Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence

Computing Preferences Based on Agents' Beliefs *

Jian Luo, Fuan Pu, Yulai Zhang, Guiming Luo
School of Software, Tsinghua University
Tsinghua National Laboratory for Information Science and Technology
Beijing 100084, China
{j-luo10@mails, pfa12@mails, zhangyl08@mails, gluo@mail}.tsinghua.edu.cn

Abstract

The knowledgebase uncertainty and the argument preferences are considered in this paper. The uncertainty is captured by weighted satisfiability degree, while a preference relation over arguments is derived by the beliefs of an agent.

Introduction

Argumentation is a reasoning model based on the construction and the evaluation of arguments (Chesñevar, Maguitman, and Loui 2000; Prakken and Vreeswijk 2002). In logical argumentation, arguments are usually constructed from a knowledgebase Δ, possibly including conflict information (Besnard and Hunter 2006; 2008). There may be some uncertain information in the knowledgebase. For example, *it will rain in New York tomorrow*. Most information is uncertain to some degree. So, *how to quantify uncertain information?* In order to decide on a pair wise basis whether one argument defeats another, preferences over arguments have been harnessed in argumentation theory. In (Simari and Loui 1992; Benferhat, Dubois, and Prade 1993; Prakken and Sartor 1997) different preference relations between arguments have been defined. A preference relation captures differences in arguments' strengths. *However, it is not always clear what the preferences mean or where these preferences come from?*

Whether an argument is believable depends not only on the intrinsic merits of the argument – of course, it needs to be based on plausible premises and must be sound – but also on the audience to which it is addressed. Hunter introduced the beliefs of an agent in argument judgement (Hunter 2004; Besnard and Hunter 2008). However, only certain information was considered in the knowledgebase. In this paper, we define *weighted satisfiability degree* to capture the uncertainty of knowledgebase. Based on the beliefs of an agent, a preference relation between arguments is defined. Finally, a preference-based argumentation framework can be derived.

*This work was supported by the Funds NSFC61171121 and the Science Foundation of Chinese Ministry of Education - China Mobile 2012.

Satisfiability Degree

Let $P = \{p_1, p_2, \dots, p_n\}$ be a finite and non-empty set of atoms, we use $\mathcal{L}_P$ to denote the set of all propositional formulae over P, formed from the logical connectives of $\vee, \wedge, \neg$ and $\rightarrow$. Let $p, q, r, \dots$ represent atoms, $\alpha, \beta, \gamma, \dots$ denote formulae, and $\Delta, \Phi, \Psi, \dots$ denote sets of finite formulae. The tautology is denoted by $\top$ and $\bot$ represents the contradiction or falsity. We start introducing the concepts of satisfiability degree by Luo et al. (Luo, Yin, and Hu 2009).

An assignment of truth values to the elements of a set $P = \{p_1, p_2, \dots, p_n\}$ is called interpretation relative to P. Let $\Omega_P = \{0, 1\}^n$ be the set of all 2^n different interpretations. Thus, any formula $\alpha \in \mathcal{L}_P$ and $\omega \in \Omega_P$, $\alpha(\omega) \in \{0, 1\}$. For any $\alpha \in \mathcal{L}_P$, a subset $\Omega_\alpha \subseteq \Omega_P$ is defined as below:

$$\Omega_\alpha = \{\omega \mid \alpha(\omega) = 1, \omega \in \Omega_P\}$$

Definition 1. Given a propositional formula set $\mathcal{L}_P$ and the global interpretation field Ω_P, the subset Ω_α is defined as above. Then, function $\wp : \mathcal{L}_P \rightarrow [0, 1]$ is called the satisfiability degree (s.d.) on Ω_P, if for $\forall\, \alpha \in \mathcal{L}_P$:

$$\wp(\alpha) = \frac{\mathrm{card}(\Omega_\alpha)}{\mathrm{card}(\Omega_P)} \tag{1}$$

where $\mathrm{card}(X)$ denotes the cardinality of the set X.

For more details about satisfiability degree, including algorithms and properties, please refer to (Luo, Yin, and Hu 2009; Luo, Luo, and Xia 2011). We write $\Phi \vdash \alpha$ to mean that the set of formulae Φ entails the formula α. The notion $\bigwedge \Phi$ will be used to denote the conjunction of all formulae in Φ, specially $\bigwedge \emptyset = \top$.

Definition 2. An argument is a pair $\langle \Phi, \alpha \rangle$ s.t. $\Phi \subseteq \Delta$ is a consistent, finite set of formulae, α is a formula such that $\Phi \vdash \alpha$, and no proper subset of Φ entails α.

The set of all arguments generated from Δ is denoted as $\mathscr{A}(\Delta)$. If $A = \langle \Phi, \alpha \rangle$ is an argument for α, we use the function $\mathrm{support}(A) = \Phi$ to denote the support of A and $\mathrm{claim}(A) = \alpha$ to denote the claim of A.

Preferences Analysis

The set of atoms contained in α is given by $\mathrm{Atoms}(\alpha)$, and the set of atoms used in Φ is computed by $\mathrm{Atoms}(\Phi) = \bigcup_{\alpha \in \Phi} \mathrm{Atoms}(\alpha)$. The weight vector ϖ of Δ represents the weight assignments for atoms in $\mathrm{Atoms}(\Delta)$. Consider a formula α as a whole with two interpretations true and false. Assign some weights to both interpretations as $\mathrm{weight}(1)$ and $\mathrm{weight}(0)$ representing the weights for the true interpretation and the false interpretation such that

$$\mathrm{weight}(1) + \mathrm{weight}(0) = 1$$

where $\mathrm{weight}(1), \mathrm{weight}(0) \in [0, 1]$.

Let $P = \{p_1, p_2, \ldots, p_n\}$ where each p_i has weighted interpretations $\mathrm{weight}^{\varpi}_{p_i}(1)$ and $\mathrm{weight}^{\varpi}_{p_i}(0)$ such that $\mathrm{weight}^{\varpi}_{p_i}(1) = \varpi(p_i)$, $\mathrm{weight}^{\varpi}_{p_i}(0) = 1 - \varpi(p_i)$.

Definition 3. Let Ω_P be the global field and ϖ a weight vector on P, the weight of $\omega \in \Omega_P$, denoted as $\mathrm{weight}^{\varpi}(\omega)$, is defined:

$$\mathrm{weight}^{\varpi}(\omega) = \prod_{i=1}^{n} [\omega_i \cdot \varpi(p_i) + (1 - \omega_i) \cdot (1 - \varpi(p_i))] \quad (2)$$

where $\omega = (\omega_1, \ldots, \omega_n)$, and $\omega_i \in \{0, 1\} (1 \le i \le n)$.

Proposition 1. *Given a weight vector ϖ on an atom set P and the global field Ω_P for interpreting $\mathcal{L}_P$, then*

$$\sum_{\omega \in \Omega_P} \mathrm{weight}^{\varpi}(\omega) = 1$$

Definition 4. Given a weight vector ϖ on P, and the global interpretation field Ω_P. The weighted satisfiability degree (w.s.d.) of $\alpha \in \mathcal{L}_P$ is defined as $\wp_{\varpi}(\alpha)$:

$$\wp_{\varpi}(\alpha) = \sum_{\omega \in \Omega_\alpha} \mathrm{weight}^{\varpi}(\omega) \quad (3)$$

The w.s.d. describes the satisfiable extent of any formula in $\mathcal{L}_P$ given ϖ. Specially, there is a weight vector ϖ_m such that $\varpi_m(p_i) = 1/2$ for any $p \in P$, then $\mathrm{weight}^{\varpi}(\omega) = 1/2^n$ ($\mathrm{card}(P) = n$) for any $\omega \in \Omega_P$. In this case, the w.s.d. of α is the same as the s.d. of α.

Proposition 2. *Given a weight vector ϖ on atom set P and any formulae $\alpha, \beta \in \mathcal{L}_P$, the w.s.d. satisfies:*

- $\wp_{\varpi}(\bot) = 0$, $\wp_{\varpi}(\top) = 1$, $\wp_{\varpi}(\alpha) + \wp_{\varpi}(\neg\alpha) = 1$
- $\wp_{\varpi}(\alpha \vee \beta) + \wp_{\varpi}(\alpha \wedge \beta) = \wp_{\varpi}(\alpha) + \wp_{\varpi}(\beta)$

Definition 5. Given a weight vector ϖ on atom set P, $\alpha, \beta \in \mathcal{L}_P$ and $\wp_{\varpi}(\beta) > 0$, the weighted conditional satisfiability degree (w.c.s.d.) of α given β is

$$\wp_{\varpi}(\alpha \mid \beta) = \frac{\wp_{\varpi}(\alpha \wedge \beta)}{\wp_{\varpi}(\beta)}$$

Proposition 3. *Let ϖ be a weight vector on atom set P. For any formulae $\alpha, \beta, \gamma \in \mathcal{L}_P$ and $\wp_{\varpi}(\gamma) > 0$, then*

- $\wp_{\varpi}(\alpha \mid \top) = \wp_{\varpi}(\alpha)$, $\wp_{\varpi}(\gamma \mid \gamma) = 1$
- $\wp_{\varpi}(\top \mid \gamma) = 1$, $\wp_{\varpi}(\bot \mid \gamma) = 0$
- $\wp_{\varpi}(\alpha \vee \beta \mid \gamma) + \wp_{\varpi}(\alpha \wedge \beta \mid \gamma) = \wp_{\varpi}(\alpha \mid \gamma) + \wp_{\varpi}(\beta \mid \gamma)$

Example 1. Let $P = \{p_1, p_2, p_3\}$, and then Ω_P contains the following 8 interpretations:

$$\omega_1 = (1,1,1), \omega_2 = (1,1,0), \omega_3 = (1,0,1), \omega_4 = (1,0,0),$$
$$\omega_5 = (0,1,1), \omega_6 = (0,1,0), \omega_7 = (0,0,1), \omega_8 = (0,0,0).$$

For $\alpha = (p_1 \vee p_2) \wedge \neg p_3$ and $\beta = p_1 \wedge p_2 \to p_3$, we have $\Omega_\alpha = \{\omega_2, \omega_4, \omega_6\}, \Omega_\beta = \Omega_P \setminus \{\omega_2\}$, and $\Omega_{\alpha\wedge\beta} = \{\omega_4, \omega_6\}$. Consider the weight vector $\varpi(p_1, p_2, p_3) = (1/3, 1/2, 3/4)$, the evaluating results are given as follows:

$$\wp_{\varpi}(\alpha) = 1/6,\ \wp_{\varpi}(\beta) = 23/24,\ \wp_{\varpi}(\beta \mid \alpha) = 3/4.$$

So, β is more satisfiable than α due to $\wp_{\varpi}(\beta) > \wp_{\varpi}(\alpha)$.

A beliefbase $\mathcal{B}$ on Δ is a set of formulae such that $\wp_{\varpi}(\bigwedge \mathcal{B}) > 0$. The *preference* defined below reflects how much an agent supports an argument under his/her beliefs.

Definition 6. Let $A \in \mathscr{A}(\Delta)$, ϖ on $\mathrm{Atoms}(\Delta)$, and $\mathcal{B}$ be a beliefbase. The preference for A under $\mathcal{B}$, denoted as $\kappa(A, \mathcal{B})$, is defined:

$$\kappa(A, \mathcal{B}) = \wp_{\varpi}(\bigwedge \mathrm{S}(A) \mid \bigwedge \mathcal{B}) \quad (4)$$

Example 2. Consider the argument $A_1 = \langle\{\neg p \wedge \neg q\}, \neg(p \wedge q)\rangle$ and the weight vector $\varpi(p, q) = (1/2, 2/3)$. Given the beliefbase $\mathcal{B} = \{p \vee q\}$, then the preference $\kappa(A_1, \mathcal{B}) = 1/9$.

Proposition 4. *Given a knowledgebase Δ, ϖ on $\mathrm{Atoms}(\Delta)$ and a beliefbase $\mathcal{B}$, a preference-based argumentation framework $(\mathcal{A}, \mathcal{R}, prefs)$ can be derived where: 1) $\mathcal{A} \subseteq \mathscr{A}(\Delta)$; 2) $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$ and $\forall (A_1, A_2) \in \mathcal{R}$, $\wp_{\varpi}(\bigwedge \mathrm{support}(A_2) \mid \mathrm{claim}(A_1)) = 0$; 3) $prefs \subseteq \mathcal{A} \times \mathcal{A}$ s.t. $\forall (A_1, A_2) \in prefs$ iff $\kappa(A_1, \mathcal{B}) > \kappa(A_2, \mathcal{B})$.*

Conclusions

In this paper, classical logic is used for generating arguments. A weight vector is given to represent the weight assignments for atoms in the uncertain knowledgebase. This vector is then used to induce the weighted satisfiability degree. Based on the beliefs of an agent, preference is proposed to measure how much an agent supports an argument. Finally, a preference-based argumentation framework can be derived from the knowledgebase given the beliefbase.

References

Benferhat, S.; Dubois, D.; and Prade, H. 1993. Argumentative in ference in uncertain and in consistent knowledge bases. In *Proceedings of UAI'93*, 411–419.

Besnard, P., and Hunter, A. 2006. Knowledgebase compilation for efficient logical argumentation. In *KR'06*, 123–133.

Besnard, P., and Hunter, A. 2008. Elements of argumentation. volume 47. MIT press Cambridge.

Chesñevar, C. I.; Maguitman, A. G.; and Loui, R. P. 2000. Logical models of argument. *ACM Comput. Surv.* 337–383.

Hunter, A. 2004. Making argumentation more believable. In *AAAI'04*, 269–274.

Luo, J.; Luo, G.; and Xia, M. 2011. An algorithm for satisfiability degree computation. In *IJCCI (ECTA-FCTA)'11*, 501–504.

Luo, G.; Yin, C.; and Hu, P. 2009. An algorithm for calculating the satisfiability degree. In *Proceedings of FSKD (2)'09*, volume 7, 322–326.

Prakken, H., and Sartor, G. 1997. Argument-based extended logic programming with defeasible priorities. *J. of Applied Non-Classical Logics* 7:25–75.

Prakken, H., and Vreeswijk, G. 2002. Logics for defeasible argumentation. *Handbook of philosophical logic* 4(5):219–318.

Simari, G., and Loui, R. 1992. A mathematical treatment of defeasible reasoning and its implementation. *Artificial Intelligence J.* 53:125–157.