacks³ an argument α iff concl $(\beta) = \neg(s_1 \land \cdots \land s_n)$, with $\{s_1, \ldots, s_n\} \subseteq$ supp (α) .

We define the relation of *defense* among arguments:

Definition 5.3. An argument β defends an argument α iff concl $(\beta) = s_1 \wedge \cdots \wedge s_n$, with $\{s_1, \ldots, s_n\} \subseteq \text{supp }(\alpha)$.

We are interested by constructing new arguments from both a set of formulas, and the attitudes of an agent regarding this set. We use a new notion of *answer* in order to link an agent with an argument: roughly, an answer is a set of formulas fixed by the attitude of the agent. For instance, if $K \not\vdash_X A$ then $(\exists x \in X \setminus \overline{K})(K \cup \{\neg x\} \vdash \neg A)$ (with a consistent A): $K \cup \{\neg x\}$ is called an answer to A.

Definition 5.4. An answer of the agent [K, X] to a consistent set A of formulas is a consistent set R of formulas such that: 1) $R = K' \cup \bigcup_{x \in X'} \{\neg x\}$; 2) $K' \subseteq K$ and $X' \subseteq X$; 3) $K' \nvDash_{\{\bot\} \cup X'} A$. The set of answers givent by [K, X] to A is written $\mathcal{R}^{A}_{[K, X]}$.

Remark. About definition 5.4:

Considering $K' \nvDash_{X'} A$ instead of $K' \nvDash_{\{\bot\} \cup X'} A$ would imply that an answer would not be constructed only from the knowledge of the agent. For instance, in this case, $\{a, a \Rightarrow b\}$ would not be considered an answer of the agent $[\{a, a \Rightarrow b\}, \{\bot\}]$ to $\{\neg b\}$, contrary to $\{a, a \Rightarrow b, \top\}$. Introducing the contradiction allows to construct shorter answers.

Condition 3) of definition 5.4 may suggest that an answer is based on a pure eristic approach. However, remind that beeing against A is beeing for $\neg A$ (*cf.* property 4.3).

Property 5.5. Any answer of an agent to a given set of formulas is inconsistent with this set. The opposite does not hold.

Property 5.6. If R is a subset of the knowledge of an agent, such that it is inconsistent with a consistent set A of formulas, then R is an answer of the agent to the set A.

Property 5.7. If a consistent set of formulas is non admissible by an agent, then there exists an answer of this agent to this set. The opposite does not hold.

Property 5.8. The answers of an agent [K, X'] to a set of formulas A' are answers of the agent [K, X] to A if X contains X', and if A contains A'.

Corollary 5.9. *R* is an answer of the agent $[K, \{\bot\}]$ to *A* iff *R* is inconsistent with *A*.

Since an argument satisfies a minimality criterion, we are interested by the shortest answers that can support the future conclusions of an argument.

Definition 5.10. An answer of the agent Φ to a set A of formulas is called relevant iff it does not contain any other answer of \mathcal{R}_{Φ}^{A} . The set of relevant answers given by Φ to A is written $\mathcal{R}p_{\Phi}^{A}$.

Let us now consider an agent [K, X] having to produce a relevant answer to a consistent set A of formulas. Several cases have to be considered: (1) the set A contradicts a constraint of the agent, which means that the agent has to produce this constraint; (2) the set A contradicts a knowledge of the agent; (3) the agent will have to use both his knowledge and his constraints in order to construct an answer:

Theorem 5.11. 1) If $\nvdash_X A$, then $\{\neg x \mid x \in \overline{A} \cap X \setminus \overline{\top}\} \subseteq \mathcal{R}p^A_{[K,X]};$ 2) If $K \cup A \vdash \bot$, then $\min([A, \{\bot\}] \mid_f K) \subseteq \mathcal{R}p^A_{[K,X]};$ 3) If $K' \in \min([A, X \setminus \overline{K}] \mid_f K)$ and $K' \notin \mathcal{R}p^A_{[K,X]},$ then $\{K' \cup \{\neg x\} \mid x \in \overline{K' \cup A} \cap X\} \subseteq \mathcal{R}p^A_{[K,X]}$

Note that 3) does not allow to capture the complete set of relevant answers related both to the knowledge and to the constraints of an agent. Answers of this category can be interpreted as lies, and we do not study them in this paper.

Theorem 5.12. The relevant answers made by the agent $[K, \{\bot\}]$ to A are the minimal subsets of K that are non-admissible by the agent $[A, \{\bot\}]$.

Definition 5.13. An argument of an agent Φ is an argument $\langle R, \neg A \rangle$ such as R is a relevant answer of Φ to A. The set of arguments of an agent Φ is written Arg_{Φ} .

This definition refines definition 5.1: an argument of an agent is an argument in the sense of definition 5.1, but in addition it is now possible to take in account the constraints of an agent for constructing arguments.

Definition 5.14. The set of arguments of an agent Φ attacking (resp. defending) an argument α is written $Arg_{\Phi}^{\mathsf{att}(\alpha)}$ (resp. $Arg_{\Phi}^{\mathsf{def}(\alpha)}$).

From the definitions 5.13 and 5.14: $\mathcal{A}rg_{\Phi}^{\operatorname{att}(\alpha)} = \{\langle R, \neg A \rangle \mid R \in \mathcal{R}p_{\Phi}^{A}, A \subseteq \operatorname{supp}(\alpha)\} ; \mathcal{A}rg_{\Phi}^{\operatorname{déf}(\alpha)} = \{\langle R, \bigwedge_{a \in A} a \rangle \mid R \in \mathcal{R}p_{\Phi}^{\{\neg A\}}, A \subseteq \operatorname{supp}(\alpha)\}.$

The attitude of an agent regarding a set of formulas can now be linked to the existence of an argument justifying this attitude.

Theorem 5.15.

- $A \in \Phi \mid_{\texttt{f}} \mathsf{supp}\left(\alpha\right) \; \Rightarrow \; \exists \langle R, \neg A \rangle \in \mathcal{A}rg_{\Phi}^{\mathsf{att}(\alpha)}$
- $A \in \Phi \mid_{-} \operatorname{supp}(\alpha) \Rightarrow \exists \langle R, \neg A \rangle \in \mathcal{A}rg_{\Phi}^{\operatorname{att}(\alpha)}$
- $\bullet \ A \in \Phi \mid_{+} \mathsf{supp} \left(\alpha \right) \ \Rightarrow \ \exists \langle R, \bigwedge_{a \in A} a \rangle \in \mathcal{A} rg_{\Phi}^{\mathsf{def}(\alpha)}$

³Besnard and Hunter use the expression "is an undercut" of *alpha*

•
$$A \in \Phi \mid_{p} \operatorname{supp}(\alpha) \Rightarrow \begin{cases} \exists \langle R_{1}, \neg A \rangle \in \mathcal{A}rg_{\Phi}^{\operatorname{att}(\alpha)} \\ \exists \langle R_{2}, \bigwedge_{a \in A} a \rangle \in \mathcal{A}rg_{\Phi}^{\operatorname{def}(\alpha)} \end{cases}$$

The fact that an agent is neutral about a subset of the support of an argument does not allow this agent to construct any answer, and hence does not allow him to generate an argument.

6 An application : the generation of maximal conservative undercuts

In order to collect arguments and counter-arguments for or against an initial thesis, Besnard and Hunter [2] construct argumentative trees in which the arguments are embedded, hence simulating a debate. The union of all argumentative trees inside an argumentative structure allows to measure how much credit can be given in favour of this thesis.

These trees only contain *maximal conservative undercuts* (MCU), that are the kind of arguments kept by Besnard and Hunter for their relevance. We show how to generate such arguments.

Definition 6.1. [2] An argument α is more conservative than an argument β iff supp $(\alpha) \subseteq$ supp (β) and concl $(\beta) \vdash$ concl (α) .

Definition 6.2. [2] An argument β is a MCU of the argument α iff β is attacking α such as no other argument attacking α is strictly more conservative than β .

Theorem 6.3. [2] If $\langle B, \neg(a_1 \land \cdots \land a_n) \rangle$ is a MCU of the argument α , then supp $(\alpha) = \{a_1, \ldots, a_n\}$.

In order to go back to the notion of agent, we consider that the knowledge contained in the unique set Δ considered by Besnard and Hunter belongs to two distinct agents before an operation of union. Since the notion of constraint is not defined in this argumentative system, the agent's set of constraints is limited to $\{\bot\}$. In order to express the notion of maximal conservative undercut inside our framework, we make it relative to the knowledge of an agent.

Theorem 6.4.

 $\langle R, \neg \mathsf{supp}(\alpha) \rangle$ is a MCU of α iff $R \in \mathcal{R}p^{\mathsf{supp}(\alpha)}_{[K, \{\bot\}]}$

Example 6.5. Consider the first argument tree in the example 9.3 of [2]. We search the MCU of the root (*i.e.* $\alpha = \langle \{b \Rightarrow a, b\}, a \rangle$). Consider the two following agent from Δ : $P = [\{a \Leftrightarrow \neg d, b \Rightarrow a, g \land \neg b, d\}, \{\bot\}]$ and $O = [\{a \Leftrightarrow \neg d, b, b \Rightarrow a, \neg g, \neg d\}, \{\bot\}].$

Following both theorem 6.4 and theorem 5.12 we have to search min ([supp $(\alpha), \{\bot\}$] $|_{\not f} K_P$): {{ $g \land \neg b$ }, { $d, a \Leftrightarrow \neg d$ }} is found. Hence we get two MCU: {{ $g \land \neg b$ }, \neg supp (α) } and {{ $d, a \Leftrightarrow \neg d$ }, \neg supp (α) }, that corresponds to the example of [2]. The next nodes are obtained similarly.

7 Conclusion and future work

This work investigates the question of the generation of new arguments. To do this, and from the paraconsistent framework defined by X-logics, we first enumerated all the possible attitudes of an agent facing a given set of formulas. We further introduced the notion of confrontation operators: these operators yield a partition of the langage from the attitudes of an agent, and the correpondings sets of formulas look like new classes of acceptability. The notion of answer, linked to the support of an adverse argument, allowed us to define the construction of new arguments. An application is the generation of maximal conservative undercuts inside the framework of Besnard and Hunter. The diversity of the attitudes an agent may adopt leads us to consider the modelisation of the notion of argumentative strategy as a future direction of research. Especialy, our framework should allow us to represent a form of lie. Further, the interpretation of the attitudes as truth values in Belnap's logic FOURrepresents another direction of research.

References

- L. Amgoud and C. Cayrol. On the acceptability of arguments in preference-based argumentation. In UAI, pages 1– 7, 1998.
- [2] P. Besnard and A. Hunter. A logic-based theory of deductive arguments. Artificial Intelligence, 128(1-2):203–235, 2001.
- [3] A. Bochman. Brave nonmonotonic inference and its kinds. Annals of Mathematics and Artificial Intelligence, 39(1– 2):101–121, 2003.
- [4] C. I. Chesñevar, A. G. Maguitman, and R. P. Loui. Logical models of argument. ACM Comput. Surv., 32(4):337–383, 2000.
- [5] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–358, 1995.
- [6] M. Fitting. Bilattices and the semantics of logic programming. J. Log. Program., 11(1&2):91–116, 1991.
- [7] F. Lin and Y. Shoham. Argument systems: A uniform basis for nonmonotonic reasoning. In *KR*, pages 245–255, 1989.
- [8] N. Rescher. Dialectics, A Controversy-Oriented Approach to the Theory of Knowledge. State University of New York Press, Albany., 1977.
- [9] P. Siegel and L. Forget. A representation theorem for preferential logics. In *Proceedings of the fifth International Conference on the Principles of Knowledge Representationand Reasoning, KR'96*, pages 453–460. Morgan Kaufmann, 1996.
- [10] G. R. Simari and R. P. Loui. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence*, 53(2-3):125–157, 1992.
- [11] G. Vreeswijk. Reasoning with defeasible arguments: Examples and applications. In D. Pearce and G. Wagner, editors, *Logics in AI: Proc. of the European Workshop JELIA*'92, pages 189–211. Springer, Berlin, Heidelberg, 1992.