An approach for identifying design principles in argument systems[*]
—a preliminary report—

M. Capobianco C. I. Chesñevar G. R. Simari

Artificial Intelligence Research and Development Laboratory
Department of Computer Science and Engineering
Universidad Nacional del Sur – Av. Alem 1253, (8000) Bahía Blanca, ARGENTINA
EMAIL: {mc,cic,grs}@cs.uns.edu.ar

ABSTRACT

Artificial Intelligence deals with the challenge of modeling commonsense reasoning. In the last fifteen years argument-based systems have came forward to meet this challenge as knowledge representation and reasoning tools able to model incomplete and conflicting information. We believe that the argumentation field has sufficiently matured in the last years to deserve a formal analysis of properties of the main existing argumentation systems.

In this work we propose a set of general principles for argumentation enunciated in abstract terms, that can be analyzed in the context of several argumention formalisms. We intend that these principles help to achieve a unified formal view of argumentation properties.

1 INTRODUCTION

Artificial Intelligence deals with the challenge of modeling commonsense reasoning. In the last fifteen years argument-based systems have came forward to meet this challenge as knowledge representation and reasoning tools able to model incomplete and potentialy inconsitent information.

The advantages of these systems when modelling common sense reasoning have sprung a new set of argumentation-based applications. In particular, agent systems, and intelligent agents may profit from the use argumentation formalisms, since knowledge representation issues play a major role in these systems. Well known problems in the agent community involve the need of complex abilities for reasoning, planning and acting in dynamic environments [13]. Specially tailored argumentation systems have also been proposed for agents interacting in such environments [1].

Argumentation has also gained wide acceptance in the multiagent systems (MAS) area by providing tools for designing and implementing different features which characterize interaction among rational agents [8]. The key to this success relies in a straightforward formalization of common-sense reasoning that can be used in decision-making systems. Furthermore, the argument-based approach can also explain the reasons that motivated a particular recommendation.

Argumentation systems have recently been used in prototypes of real world applications, such as clustering algorithms [5], intelligent web search [2], reasoning in multiagent systems [9, 1], It is also expected that in a near future argumentation may be integrated with other applications such as intelligent interfaces, e-commerce agents, and automatic information processing, among others. This seemd to indicate that the argumentation field has evolved considerable in the last years.

We believe that the field has sufficiently matured to deserve a formal analysis of properties for the main existing argumentation systems. In these analysis we ought to seek for the similarities and differences among the existing systems, and determine their consequences.

In what follows we propose a set of general principles for argumentation. These principles are enunciated in abstract terms that can be analyzed in the context of several argumention formalisms. We intend that these principles help to achieve a unified formal view of argumentation properties.

2 GENERAL PRINCIPLES FOR ARGUMENT SYSTEMS

To analyze properties of argumentation systems we identify two key elements in these formalisms: a process for obtaining arguments and a classification process that determines whether arguments are warranted. Only warranted arguments can be used to obtain inferences. This inference pattern is generally used by argumentation systems despite their differences.

In this section we enunciate a set of abstract principles we think should hold on argument systems. These principles are naturally associated

---

[*] Partially supported by Agencia Nacional de Promoción Científica y Tecnológica (PAV 2003 Nro. 076, PICT 2002 Nro. 13096, PICT 2003 Nro. 15043), CONICET (*Consejo Nacional de Investigaciones Científicas y Técnicas de la República Argentina*), and CIC (*Comisión de Investigaciones Científicas de la Provincia de Buenos Aires*).

with a set of postulates.

2.1 Argument level

The argument construction level is what Prakken and Vreeswijk call *logical level.* In this level arguments are obtained from the knowledge base according to some specific process. Generally, the knowledge base contains some kind of *schemes* that can be considered as domain-independent reasoning patterns. These schemes resemble rules of inference in deductive logics, but are not restricted to them. They can include non-deductive logics and are not restricted to classical logical inference. This is the first process that suffers the information in the knowledge base.

Every argument requires some kind of derivation or a general notion of consequence that connects its conclusion with a set of premises and schemes obtained from the the knowledge base. Some formalisms (such as [11, 4, 7]) also require a set of restrictions on this derivation so it can be considered an argument. Others prefer to take an unrestricted approach (for example, [6, 12]).

In this work we present arguments in favor of the first view. We show that using unrestricted arguments leads to a set of undesirable situations in the inference process. We believe that at least two restrictions should be supported: consistency and minimality. Consistency is by far the most important. Generally, an argument is considered consistent if it does not support contradictory literals.

Consistency avoids *self-defeating arguments*[6]. An argument is self-defeating if it defeats itself according to the system's defeat definition. Pollock [6] mentions that in his system this type of arguments collapses the inference process, an devices a set of conditions to avoid them in the defeat graph construction. This and other problems will arise in any argumentation system if self-defeating arguments are allowed.

We believe that self defeating arguments should be avoided in every argument system. Nevertheless, even in the case that self defeat is allowed, it should be evident that no self defeating argument should be warranted.

Postulate 2.1 (*Self defeating*) For an argumentation system AS, its inference procedure should be warrant any self defeating argument. ■

If self-defeating arguments are simply avoided this postulate is trivially satisfied. This happens when arguments are required to be consistent, since argumentation systems use the notion of consistency to define attack or conflict among arguments. An argument A attacks another argument B if some element in B cannot coexist with some element in A, given that their mutual aceptance produces an inconsistency. Therefore, a notion of internal consistency of arguments avoids any of such conflicts.

Another usual restriction is minimality of arguments [11, 4, 7]. This notion varies according to the argument definition. If arguments are defined as sets minimality is a fundamental requirement. In such cases minimality avoids unnecessary rules. Note that unnecessary rules present in an argument only weaken it, creating more attack points. An argument B could conflict with another argument A attacking one of the superfluous rules in A. Even this seems to be an obvious requirement, there are argumentation systems that do not avoid this kind of attacks, such as [12].

2.2 Inference engine

To support an abstract analysis of inference properties, we endorse viewing argumentative inference engines as composed of four fundamental relations among arguments:

- Subargumentation: This relation denotes a kind of *inclusion* between arguments. In fact when arguments are defined as sets it is simply defined as set inclusion. If arguments are defined as inference trees, they are defined as sub-trees. It is highly dependant on argument definition, but the general idea remains the same in every case. Let AS be an argumentation system, then we denote its associated sub-argument relation as S_{AS}.
- Attack: This notion has been defined in many ways in argumentation systems. The general idea is to characterize conflicting pairs of arguments. Let AS be an argumentation system, then we denote its associated attack relation as A_{AS}.
- Concordance: This relation is the dual of attack. Two arguments are concordant if their simultaneous acceptance does not lead to a contradiction. This relation, such as subargumentation, can be considered an agreement relation among arguments. Let AS be an argumentation system, then we denote its associated concordance relation as C_{AS}.
- Defeat: This relation is a refinement of attack. It defines a way of deciding conflicts between arguments. It is usually based on an preference criterion. Let AS be an argumentation system, then we denote its associated defeat relation as D_{AS}.

These four relations are key elements in argumentation systems. Next we formulate a set of postulates using these relations and the following definitions:

Definition 2.1 (Internal attacks) Let AS be an argumentation system. The relation *internal attacks*, denoted as IA_{AS} is defined as $A_{AS} \cap S_{AS}$. ■

Intuitively, this relation represent attacks occurring in self defeating arguments. If no self defeat is allowed, $IA_{AS} = \emptyset$.

Definition 2.2 (Concordant family) Let AS be an argumentation system and A an argument in the AS system. The family of arguments concordant with A, denoted as $\mathsf{concordant}(\mathcal{A})$ is defined as: $\mathsf{concordant}(\mathcal{A}) = \{B | (A, B) \in C_{AS}\}$. ■

The first postulate relates subargumentation and concordance.

Postulate 2.2 (*Inclusion*) Let AS be an argument systems. Then $S_{AS} - IA_{AS} \subset C_{AS}$, that is to say, the subargumentation relation minus the internal attacks relation should be a subset of the concordance relation. ■

If no self defeating arguments are allowed, the sub-argumentation relation should a subset of the concordance relation. This can be seen as a stronger version of this postulate that holds in this special class of argument systems.

If self defeating is allowed, the inclusion can still hold if we first eliminate the subargumentation pairs that involve self defeating arguments. This is reflected in postulate 2.2.

Postulate 2.3 (*Disjoints sets*) Let AS be an argument systems. Then $A_{AS} \cap C_{AS} = \emptyset$, that is to say, attack and concordance should not have common pairs. ■

This is also an intuitive principle: an argument A is either attacking B or is concordant with B, but should not do both at the same time.

There are other basic postulates, for example, sub-argumentation should be an order relation among arguments. Concordance should be symmetric and reflexive (only when self-defeating is not allowed) but not transitive.

Concordance can also be used to control the way arguments reinstate, avoiding fallacious situations, as stated in the following postulate:

Postulate 2.4 (*Restricted reinstatement*) Given an argument A, only arguments in its concordant family, $\mathsf{concordant}(\mathcal{A})$, should be used to reinstate A. ■

Reinstatement is a common concept in most modern argument systems. It defines a kind of support. It is reasonable that only arguments concordant with A can support or defend A from other attacks. Otherwise, an argument B could be supporting and attacking an argument A.

The next postulate concerns the defeat relation. In general, this relation relies on the on the preference criteria used by the system. Different criteria have been proposed in the literature, and some authors have even chosen to parametrize the systems with respect to the preference criterion to make it more flexible [4, 7].

In this setting, different criteria can be used according to the particular domain under consideration. Even tough this is an interesting approach, we consider that some general restrictions should be used to prevent inadequate criteria from being used. In this sense, Vreeswijk [12] and Simari *et al.* [10] have proposed different sets of restrictions. In what follows we analyze both approaches to obtain a general postulate regarding preference criteria.

Simari *et al.* propose that the preference criterion should establish a partial order in the set of arguments. Vreeswijk enunciated a set of restrictions to ensure the coherence of the defeat relation. To do this, a given preference criterion $<$ should:

1. be reflexive and transitive,
2. do not form infinite chains: $A_1 < A_2 < \ldots < A_n < \ldots$,
3. for every pair of arguments A and B such that A is a sub-argument of B it holds that $B \leq A$.

In this case the relation $\leq$ may not be a partial order, considering that antisymmetry does not necessarily hold. If we compare both approaches, we can see that reflexivity and transitivity are common elements.

Infinite chains are not explicitly forbidden in Simari's approach, but combining transitivity and antisymmetry this arises as a collateral result in systems with a finite set of arguments:

Proposition 2.1 Let $\leq$ be a reflexive, antisymmetric and transitive preference criterion over a finite set of arguments. Then there are no infinite chains of the form: $A_1 < A_2 < \ldots < A_n < \ldots$. ■

Proof Suppose there is an infinite chain σ. Since the set of arguments is finite, there is an argument A_1 such that $\sigma = A_1 < A_2 < \ldots < A_n < A_1 < \ldots$. Then σ is circular. By transitivity, we can affirm that $A_1 < A_n$. And by antisymmetry, since $A_n < A_1$, $A_1 \not< A_n$. This contradiction arose from the initial supposition. □

But conditions (1) and (2) of Vreeswijk's definition do not imply that $<$ is a partial order.

Therefore, requiring that the preference criterion be a partial order is more restrictive than the first and second condition of Vreeswijk's work. Nevertheless, antisymmetry seems an intuitive and adequate condition.

Regarding the third condition it is reasonable than an argument cannot be made stronger by adding more rules, but we believe that this condition should emerge as a property of the argumentation system, rather than be required by an explicit condition of the preference criterion. To sum up, we conclude the following postulate about preference criterion:

Postulate 2.5 (*Preference criteria*) Preference criteria for argumentation systems should induce a partial order on the set of arguments. ■

It is interesting to consider the interaction among the defined relations and the set of warranted arguments that are sanctioned by a given knowledge base. The first postulate in this respect is associated with the concept of conflict-free sets [3]:

Postulate 2.6 (*Conflict-free warrant*) The set of warranted arguments of a given knowledge base must be a conflict-free set, that is, there is no pair of arguments A, B such that A is a counterargument of B. ■

A natural corollary of this postulate is the consistency of the set of conclusions sanctioned by the system.

Another important postulate relates subargumentation and the set of warranted arguments, and was first enunciated by Vreeswijk [12] as a desirable property of argumentation systems.

Postulate 2.7 (*Warranted sub-arguments*) Let A be a warranted argument. Then for every argument B such that B is a sub-argument of A it should hold that B is a warranted argument. ■

In other words, when accepting an argument A, we are implicitly accepting all of its subarguments.

3 CONCLUSIONS AND FUTURE WORK

We have proposed a set of principles for the design of argument systems general enough to be applied to different formalisms. By the analysis of these principles we have identified a set of key aspects in argumentation systems. Principle 2.1 emerges among them as a fundamental issue, being a base for many coherent properties in argument systems.

Moreover, we believe that self-defeating arguments should better be eliminated from the inference mechanism. On the one hand this simplifies the system, and on the other hand it is also a philosophical matter, self defeating arguments do not have the strength to support their conclusions.

As future work, we believe that it should be interesting to test the proposed principles in the main existing argumentation formalisms.

References

[1] Capobianco, M., Chesnevar, C., and Simari, G. Argumentation and the dynamics of warranted beliefs in changing environments. *Journal of Autonomous Agents and Multi-Agent Systems*, 11 (2005), 127–151.

[2] Chesñevar, C., and Maguitman, A. An argumentative approach to assesing natural language usage based on the web corpus. In *Proc. of European Conference on Artificial Intelligence (ECAI 2004). Valencia, Spain (to appear)* (Aug. 2004), ECCAI.

[3] Dung, P. M. On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning and Logic Programming and n-person Games. *Artificial Intelligence 77* (1995), 321–357.

[4] García, A., and Simari, G. Defeasible Logic Programming: An Argumentative Approach. *Theory and Practice of Logic Programming 4*, 1 (2004), 95–138.

[5] Gomez, S., and Chesñevar, C. A Hybrid Approach to Pattern Classification Using Neural Networks and Defeasible Argumentation. In *Proc. of Intl. 17th FLAIRS Conference. Palm Beach, FL, USA* (May 2004), AAAI, pp. 393–398.

[6] Pollock, J. L. *Cognitive Carpentry: a blueprint for how to build a person.* Bradford/MIT Press, 1995.

[7] Prakken, H., and Sartor, G. A System for Defeasible Argumentation with Defeasible Priorities. In *Proceedings of the International Conference on Formal and Applied Practical Reasoning* (1996), Springer Verlag, pp. 510–524.

[8] Rahwan, I. Guest editorial: Argumentation in multi-agent systems. *Autonomous Agents and Multiagent Systems 11*, 2 (2005), 115–125.

[9] Rahwan, I., Ramchurn, S. D., Jennings, N. R., P. McBurney, S. P., and Sonenberg, L. Argumentation-based negotiation. *The Knowledge Engineering Review 18*, 4 (2003), 343–375.

[10] Simari, G. R., Chesñevar, C. I., and García, A. J. The Role of Dialectics in Defeasible Argumentation. In *Proceedings of the XIV Conferencia Internacional de la Sociedad Chilena para Ciencias de la Computación* (Nov. 1994), pp. 111–121. http://cs.uns.edu.ar/giia.html.

[11] Simari, G. R., and Loui, R. P. A Mathematical Treatment of Defeasible Reasoning and its Implementation. *Artificial Intelligence 53*, 1–2 (1992), 125–157.

[12] Vreeswijk, G. Abstract Argumentation Systems. *Artificial Intelligence 90*, 1–2 (1997), 225–279.

[13] Wooldridge, M., Ed. *An Introduction to Multiagent Systems.* John Wiley & Sons, 2002.