An Argumentation Based Framework for Defeasible and Qualitative Reasoning

Nikos I. Karacapilidis Dimitris Papadias Thomas F. Gordon

Artificial Intelligence Research Division
Institute for Applied Information Technology
GMD - German National Research Center for Computer Science
Schloss Birlinghoven, 53754 Sankt Augustin, Germany
e-mail: {karacapilidis, papadias, gordon}@gmd.de

Abstract. Multiagent settings are usually characterized by numerous goals, diverse opinions and conflicts of interest. In order to reach understanding and achieve cooperation, agents need a means of expressing their individual arguments which may contain explanations, justifications or any other kind of information. Furthermore, existing information may usually be incomplete, inconsistent and expressed in qualitative terms. In this paper, we present an argumentation-based framework that supports defeasible and qualitative reasoning in such environments. An interval-based qualitative value logic is applied, together with an inference mechanism in order to refine agents' knowledge, check consistency and, eventually, conclude the issue. The model is currently under development in Java, the aim being to deploy it on the World Wide Web.

1 Introduction

The argumentation component of the interaction between agents seems to attract the interest of researchers from various well-established areas, such as Distributed AI, nonmonotonic reasoning, decision theory, legal reasoning, as well as linguistics, philosophy, psychology and cognitive science. Stemming from legal reasoning procedures, argumentation has become an appropriate means of interaction in order to handle problematic instances, like normative conflicts and nonmonotonicity. Our goal is the development of a computational model of argumentation, by which groups of natural or artificial agents can express their claims and judgements, aiming at informing or convincing, depending on the kind of interaction. Such a model has to support rational, effective and fair decision making in ill-defined cases, where information is limited and/or not precisely known, and conflicts among agents are common. In this paper, we present a framework for defeasible argumentation and negotiation under the above conditions. It can be used in any kind of group decision making processes, and is able to handle inconsistent, qualitative and incomplete information in cases where one has to weigh reasons for and against the selection of a certain course of action.

We view such a framework as only a part of a mediating system for collaborative decision making and problem solving (see [14] for an extensive discussion on it). More specifically, the overall system also consists of a *Logic Layer*, where the notions of necessary *consequence* and *contradiction* are defined, a *Speech Act Layer*, where the space of possible kinds of actions an agent may perform is defined, and a *Protocol Layer*, where norms and rules about duties and rights of agents are specified (see also [8] and [19] for discussions on two slightly different approaches).

Formal models of argumentation have been built on various logics (see for example: [7] reconstructing Rescher's theory of formal disputation [21]; [18] based on Reiter's default logic [20]; and [12] using Geffner and Pearl's concepts of conditional entailment [11]). Whether it makes sense to use nonmonotonic, inductive or analogical logics at the logic layer is extensively discussed in [19]. In this paper, we shall not strictly specify the logic we intend for the system. The formalization of the other (more higher) layers should not assume any particular choice on the underlying logic. Similar models of defeasible argumentation have also left the subject unspecified (e.g., [26]).

The next section introduces the concepts involved in our argumentation framework and defines preference relations among competing statements. The argumentation concepts at this level result in a non-monotonic formalism founded on argumentation principles. Section 3 illustrates the procedure of concluding an issue together with the associated inference mechanisms that refine agents' knowledge and check consistency. Related and future work is discussed in Section 4.

2 The Argumentation Based Framework

2.1 Elements

Our terminology is based on that of issue-based information systems (first introduced in [15]; used in the IBIS system [27]). The argumentation elements presented below are the position, issue and argument (pro and contra)[1].

Positions are considered to be the basic objects in our framework. Any kind of data an agent wants to assert during an interaction can be used in order to represent a position. These data may have been brought up to declare alternative solutions, justify a claim, advocate the selection of a specific course of action, or avert the agents' interest from it. A position can be (or become) true or false, important or irrelevant for the corresponding problem, and may finally become acceptable or not.

Issues correspond to decisions to be made, or goals to be achieved. They consist of a set of alternative positions and a set of constraints that hold among them. An issue can be interpreted as which alternative position to prefer, if any. We don't allow more than one alternative position of an issue to be selected. At any stage of the argumentation process, an issue may be either inconsistent (due to inconsistency in the associated set of constraints), able to recommend a

[1] Due to space limitations, the formal definitions of the argumentation framework elements have been omitted in this version of the paper

solution (position) for its conclusion, or not. In fact, the last case indicates that none of the alternative positions of the issue is recommended.

Arguments are assertions about the positions which speak for or against them (multiple meanings of the term *argument* are discussed in [19]; various application areas are presented in [2]). An argument links together two positions of different issues. Agents can put forward arguments to convince their opponents or to settle an issue via a formal decision procedure. We distinguish between supporting arguments (pro) and counterarguments (con). Besides, we assume that arguments are refutable (see also [5], [6]), and two conflicting arguments can simultaneously be applied.

Various notions of an argument have been suggested in the literature. In [9], an argument consists of a *support base*, that may contain formulas which speak for or against a certain position, and a *conclusion.* A similar notion of argument has been given in [3]. All the above notions are extensions of the one proposed in [23]. Our definition differs from the above in that it does not presume that the support base of an argument is minimal. That is, a strict subset of the support base of an argument can be the support base of another argument with the same conclusion.

In most cases, we have to compare arguments in favor of the same conclusion, as well as, arguments leading to contradictory conclusions. The subjects of priority relationships and preference orders between arguments have been handled through quantitative approaches. For example, [16] and [22] have used the concepts of *penalty logic* (cost of not taking a premise into account) and *confidence factors*, respectively. Unfortunately, well defined utility and probability functions regarding properties or attributes of alternative positions (used, for instance, in traditional OR approaches), as well as complete ordering of these properties are usually absent. On the other hand, non-monotonic formalisms have defined preference relations on subtheories in various ways (see for example the concepts of *preferred subtheories* [5], *conditional entailment* [11] and *prioritized syntax-based entailment* [4]).

An argumentation system is usually defined given a set of arguments equipped with a binary relation holding among them. Since decision making and problem solving are evolving and theory-construction procedures, and also due to the problems mentioned above, an argumentation framework should allow for "weak" commitments on such relations. It is up to the agents to strengthen eventual weak relations, by providing the appropriate argumentation, and achieve the desired results.

2.2 Preference relations

Argumentation can be viewed as a special form of logic programming. As shown in [10], an argumentation system may be considered as consisting of two parts: an argument generation unit and an argument processing unit. The second one is a logic program consisting of the clauses: (i) $\mathtt{acc(x)} \leftarrow \neg\mathtt{defeat(x)}$ and (ii) $\mathtt{defeat(x)} \leftarrow \mathtt{attack(y,x)}, \mathtt{acc(y)}$, where (i) means that an argument is acceptable if it is not defeated, and (ii) means that an argument is defeated if it is attacked by an acceptable argument.

A first sketch of an appropriate underlying logic for such an argumentation framework, namely *Qualitative Value Logic*, was first proposed in [8]. This logic aims at relieving the users of the necessity of specification of exact cost values on subtheories, while offers them the possibility to reason about preferences (see also [13] exploring it together with abilities of constraint satisfaction formalisms). In this paper, we enhance this logic, in order to address the following problems that usually appear in real decision making instances: (i) A complete preference ordering among statements is not always attainable. There may be some formal properties such as transitivity and non-circularity, but still a partial ordering is what we are able to achieve. (ii) There is not always complete information for each alternative position of an issue regarding the attributes asserted by the arguments. In other words, the union of the support bases of each alternative position in an issue is not common. For instance, in order to conclude an issue with two alternative positions p_1 and p_2, we may know that "p_1 has the attributes a and b", while "p_2 has the attributes a, c and d", but no information regarding the ordering of b, c and d has been given.

We need a means for comparing alternative positions in an issue, taking into account the argumentation which has been put forward by the agents. We compare positions according to the value of their importance for the agents taking place in the communication process. The *importance value* of a position may range from $-\infty$ to $+\infty$. However, in most real world decision making instances, asking agents to attach a numeric weight or a certainty factor to default rules seems rather utopian (see also [2]). We associate with each position an *interval of importance*. Positions with (only) positive or negative importance may consist of the support base of a supporting argument or a counterargument, respectively. Note that, by using intervals, the framework also allows for a certain value of importance to be asserted (when the first and last points of the interval coincide). In other words, with such an interval-based value logic, our framework covers the cases where importance is interpreted as having either a certain value or a range of consecutive values.

A *preference relation* is a formula that relates two arguments[2]. Preference relations provide a qualitative way to weigh reasons for and against the selection of a certain course of action. Note that preference relations are also considered to be defeasible and subject to debate. We introduce the following preference relations between arguments: *SMP (strongly more preferable), SLP (strongly less preferable), EQP (equally preferable), WMP (weakly more preferable)* and *WLP (weakly less preferable).* The key difference between strong and weak preference is that in the former case the existing argumentation allows us to draw a distinction considering the entire support bases of the arguments under consideration, where

[2] To be more precise, we should say the support bases of two arguments. As mentioned above, an argument links two positions of two different issues, a support base and a conclusion; our framework allows "actions" on arguments implicitly, i.e., through their support bases.

in the latter only some parts of them[3] The above relations form two pairs of inverse relations ($SMP(p, q) \equiv SLP(q, p)$ and $WMP(p, q) \equiv WLP(q, p)$). The disjunction of all the above relations, usually known as the *universal* relation, will be denoted as *UNK* (it can be interpreted as *unknown* in our framework, since any of them may hold).

In addition, preference relations may be asserted independently of the position they refer to, i.e., independently of the conclusions of the related arguments (e.g., distinguish between the argument stating that speed is more important than price for any type of car, and another one arguing the same, but only for a particular car model).

3 Concluding an issue

As discussed above, our argumentation framework allows for defeasible reasoning in agent communication; that is, further information can trigger another course of action to appear more preferable than what seems best at the moment. Agents can put forward new preference relations at any time. Whenever that happens, the model infers all consequences by computing the transitive closure of the associated preference relations. This procedure is similar to the one first proposed in [1]. Briefly, a new relation adds a constraint in the relevant issue, which may in turn introduce new constraints between other arguments through the transitivity rules that hold among these relations. These rules are summarized in Figure 1.

	SMP	**SLP**	**EQP**	**WMP**	**WLP**
SMP	SMP	UNK	SMP	SMP	UNK
SLP	UNK	SLP	SLP	UNK	SLP
EQP	SMP	SLP	EQP	WMP	WLP
WMP	SMP	UNK	WMP	WMP	UNK
WLP	UNK	SLP	WLP	UNK	WLP

Fig. 1. The transitivity rules for the preference relations.

3.1 Inference mechanism

A constraint satisfaction problem is implicitly deployed. The corresponding constraint graph is being formed as the communication evolves. Each issue is actually a complete sub-graph of it. Applying path consistency algorithms together

[3] Formal definitions of these relations have been also left out in this version of the paper. However, informally speaking, $WMP(p, q)$ should be interpreted as "p could be better than q (we doubt due to lack of the appropriate information), but q could never be better than p"; $SMP(p, q)$ should be interpreted as "p is always better than q".

with the transitivity rules shown in Figure 1, we refine the agents' knowledge about the preference relations and check consistency. Initially (with no constraints yet asserted), all pairs of positions in each issue "receive" the *UNK* relation. When a new preference relation is brought up, all consequences are computed (through the computation of the transitive closure of the associated preference relations, for more details see [1]). Let $\mathtt{R}(p_k, p_i) \otimes \mathtt{R}(p_i, p_j)$ denote the composition function for lists of preference relations holding among (p_k, p_i) and (p_i, p_j), and *CheckQueue* denote a queue of "tests" to be done. The path consistency algorithm is as follows:

```
for each pair of positions (p_i, p_j)
       push (p_i, p_j) to CheckQueue
end-for
while not-empty-CheckQueue
       pop-CheckQueue (p_i, p_j);
              for each position p_k such that p_k ≠ p_i and p_k ≠ p_j
                     NewRel(p_k, p_j) = R(p_k, p_j) ∩ (R(p_k, p_i) ⊗ R(p_i, p_j))
                     if NewRel(p_k, p_j) = ∅
                            then return inconsistency;
                            else if    NewRel(p_k, p_j) ⊂ R(p_k, p_j)
                                       and not-in-CheckQueue(p_k, p_j)
                                              then push-CheckQueue(p_k, p_j);
                     NewRel(p_i, p_k) = R(p_i, p_k) ∩ (R(p_i, p_j) ⊗ R(p_j, p_k))
                     if NewRel(p_i, p_k) = ∅
                            then return inconsistency;
                            else if    NewRel(p_i, p_k) ⊂ R(p_i, p_k)
                                       and not-in-CheckQueue(p_i, p_k)
                                              then push-CheckQueue(p_i, p_k);
              end-for
end-while
```

Finally, what is needed is a classification of the positions in each issue. Aggregating the final set of preference relations, i.e., the one resulting after the application of the path consistency algorithm, we can eventually draw some comparisons among positions, and assign a *qualification label* to each of them. We allow for the labels *SBTA* (strongly better than all), *SBTS* (strongly better than some), *WBTA* (weakly better than all), *WBTS* (weakly better than some), *ETA* (equal to all), *SWTS* (strongly worse than some), *SWTA* (strongly worst than all), *WWTS* (weakly worse than some), and *WWTA* (weakly worst than all). Similar ideas are discussed in [9], based on a classification of arguments to those with an *empty support*, those which are not *rebutted*, and those which are not *undercut* by another argument.

3.2 Examples

In this section we present two examples in order to demonstrate the features of our framework. First consider the following: the goal at a part of a discussion is to find a constructor for a part of a car engine. Assume the following three alternatives: *C1, C2* and *C3*. The asserted attributes concern the quality, service, delivery time and cost that each of the alternatives provide. Figure 2 illustrates the discussion graph: positions are denoted with ellipses, issues with rectangles and arguments with straight lines (counterarguments are distinguished with a small horizontal line crossing the diagonal ones). The shadowed position of each issue is the system-recommended one. Figure 3a summarizes the existing knowledge about the attributes that each alternative has (or not). Assume also that the following preference relations have been asserted so far:

`SMP(fair cost, good quality)`, and
`SLP(meet due date, fair cost)`.

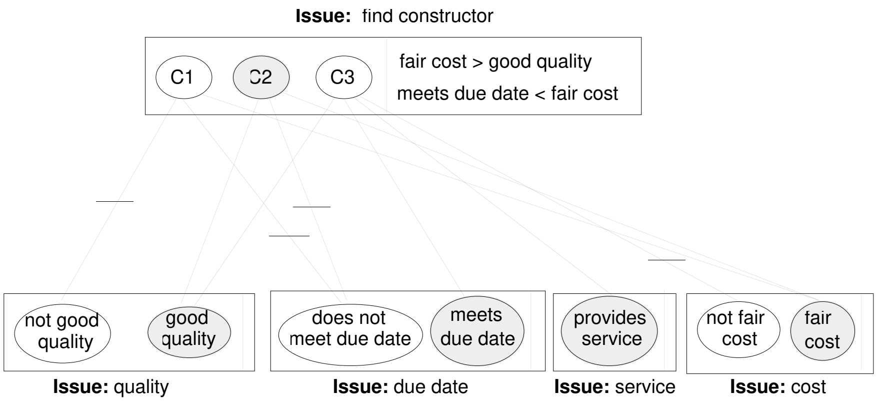


Fig. 2. Utility example.

As shown, there is not a complete linking between each alternative of an issue and every asserted attribute. For instance, there is no argumentation at the moment about the service provided for the *C1* and *C2* alternatives. The preference relations between the alternative constructors, resulting after the application of the inference mechanism illustrated above, are given in Figure 3b. One can see that *C2* is the recommended solution of the "find constructor" issue (*WBTA(C2)*). *C1* is strictly less preferable than each of *C2, C3*, i.e., *SLP(C1, C2)* and *SLP(C1, C3)*, while *C2* is weakly more preferable than *C3*, i.e., *(WMP(C2, C3)*.

The second example is a slightly modified version of the classical Tweety one. The example demonstrates how our framework can handle *specificity*. The existing argumentation, which has been brought up by two interacting agents (N and T are their names), is given in Figure 4, together with the corresponding discussion graph.

Experiments with examples from diverse types of communication indicate that, according to their content, we can distinguish between two types of issues:

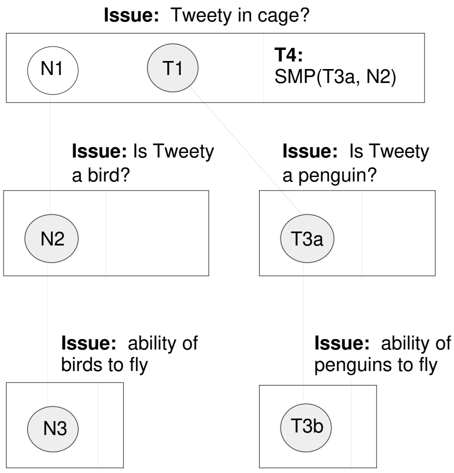
	good quality	*meets due date*	*service provided*	*fair cost*
C 1	no	no	??	yes
C 2	yes	no	??	yes
C 3	yes	yes	yes	no

(a) data

	C1	**C2**	**C3**
C 1	EQP	SLP	SLP
C 2	SMP	EQP	WMP
C 3	SMP	WLP	EQP

(b) results

Fig. 3. Utility example data and results.



N1: Put Tweety in a cage.
T1: No, why?
N2: Because Tweety is a bird.
T2: So what?
N3: Birds fly.
T3a: But Tweety is a penguin;
T3b: Penguins don't fly.
N4: So what?
T4: My argument is stronger than yours *(specificity)*.
N5: OK. So do we put it in a cage or not?
T5: Let Tweety be free *(equal to T1)*

Fig. 4. Tweety example.

(i) *boolean* issues, in which only a position and its negation can be included. However, it is not obligatory that both of them will be included, since there is not always an automatic insertion of the position $\neg p$ after the assertion of p (consider the issues "Tweety in cage?" and "Is Tweety a bird?" in the second example, and "quality"[4], "due date", etc. in the first example); (ii) Selection issues, in which alternative positions can be included and no alternative is the negation of another one (consider the issue "find constructor" in the first example).

4 Discussion

Approaches to reasoning in the presence of inconsistency can be distinguished to these based on the notion of *maximal consistent subbase*, where coherence is restored by selecting consistent subbases, and those based on *argumentation principles*, where inconsistency is "accepted" by providing arguments for each

[4] The reader should interpret "quality" as an abbreviation title for the issue "what kind of quality does Constructor *Cx* provides?", while *x=1,2,3*. In fact, our model retains three "copies" of this issue, one for each Constructor.

conclusion [9]. Non-monotonic inference relations are basically used in the first category, defined by methods of generating "preferred" belief subbases and classical inference [5]. The second category, inspired from work done on AI and Law, views the validity and priority of belief bases as subject to debate. The role of agents here is to construct theories, that are robust enough to defend a statement. Throughout this paper, we have taken into account belief bases that have been explicitly stated by an agent.

Among related, well-tried concepts and theories that have addressed the problems of practical and substantial reasoning, we only mention here Toulmin's early *second theory of argumentation* that considers logic as generalized jurisprudence [25], Pollock's OSCAR model of defeasible reasoning [17], Rescher's theory of *formal disputation* [21], Sycara's PERSUADER model of *goal conflict resolution* [24], and the IBIS *rhetorical method* developed at MCC [27]. These approaches have been attempting to account for how humans combine deductive, defeasible, inductive and probabilistic reasoning. The main target of our approach is to jointly exploit the linguistic model of argumentation, as it has been evolved among logicians and computer scientists, with quantitative models coming from disciplines like Game Theory and Operations Research.

The model presented is currently under development in Java. We experiment with various user interfaces on the World Wide Web, the target being a high quality visualization of the structure of the communication as it evolves, and the easy retrieval and contribution of information.

Acknowledgements: The authors thank Gerhard Brewka, David Kerr, Kia Toubekis and Hans Voss for helpful discussions on various topics of this paper. This work was partially funded by the European Commission (DG XIII), under the GEOMED project. Dimitris Papadias is financed by the Commission of the European Communities through the ERCIM Fellowship Programme.

References

1. Allen, J.F.: Maintaining Knowledge about Temporal Intervals. *Communications of the ACM* 26 (11), 1983, pp. 832-843.
2. Bench-Capon, Tr.: Argument in Artificial Intelligence and Law. In J.C. Hage, T.J.M. Bench-Capon, M.J. Cohen and H.J. van den Herik (eds.), *Legal knowledge based systems - Telecommunication and AI & Law*, Koninklijke Vermande BV, Lelystad, 1995, pp. 5-14.
3. Benferhat, S., Dubois, D., Prade, H.: How to infer from inconsistent beliefs without revising? In *Proceedings of the 14th IJCAI*, Montreal, 1995, pp. 1449-1455.
4. Benferhat, S., Cayrol, C., Dubois, D., Lang, J., Prade, H.: Inconsistency Management and Prioritized Syntax-Based Entailment. In *Proceedings of the 13th IJCAI*, Chambery, 1993, pp. 640-645.
5. Brewka, G.: Preferred Subtheories: An extended logical framework for default reasoning. In *Proceedings of the 11th IJCAI*, Detroit, 1989, pp. 1043-1048.
6. Brewka, G.: Reasoning about Priorities in Default Logic. In *Proceedings of the 12th AAAI*, Seattle, 1994, pp. 940-945.

7. Brewka, G.: A Reconstruction of Rescher's Theory of Formal Disputation Based on Default Logic. In *Working Notes of the 12th AAAI Workshop on Computational Dialectics*, Seattle, 1994, pp. 15-27.
8. Brewka, G., Gordon, T.: How to Buy a Porsche: An Approach to Defeasible Decision Making. In *Working Notes of the 12th AAAI Workshop on Computational Dialectics*, Seattle, 1994, pp. 28-38.
9. Cayrol, C.: On the Relation between Argumentation and Non-monotonic Coherence-Based Entailment. In *Proceedings of the 14th IJCAI*, Montreal, 1995, pp. 1443-1448.
10. Dung, P.M.: On the acceptability of arguments and its fundamental role in non-monotonic reasoning and logic programming. In *Proceedings of the 13th IJCAI*, Chambery, 1993, pp. 852-857.
11. Geffner, H., Pearl, J.: Conditional Entailment: Bridging two Approaches to Default Reasoning. *Artificial Intelligence* 53 (2-3), 1992, pp. 209-244.
12. Gordon, T.: The Pleadings Game: An Exercise in Computational Dialectics. *Artificial Intelligence and Law* 2(4), 1994, pp. 239-292.
13. Karacapilidis, N.I.: Planning under Uncertainty: A Qualitative Approach. In C. Pinto-Ferreira and N.J. Mamede (eds.), *Progress in Artificial Intelligence*, Lecture Notes in Artificial Intelligence 990, Springer-Verlag, 1995, pp. 285-296.
14. Karacapilidis, N.I., Gordon, T.: Dialectical Planning. In *Proceedings of the 14th IJCAI Workshop on Intelligent Manufacturing Systems*, Montreal, 1995, pp. 239-250.
15. Kunz, W., Rittel, H.W.J.: Issues as Elements of Information Systems. Working Paper 131, Universität Stuttgart, Institut für Grundlagen der Plannung, 1970.
16. Pinkas, G.: Propositional Non-Monotonic Reasoning and Inconsistency in Symmetric Neural Networks. In *Proceedings of the 12th IJCAI*, Sydney, 1991, pp. 525-530.
17. Pollock, J.: Defeasible Reasoning. *Cognitive Science* 11, 1988, pp. 481-518.
18. Prakken, H.: Logical tools for modelling legal argument. Ph.D. Dissertation, Free University of Amsterdam, 1993.
19. Prakken, H.: From Logic to Dialectics in Legal Argument. In *Proceedings of the 5th International Conference on AI and Law*, ACM Press, 1995, pp. 165-174.
20. Reiter, R.: A Logic for Default Reasoning. *Artificial Intelligence* 13, 1980, pp. 81-132.
21. Rescher, N.: *Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge*. State University of New York Press, Albany, 1977.
22. Sian, S.S.: Adaptation based on cooperative learning in multi-agent systems. In Y. Demazeau and J.P. Müller (eds.), *Decentralized AI 2*, Elsevier Science Publishers B.V., 1991, pp. 257-272.
23. Simari, G.R., Loui, R.P.: A Mathematical Treatment of Defeasible Reasoning and its Implementation. *Artificial Intelligence* 53 (2-3), 1992, pp. 125-157.
24. Sycara, K.: Resolving Goal Conflicts via Negotiation. In *Proceedings of the 7th AAAI*, Saint Paul, Minnesota, 1988, pp. 245-250.
25. Toulmin, S.E.: *The Uses of Argument*. Cambridge University Press, 1958.
26. Vreeswijk, G.: Studies in Defeasible Argumentation. Ph.D. Dissertation, Free University of Amsterdam, 1993.
27. Yakemovic, K.C.B., Conklin, E.J.: Report on a Development Project Use of an Issue-Based Information System. In F. Halasz (ed.), *Proceedings of CSCW 90*, LA, 1990, pp. 105-118.

This article was processed using the LaTeX macro package with LLNCS style