A Survey on Probabilistic Argumentation

Rolf Haenni[1], Bernhard Anrig[2], Jürg Kohlas[2], and Norbert Lehmann[2]

[1] Computer Science Deptartment, University of California, Los Angeles, CA 90095
rolf.haenni@gmx.net,
WWW home page: http://haenni.shorturl.com

[2] Department of Informatics, University of Fribourg, CH-1700 Fribourg, Switzerland
{bernhard.anrig,juerg.kohlas,norbert.lehmann}@unifr.ch,
WWW home page: http://www2-iiuf.unifr.ch/tcs/

Abstract. Most formal approaches to argumentative reasoning focus on the analysis of qualitative aspects. An exeption is the framework of probabilistic argumentation systems. Its philosophy is to include both qualitative and quantitative aspects of argumentation through a simple way of combining logic and probability theory. Probabilities are used to weigh arguments for and against a particular conclusion. The aim of this paper is to help establishing probabilistic argumentation as a legitimate formal model and a powerful tool for argumentation.

1 Introduction

In the last couple of years, *argumentation* has gained growing recognition as a new and promising research direction in artificial intelligence. As a consequence of this increasing interest, different authors have investigated argumentation and its application in various domains. By looking at today's literature on this subject, one realizes that argumentation is understood in fairly different ways. The variety of attempts to study the nature of arguments and the process of argumentation is therefore characterized by its broad diversity.

Many authors consider argumentation as a *dialectical process* during disputations (e.g. in [33]). Some authors also speak about *negotiation* (e.g. in [22]). Often, their intended field of applications are legal cases (e.g. in [26]). Other authors focus their studies on the problem of the *acceptability* and the *comparison* of arguments (e.g. in [13, 1]). A recent contribution in this area is the *logic-based theory of deductive arguments* [11]. The common feature of all these approaches is their restriction to different types of logic. As a consequence, they are all limited in the way they combine arguments for and against a particular conclusion. A comprehensive overview of modeling argumentation in logics can be found in [27].

The approach we present in this paper is known as *probabilistic argumentation systems* (PAS) [15]. This technique has been developed independently from the approaches mentioned above. Its philosophy is somehow different, as it is not uniquely based on logic, but also on probability theory. Probabilities are used to weigh arguments for and against a particular conclusion. The result is a pair of quantitative aggregation functions that can be used to determine whether the conclusion is taken to hold or not. The strength of our method comes certainly from this simple way of combining logic and probability theory. In fact, this demonstrates how non-monotonicity can be obtained within the framework of classical logic, thus making the necessity of developing non-monotonic logics (such as default logic or circumscription) questionable. Furthermore, it provides a new way of interpreting and understanding Dempster-Shafer theory and its connection to probability theory and Bayesian networks [20].

In our opinion, the PAS framework is the most sophisticated formal model found in the literature so far, that covers both, qualitative and quantitative aspects of non-monotonic reasoning under uncertainty. Moreover, compared to most other formal models of argumentation or non-monotonic logics, the strength of PAS comes also from the existence of efficient computational methods [15, 14]. This makes PAS not only interesting as a theoretical model, but also as a candidate for serious applications in practice. The aim of this paper is to summarize the state of the art. Our hope is to contribute to a better understanding of the nature of arguments and to establish PAS as a legitimate formal model and a powerful tool for argumentation.

2 The Basic Idea

The idea of the PAS framework goes back to the concept of *assumption-based truth maintenance systems* (ATMS) [12]. It is also closely related to *abduction* [10, 24] and the *independent choice logic* [25]. Thus, the goal is not to describe argumentation as a dialectical process, but rather to serve as a *deductive tool* that helps to judge *hypotheses*, that is open questions about the unknown or future world, in the light of the given uncertain and partial knowledge. From a qualitative point of view, the problem is to derive *arguments* in favor and *counter-arguments* against the hypothesis of interest. An argument is a defeasible proof built on uncertain assumptions, that is a chain of deductions based on assumptions that makes the hypothesis true. Efficient algorithms are obtained by focussing the search on the most relevant arguments [15, 14].

In a second step, a quantitative judgement of the situation is obtained by considering these probabilities that the arguments are valid. This is done in a similar way as in [28, 7] (see [18] for a discussion). The *credibility* of a hypothesis can then be measured by the total probabilities that it is supported or defeated by arguments. Conflicts are handled through conditioning. The resulting *degrees of support* and *possibility* correspond to *belief* and *plausibility*, respectively, in the Dempster-Shafer theory of evidence [29, 19, 32]. Although a qualitative judgement may be valuable, a quantitative judgement is often more useful and helps to decide whether a hypothesis can be accepted, rejected, or whether the available knowledge does not permit to decide. The complete situation is illustrated in Figure 1.

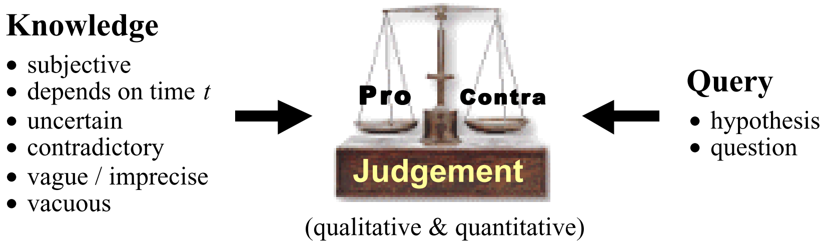

Fig. 1. Arguments for and against a hypothesis are placed on a balance and weighed by corresponding probabilities.

A fundamental property of PAS is that additional knowledge causes the quantitative judgement of the situation to change non-monotonically. Clearly, the property of *non-monotonicity* is required in any mathematical formalism for reasoning under uncertainty. It reflects a natural property of how a human's conviction or belief can change when new information is added.

3 The Formal Model

For the construction of a probabilistic (propositional) argumentation system, consider two disjoint sets $A = \{a_1, \ldots, a_m\}$ and $P = \{p_1, \ldots, p_n\}$ of propositions. The elements of A are called *assumptions*. $\mathcal{L}_{A \cup P}$ denotes the corresponding propositional language. If ξ is an arbitrary propositional sentence in $\mathcal{L}_{A \cup P}$, then a triple (ξ, P, A) is called *(propositional) argumentation system*. ξ is called *knowledge base* and is often specified by a conjunctively interpreted set $\Sigma = \{\xi_1, \ldots, \xi_r\}$ of sentences $\xi_i \in \mathcal{L}_{A \cup P}$ or, more specifically, clauses $\xi_i \in \mathcal{D}_{A \cup P}$, where $\mathcal{D}_{A \cup P}$ denotes the set of all (proper) clauses over $A \cup P$.

The assumptions play an important role for expressing uncertain information. They are used to represent uncertain events, unknown circumstances and risks, or possible outcomes. Conjunctions of literals of assumptions are of particular interest. They represent possible scenarios or states of the unknown or future world. $\mathcal{C}_A$ denotes the set of all such conjunctions. Furthermore, $N_A = \{0, 1\}^{|A|}$ represents the set of all possible interpretations relative to A. The elements $s \in N_A$ are called *scenarios*. Our theory is based on the idea that one particular scenario $\hat{s} \in N_A$ is the *true* scenario.

Consider now the case where a second propositional sentence $h \in \mathcal{L}_{A \cup P}$ called *hypothesis* is given. Hypotheses represent open questions or uncertain statements about some of the propositions in $A \cup P$. What can be inferred from ξ about the possible truth of h with respect to the given set of unknown assumptions? Possibly, if some of the assumptions are set to *true* and others to *false*, then h may be a logical consequence of ξ. In other words, h is *supported* by certain scenarios $s \in N_A$ or corresponding *arguments* $\alpha \in \mathcal{C}_A$. Note that *counter-arguments* refuting h are arguments supporting $\neg h$.

More formally, let $\xi_{\leftarrow s}$ be the formula obtained from ξ by instantiating all the assumptions according to their values in s. We can then decompose the set of scenarios N_A into three disjoint sets $I_A(\xi) = \{s \in N_A : \xi_{\leftarrow s} \models \bot\}$, $SP_A(h,\xi) = \{s \in N_A : \xi_{\leftarrow s} \models h,\, \xi_{\leftarrow s} \not\models \bot\}$, and $RF_A(h,\xi) = \{s \in N_A : \xi_{\leftarrow s} \models \neg h,\, \xi_{\leftarrow s} \not\models \bot\}$ of *inconsistent*, *supporting*, and *refuting scenarios*, respectively. Furthermore, if $N_A(\alpha) \subseteq N_A$ denotes the set of models of a conjunction $\alpha \in \mathcal{C}_A$, then we can define corresponding sets of *inconsistent*, *supporting*, and *refuting arguments* of h relative to ξ by $I(h,\xi) = \{\alpha \in \mathcal{C}_A : N_A(\alpha) \subseteq I_A(h,\xi)\}$, $SP(h,\xi) = \{\alpha \in \mathcal{C}_A : N_A(\alpha) \subseteq SP_A(h,\xi)\}$, and $RF(h,\xi) = \{\alpha \in \mathcal{C}_A : N_A(\alpha) \subseteq RF_A(h,\xi)\}$, respectively. Often, since $I(h,\xi)$, $SP(h,\xi)$, and $RF(h,\xi)$ are *upward-closed* sets, only corresponding *minimal* arguments are considered.

So far, hypotheses are only judged qualitatively. A quantitative judgment of the situation becomes possible if every assumption $a_i \in A$ is linked to a corresponding *prior probability* $p(a_i) = \pi_i$. Let $\Pi = \{\pi_1, \ldots, \pi_m\}$ denote the set of all prior probabilities. We suppose that the assumptions are mutually independent. A quadruple (ξ, P, A, Π) is then called *probabilistic argumentation system* [15].

In order to judge h quantitatively, consider now the conditional probability that the true scenario $\hat{s}$ is in $SP_A(h,\xi)$ but not in $I_A(\xi)$. In the light of this remark, $dsp(h,\xi) = p(\hat{s} \in SP_A(h,\xi) \mid \hat{s} \notin I_A(\xi))$ is called *degree of support* of h relative to ξ. It is a value between 0 and 1 that represents quantitatively the support that h is true in the light of the given knowledge. Clearly, $dsp(h,\xi) = 1$ means that h is certainly true, while $dsp(h,\xi) = 0$ means that h is not supported (but h may still be true). Note that degree of support is equivalent to the notion of (normalized) *belief* in the Dempster-Shafer theory of evidence [29, 32]. It can also be interpreted as the probability of the provability of h [23, 31].

A second way of judging the hypothesis h is to look at the corresponding conditional probability that the true scenario $\hat{s}$ is not in $RF_A(h,\xi)$. It represents the probability that $\neg h$ can not be inferred from the knowledge base. In such a case, h remains *possible*. Therefore, the conditional probability $dps(h,\xi) = p(\hat{s} \notin RF_A(h,\xi) \mid \hat{s} \notin I_A(\xi)) = 1 - dsp(\neg h,\xi)$ is called *degree of possibility* of h relative to ξ. It is a value between 0 and 1 that represents quantitatively the possibility that h is true in the light of the given knowledge. Clearly, $dps(h,\xi) = 1$ means that h is completely possible (there are no counter-arguments against h), while $dps(h,\xi) = 0$ means that h is false. Degree of possibility is equivalent to the notion of *plausibility* in the Dempster-Shafer theory. We have $dsp(h,\xi) \leq dps(h,\xi)$ for all $h \in \mathcal{L}_{A\cup P}$ and $\xi \in \mathcal{L}_{A\cup P}$. Note that the particular case of $dsp(h,\xi) = 0$ and $dps(h,\xi) = 1$ represents total ignorance over h.

4 Algorithms and Implementation

The existing computational methods for PAS belong to two different categories, depending whether the goal is a pure quantitative or a hybrid (qualitative and quantitative) analysis. The methods in the first category are based on the idea of transforming the given PAS into a corresponding family of Dempster-Shafer belief functions, and then to apply existing techniques such as *local computation* in join trees [20, 30]. Pure numerical methods are generally more efficient, but many problems remain intractable. However, a new concept of *incomplete belief functions* leads to convenient and efficient approximation algorithms (cf. forthcoming publication).

In the second category, the process consists of two succeeding steps. First, the sets of supporting and refuting arguments are derived from the given knowledge base through a resolution-based process of *variable eliminations* [15]. In a second step, the corresponding numerical degrees of support and possibility are then computed by methods originated from corresponding problems in reliability theory [3, 8, 19]. As the number of supporting and refuting arguments often grows exponentially with the size of the given knowledge base, an exact symbolic analysis is of course only feasible for relatively small problems. It is also not clear what to do if the sets of arguments get too large. However, in the same way as people naturally focus on the most relevant arguments, it is possible to measure the relevance of supporting and refuting arguments by corresponding cost or utility functions and to adapt the above-mentioned elimination procedure for a restricted search [14]. Moreover, the same idea leads to convenient anytime algorithms (cf. forthcoming publication).

A particular system called ABEL [5, 6] (http://www-iiuf.unifr.ch/tcs/abel) is an example of implementing probabilistic argumentation systems. It includes an appropriate modeling and query language, as well as corresponding inference mechanisms. Problems from a broad spectrum of application domains show that ABEL is very general and powerful [4]. The system has an open architecture, that permits the later inclusion of further deduction techniques.

5 Outlook

In this paper, we have only discussed *propositional* argumentation systems. However, note that the concept of argumentation systems can easily be generalized, for example to set constraint logic [16], linear gaussian systems [21], or even more general cases of assumption-based constraint logics. Furthermore, several publications are available or in progress in order to discuss the relation of argumentation systems to other concepts, such as valuation algebras [17], model-based diagnostic [2], default logic (cf. forthcoming thesis), abduction [10], or theorist [9].

References

1. L. Amgoud and C. Cayrol. On the acceptability of arguments in preference-based argumentation. In G. F. Cooper and S. Moral, editors, *Proc. of the 14th Conf. on Uncertainty in Artif. Intell. (UAI-98)*, pages 1–7, San Francisco, 1998. Morgan Kaufmann.
2. B. Anrig. *Probabilistic Model-Based Diagnostics.* PhD thesis, University of Fribourg, Institute of Informatics, 2000.

3. B. Anrig and F. Beichelt. Disjoint sum forms in reliability theory. *South African J. on Operations Research ORION*, 16(1), 2001.
4. B. Anrig, R. Bissig, R. Haenni, J. Kohlas, and N. Lehmann. Probabilistic argumentation systems: Introduction to assumption-based modeling with ABEL. Technical Report 99-1, Institute of Informatics, University of Fribourg, 1999.
5. B. Anrig, R. Haenni, J. Kohlas, and N. Lehmann. Assumption-based modeling using ABEL. In D. Gabbay, R. Kruse, A. Nonnengart, and H. J. Ohlbach, editors, *First Internat. Joint Conf. on Qualitative and Quantitative Practical Reasoning; ECSQARU–FAPR'97*. Lecture Notes in Artif. Intell., Springer, 1997.
6. B. Anrig, R. Haenni, and N. Lehmann. ABEL – a new language for assumption-based evidential reasoning under uncertainty. Technical Report 97–01, University of Fribourg, Institute of Informatics, 1997.
7. K. B. and P. E. Lehner. Assumptions, beliefs and probabilities. *Artificial Intelligence*, 41(1):65–77, 1989.
8. R. Bertschy and P. A. Monney. A generalization of the algorithm of Heidtmann to non-monotone formulas. *Journal of Computational and Applied Mathematics*, 76:55–76, 1996.
9. D. Berzati and B. Anrig. Propositional argumentation systems vs. Theorist. Technical Report 01–04, Department of Informatics, University of Fribourg, 2001.
10. D. Berzati, R. Haenni, and J. Kohlas. Probabilistic argumentation systems and abduction. In C. Baral and M. Truszczynski, editors, *Proc. of the 8th Internat. Workshop on Non-Monotonic Reasoning, Breckenridge Colorado*, 2000.
11. P. Besnard and A. Hunter. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128:203–235, 2001.
12. J. de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28:127–162, 1986.
13. P. M. Dung. On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2):321–357, 1995.
14. R. Haenni. Cost-bounded argumentation. *International Journal of Approximate Reasoning*, 26(2):101–127, 2001.
15. R. Haenni, J. Kohlas, and N. Lehmann. Probabilistic argumentation systems. In J. Kohlas and S. Moral, editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems, Volume 5: Algorithms for Uncertainty and Defeasible Reasoning*. Kluwer, Dordrecht, 2000.
16. R. Haenni and N. Lehmann. Building argumentation systems on set constraint logic. In B. Bouchon-Meunier, R. R. Yager, and L. A. Zadeh, editors, *Information, Uncertainty and Fusion*, pages 393–406. Kluwer Academic Publishers, 2000.
17. J. Kohlas. Valuation algebras: Generic architecture for reasoning. draft, 2001.
18. J. Kohlas, B. Anrig, R. Haenni, and P. Monney. Model-based diagnostics and probabilistic assumption-based reasoning. *Artificial Intelligence*, 104:71–106, 1998.
19. J. Kohlas and P. A. Monney. *A Mathematical Theory of Hints. An Approach to the Dempster-Shafer Theory of Evidence*, volume 425 of *Lecture Notes in Economics and Mathematical Systems*. Springer, 1995.
20. N. Lehmann. *Argumentation Systems and Belief Functions*. PhD thesis, University of Fribourg, Switzerland, 2001.
21. P. Monney. Assumption-based reasoning with functional models. Habilitation Thesis, University of Fribourg, Seminar of Statistics, 2000.
22. S. Parsons, C. Sierra, and N. Jennings. Agents that reason and negotiate by arguing. *J. of Logic and Computation*, 8(3):261–292, 1998.
23. J. Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, 1988.

24. D. Poole. Probabilistic Horn abduction and Bayesian networks. *Artificial Intelligence*, 64:81–129, 1993.
25. D. Poole. The independent choice logic for modelling multiple agents under uncertainty. *Artificial Intelligence*, 94(1–2):7–56, 1997.
26. H. Prakken and G. Sartor. A dialectical model of assessing conflicting arguments in legal reasoning. *Artif. Intell. and Law*, 4:331–368, 1996.
27. H. Prakken and G. Vreeswijk. Logical systems for defeasible argumentation. In D. Gabbay, editor, *Handbook of Philosophical Logic, 2nd edition*, volume D7: Semi Classical Logics 7. Kluwer Academic Publishers, 2001.
28. G. M. Provan. A logic-based analysis of Dempster-Shafer theory. *International Journal of Approximate Reasoning*, 4:451–495, 1990.
29. G. Shafer. *The Mathematical Theory of Evidence.* Princeton University Press, 1976.
30. P. P. Shenoy. Binary join trees for computing marginals in the Shenoy-Shafer architecture. *International Journal of Approximate Reasoning*, 17(2–3):239–263, 1997.
31. P. Smets. Probability of provability and belief functions. In M. Clarke, R. K. R., and S. Moral, editors, *Proc. of the ECSQARU'93 conf.*, pages 332–340. Springer-Verlag, 1993.
32. P. Smets and R. Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–234, 1994.
33. G. Vreeswijk. Defeasible dialectics: A controversy-oriented approach towards defeasible argumentation. *J. of Logic and Computation*, 3:317–334, 1993.