Relatório Técnico

Núcleo de Computação Eletrônica

Modelling the Cooperative Information Filtering Problem[1]

Sueli Bandeira Teixeira Mendes[2]
Claudia Lage Rebello da Motta[3]

(1) Partially supported by CNPq - National Council fpr Scientific Technological Development;
(2) Consultant of Soft Consultoria; (3) Researcher at Núcleo de Computação Eletrônica (UFRJ).

NCE - 12/99

Universidade Federal do Rio de Janeiro

Modelling the Cooperative Information Filtering Problem[1]

Sueli Bandeira Teixeira Mendes[2]

Claudia Lage Rebello da Motta[3]

ABSTRACT

In this paper, we present an original approach for modelling the cooperative information filtering problem using Logic of Information Flow and Situation Theory. We describe the fundamental concepts of these theories and discuss crucial issues that arise during the exposition, such as defining Situation Theory as a Heyting algebra and showing that in our application this theory is well-founded.

Keywords

Logic of Information Flow, Situation Theory, information flow, information filtering, social filtering, cooperative information filtering, filters, model, distributive lattice, Heyting algebra's, well-founded theories, replacement systems.

INTRODUCTION

In this paper, although we have two original approaches - the application and the formalization - we choose to focus on the latter because it may be a powerful tool for many researchers that deal with information, in our case, with cooperative information filtering.

There are many information filtering systems [9, 10, 11, 13] (also called recommender systems)[12] but all of them use the results of the filtering individually, i.e., each user uses the results for his or her own purpose. In our system, a cooperative information filtering

---

[1] This research was partially supported by CNPq – National council for Scientific Technological Development.

[2] Sueli Bandeira Teixeira Mendes is a consultant of Soft Consultoria.

[3] Claudia Lage Rebello da Motta is a researcher at Núcleo de Computação Eletrônica of Universidade Federal do Rio de Janeiro

teamwork oriented, we use the results to improve the productivity of teamwork itself. Relevant and important information that is useful for the members of the team to achieve their common goal is shared. All members are aware that knowledge must be spread by all in order to homogenize it.

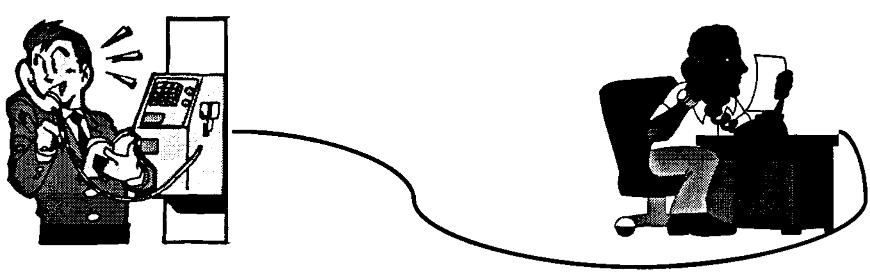
Because of the complexity of this problem, we feel the necessity to formalize it. We tried hard to find a tool to help us in this formalization and we found in Logic of Information Flow and in the Situation Theory powerful and efficient tools to do our modelling. After this, we could focus clearly on the real problem and work for which it is hard to find a solution. During this formalization, we found that many other information systems could be formalized as well use these models. In this paper we will give more details of the formalization itself and show the potentiality of this logic.

LOGIC OF INFORMATION FLOW

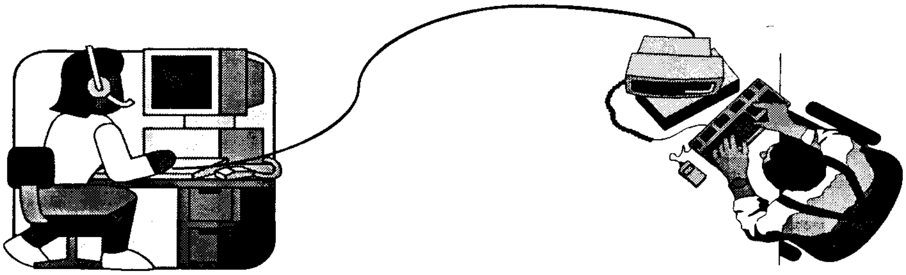
The Logic of Information Flow (LIF) was basically created by Jon Barwise [3, 5] based on the Situation Theory. The fundamental interest for those that study this logic is to offer a formal theory of information.

The Logic of Information Flow primitives are Sites and Channels: Sites are places or situations and channels are ways that connect the sites. Here we say that S denotes a set of sites and C denotes a set of channels. For instance, we can think of something physical where Sites are telephones, computers or "sites" of the Internet, and channels are telephone cables or optic fibbers, coaxial cables in a computer network, etc. However, in our model, we will work with something more abstract.

Sites and Channels as physical concepts:



Site Channel Site

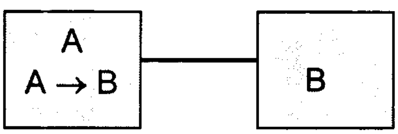


Site Channel Site

Sites and Channels as abstractions:

Classic Logic

$$\frac{A,\ A \rightarrow B}{B}$$



site channel site

In some cases, the set of channels may be included in S. This means that in this case channel C may be a site. It occurs when we model the Classic Logic where sites and channels are the same thing. However, in most cases, sites and channels may be disjunct sets, where $S \cap C = \emptyset$.

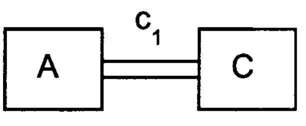
Why are sites and channels the same thing in Classic Logic? If we consider those sentences, (as in prepositional calculus) sites, - as Barwise did-, we have a set of truths called axiomas and theorems. The latter have been proved by using inference rules. These rules are channels because the information that is in one site is going to flow to another site, by the application of an inference rule. However, as we can express the inference rules in the same language as the language we use for the sites, in fact, C is identical to S.

Constraints

One of the fundamental relations between sites and channels is what we call constraint. It is used here as something that generalizes every kind of inference rule, making it possible that from one proposition we infer another. Constraints, in general, are regularities that, in a general sense, taking our world as an example, make it uniform, by means of physical or natural laws. Example: for all thermometers that work well, the height of the mercury will always be the same for a certain temperature. There are kinds of constraints that are conventions. In the Classic Logic the inference rules are the constraints.

In our case, a constraint gives the type of channel where the information flows, i.e., it is the filter that determines which information will passes from one site to another. In other words, it is the criterion that determines which kind of information flows into the channel. These criteria, that could be any one, are pre-defined by the user like rates, gradesand others. Example of constraint:

constraint



$c_1 : A \rightarrow C$

Where c_1 is the type of channel that carries the information from A to C.

MODELLING THE INFORMATION FILTERING PROBLEM

Some observations are welcome for better understanding of the modelling bellow. When we are modelling we need to choose a way that follows our intuition, where we can see the problem more clearly. The choice of the right primitives is essential to the success of any formalization.

☑ For us, sites are the set of user's profiles.

☑ By using our abstract point of view, profiles can be of an individual user or a set of users;

☑ The modelling is independent of the type of filter or site implemented;

☑ Using the formal language of the Situation Theory, we can describe sites, i.e., people or a set of people, as an information base or infobase.

There are two parts of the modelling:

❶ The Situation Theory where the facts are presented ;

❷ The Logic of Information Flow where the most general idea of Barwise is observed. It says that independently of the logic or theory, in the portion of the information flow, we always have the sites of information and channels where this information flows between the sites.

In particular, we see these sites as these infobases described in the Situation Theory's language, that is a user or a set of users ratings. In addition, we see the channels as filters. Which is exactly the idea of Barwise. It also explains why the information goes from one place to another. Finally, the constraint indicates the type of the channel.

Sites - Situations

The Situations is normally like a label of an infobase where the facts, i.e., everything we say that is true or false, are to be found in that situation. In the Situation A below, the facts are:

Situation A

<u>Fact1</u>

<u>Fact2</u>

<u>Fact3</u>

...

<u>Factn</u>

These facts are called infons.Using the Situation Theory, these facts are expressed in terms of information. The infons below are supported by the situation A, i.e., they are valid in this situation.

Situation A

(infon1) 1 ________

(infon2) 2 ________

(infon3) 3 ________

...

(infonn) n ________

The site there may be more than one situation, i.e., more than one infobase. We can simplify it and think of a situation as the classification of the documents or divide the situations by the evaluation criteria. For example:

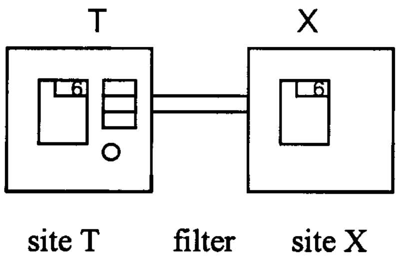
Situation A - all documents are classified by grades from 0 (zero) to 10 (ten);
Situation B - all documents are classified by faces ☺/😐/☹.
Situation C - all documents are classified by concepts from "A" to "E".
Situation D - all documents are classified by annotations like: "loveit" or "hateit".
Situation E - all documents are classified by graphic bars, the longer the better.

These situations may be compound on a site or they may be divided on several sites, for instance, one site for each rating. The facts that compound the situation are facts that indicate how people classify the documents. Later, people define the constraints.

Filters

In modelling, we think of criteria that are going to define the filters. We are not interested in the content of the documents; we want to know the constraints that define the filters. A change of model also changes the filter and the constraint.

Example: In a site T, we have documents and a place to store their evaluation. Suppose the criterion is a grade from zero to ten. We can have a constraint like "all documents, in site T, which are classified better than 5 (five), goes to site X.



Tapestry [9] uses annotations as "likeit" and "hateit" to classify their documents. In GroupLens [11], the classifications change depending on the implemented version: grades from 1 (one) - worst- to 6 (six) - best-, concepts ("A" - best and "E" - worst) and graphic bars (the longer, the better). In Firefly [13] they use grades associated with sentences that try to capture the feeling of the user. The scores go from 0 (zero) to 7 (seven), 0 to 3 - doesn't like, 4 - indifferent, and 5 to 7 - like it.

Time and Place

Both Barwise [5] and Devlin [7] thought of a special location to put the parameters in the infon: time and place. They are parameters in the relation to the Situation Theory that may or not be used. In our model, we may use these special parameters to determine the date of the document. In the parameter places. We may also register the source of a document. If the document comes from the World Wide Web, we can store the URL (Universal Resource Location), for example.

Polarity (0 or 1)

Polarity may be used to indicate if a document was or was not evaluated by the user. If nobody evaluated a document, it is going to be zero, and a dummy for those objects that were not instantiated.

Example of an infon: (<R (n_1,...,n_r, l,t,); i> where R is a n_r + 2-uple relation, n_1,...,n_r are arguments, l is a place and t denotes time. he symbol i of the infon denotes its polarity: i is either 0 or 1.

SITUATION THEORY

Now it is important to describe Situation Theory (ST) because is not a standard theory, i.e., it is not a Classic First-Order Theory. So it brings about several problems we will deal with in this paper.

For someone interested only in Information/Recommender Systems the discussion below presents no interest. It suffices for such a researcher to understand the basic LIF and ST. He/She will agree that the formal System is simple enough and compares favorably to other tentative formalization. But for someone that works also in Logic by itself, the discussion below, certainly presents an important approach to the formalization presented.

In his paper [3], Barwise shows that Logic of Information Flow can be a model for several Formal Systems, including: First-Order Theory Classic Logic, Intuitionistic Logic, Hoare's Logic and Dretske's Logic. We do not repeat here his arguments since they can be found in [3]. The first thing we must appoint is that following [8] we consider Situation Theory as it was revised in [2] where he admits several different kinds of "support" predicates " $\models$". Consequently, the reduction of a situation to a set of infons, and the identification of a basic infon with some entity consisting of a relation, an assignment and a polarity must be relativized to the relevant predicate " $\models$".

Another important issue raised by the ST is the question of well foundness of the Theory. Considering that in [3] an infon can be substituted for another infon within an infon and the same is true for situations, ST may incur not only in paradoxes as the well known Liar Paradox, as well as in an indefinite descent a_n, a_{n-1}, ...,a_0..., i.e., it can incur in the problem of a not well founded theory. In fact as we see from our application of LIF and ST to our Information Flow on the Web by using Filters, this problem does not arise since all parameters we use can only be replaced by primitive objects such as persons, dates, rates, etc.

will only consider a restrict ST that can be described as a distributive lattice and since this algebra is a complete distributive lattice it is in fact a Heyting algebra.

To emphasizes the difference between ST and classical first-order logic we can note that in the latter there is a clear mathematical notion of what a model of information associated with sentences is, and what means one content to entail another. In classical first-order logic contents are modelled by collections of models and a collection x entails a collection y iff $x \subseteq y$. In ST there is the need of providing a crisp model of either content or entailment. In Barwise's paper, the references above the answers, show that a ST turns out to be a distributive lattice.

To better understand ST as a set of infons it must be emphasized that there is a great difference between an infon, as a piece of information and a proposition. Here, when we use the term "information" it is in the sense of "possible information", i.e., whereas in the case of a proposition it can always be true or false. In the case of infons they can be truthful or not depending on whether there is a situation that supports it.

SET OF INFONS FORMS AN ALGEBRA

Following Barwise & Etchmendy [4] we will define an Infon Algebra $I = <\text{Sit}, I, \models, \Rightarrow>$, a non-empty collection of objects Sit (situations) and a distributive lattice $< I, \Rightarrow >$ with 0 (the meet of lattice) and 1 (the join of the lattice). The lattice has the collection of objects I, infons as its domain, and a relation $\models$ on Sit X I (Cartesian product) such that for all s (situation) and γ, τ (infons) the following conditions hold:

(a) If $s \models \gamma$ and $\gamma \Rightarrow \tau$ then $s \models \tau$

(b) $s \not\models \varnothing$ and $s \models 1$

(c) If Σ is a finite set of infons, $s \models \wedge\Sigma$ if and only if $s \models \gamma$ for each $y \in \Sigma$

(d) If Σ is a finite set of infons, then $s \models \vee\Sigma$ if and only if it supports some $\gamma \in \Sigma$

In our application, we can see the infons as set of ratings, i.e., we will now describe a sorted algebra that applies to our case. In this algebra, we have a set of atomic objects each of a type, and these objects include the atomic relation rating. The objects are the following:

i) a finite set of person $p_1, ..., p_i$ (intuitively they are the researchers of a group)

ii) a finite set of documents $do_1, ..., do_j$ (these include papers, videos and so on)

iii) a finite set of ratings $r_1, ..., r_k$

iv) a finite set of dates $da_1, ..., da_l$ (they mean the date of the documents)

v) a finite set of address $a_1,...,a_m$ (these can be URL's or any other source of documents)

The atomic relation Rating forms a structured object infon of the form:

<Rating (p_i, do_j, r_k, da_l, a_m), i>

Where $i \in \{1,0\}$, i.e., the polarity of the infon meaning a document $\mathbf{do_j}$ was rated by the person $\mathbf{p_i}$, if i = 1 or not if i = 0.

We say that an infon $\gamma \Rightarrow \tau$ (following Barwise & Etchmendy [4] we read "$\Rightarrow$ " as "involves") when the content of γ is practically the same as the content of τ, i.e., τ does not bring any new idea to the infon γ. This implies that the rated document referred by γ also includes the infon $\gamma \Rightarrow \tau$, i.e., this information is given to other elements of the group to prevent them from reading the document referred by τ.

In this algebra $\varnothing$ is the incoherent information, i.e., the same infon that has polarity 1 and 0 so that is intuitively true that no situation can hold it. 1 is the null document so that 1 is vacuously true for any situation. 0 and 1 where include here so that (I, $\Rightarrow$) can be defined as a distributive lattice. In our algebra (I, $\Rightarrow$) the notion of entailment (i.e., $\gamma \Rightarrow \tau$) is as we said above interpreted as the idea of the rating of a document $\mathbf{do_i}$ implying the rating of a document $\mathbf{do_j}$, since they can be considered as having the same content.

This operator $\Rightarrow$ is reflexive straightforwardly, i.e.:

i) For all σ, $\sigma \Rightarrow \sigma$.

It is transitive, i.e.:

ii) For all σ, τ and π, if $\sigma \Rightarrow \tau$ and $\tau \Rightarrow \pi$ then $\sigma \Rightarrow \pi$.

And it is also antisymmetric, i.e.:

iii) For all σ and τ if $\sigma \neq \tau$ either $\sigma \Rightarrow \tau$ or $\tau \Rightarrow \sigma$ fails.

The algebra (I, $\Rightarrow$) is complete since the conditions c and d above hold for all sets Σ of infons in our interpretation.

This algebra is clearly distributive, i.e.:

e) For all σ, τ and π, $\sigma \wedge (\tau \vee \pi) = (\sigma \wedge \tau) \vee (\sigma \wedge \pi)$.

Since this algebra is a complete distributive lattice, it may be called a Heyting algebra.

The algebra (I, $\Rightarrow$) in our interpretation is not a Boolean algebra since:

For all σ it is not true that, $\sigma \vee \neg\sigma = 1$ (this can be easily seen considering that 1 in our interpretation is the null document).

To show informally the well foundness of our algebra, we will use Aczel's definition of a replacement system [1]. To understand the notion of a replacement system we have to introduce the notions of a signature of an algebra and the notion of an ontology. The signature of an algebra is simply the arity of the operator of this algebra. For instance, the signature of the operator rating is 5 since rating is a n-upla where n = 5.

The notion of an ontology generalizes the notion of a signature. The signature defines the form of a structured object of a domain of objects. If we have a universe of objects, we call an ontology the *forms* of the objects of this universe. The *forms* of an ontology describe the structures of all structured objects of a universe. This means that to have an ontology we must have the form, i.e., the signature of all operators that form the structured objects of our universe of objects.

Following Aczel's paper [1] we can restrict the notion of an ontology taking all possible forms of the objects of a universe U from a fixed *class* X of possible signature of objects. With this we obtain, a *form system over* X. In this special case a replacement system is simply a form system over a class A where the class of the forms of the system coincide with A and an ontology is a form system over the universal class of all objects of the metatheory. In our case the objects of the metatheory are:

i) The atoms $p_1, \ldots, p_i$; $do_1, \ldots, do_j$; $r_1, \ldots, r_k$; $da_1, \ldots, da_l$; $a_1,\ldots,a_m$ and the atomic relation *rating*.

ii) The structured objects *infons*.

iii) The structured objects situations that are simply sets of infons.

So that in our case an ontology must comprise the forms of these kinds of objects.

Definition of a form system of a class X.

Definition 1 (A,C,.) is a form system over a class X if A is a class, C is an operator such that $C: A \rightarrow \text{pow } X$ and for each $a \in A$, if $\delta : Ca \rightarrow X$ then $\delta.a \in A$ and the following holds:

1) $C(\delta.a) = \{\delta\, x \mid x \in Ca\}$

2) $\delta.a = a$ if $ax = x$ for all $x \in Ca$

3) $\delta\,.\,(\delta.a) = (\delta' \circ \delta)\,.\,a$ if $\delta' : C\,(\delta.a) \rightarrow X$

In the Definition 1 if $\delta : Ca \rightarrow X$ then it can be used on the form <u>a</u> to obtain a form $\delta.a$ with set $\{\delta x \mid x \in Ca\}$ of components. The form $\delta.a$ should be understood to be the form obtained from a by replacing each component x of a by δx.

Using a simple example we may have the form:

Rating (p_i, do_j ,r_k, da_l, a_m) where p_i, do_j ,r_k, da_l, a_m are the parameters taking from X. In this case X is the class $P \cup Do \cup R \cup Da \cup A$ where

P	=	set of researchers
Do	=	set of documents
R	=	set of ratings
Da	=	set of dates
A	=	set of addresses

Then we obtain a new object from Rating (p_i, do_j ,r_k, da_l, a_m) by replacing each parameter x of the relation Rating by δx. The new form must be a form such that each parameter type must be replaced by an object of the same type as the type the parameter stands for.

In the case of the Second order relation $\Rightarrow$ the parameters are of infon type, i.e., if we have:

$a = \sigma \Rightarrow \tau$

then $Ca = \{ \sigma, \tau\}$ and

$\delta.a = \sigma' \Rightarrow \tau'$ where δ.a is the new form obtained by replacing σ by σ' and τ by τ'. Obviously, in these examples, the conditions of Definition 1 hold.

Definition 2. (A, C, .) is a replacement system if it is a form system over A.

The objects of A can be thought either as forms for objects or as the objects themselves. Thinking of them as the objects themselves, they are to be understood as structured entities that have components which are themselves objects of A.

In our example we have:

$A_1 = P \cup Do \cup R \cup Da \cup A \cup Rating$

(those are atomic objects)

and

$A_2 = \{\Rightarrow\} \cup \{ \sigma \mid \sigma \text{ is an infon}\}$

(these are the structured objects).

Finally A is:

$A_1 \cup A_2$

knowing that each infon has the form as we said above: Rating (p_i, do_j ,r_k, da_l, a_m).

when we try to generalize a signature to obtain an ontology we must be careful so that the class of parameters X comprises all the objects of our metatheory. In Aczel's paper we avoid any dependency on X by fixing X to be a universe U that includes all the objects of our metatheory. From this we have the definition:

Definition 3. An ontology is a form system over the universal class U.

From the definitions above we have following Aczel's paper the definition of a well-founded system.

Definition 4. Let $A = (A, C, .)$ be a replacement system, then A is well-founded if for every non-empty set $x \subseteq A$ there is $a \in x$ such that Ca is disjoint from x.

To show this informally, consider the following example, from our application:

In our case we have

$A_1 = \{ P, Do, R, Da, A, Rating\}$ i.e.

the atomic objects

$A_2 = \{\Rightarrow$, infons, situations$\}$

where situations = $[x \mid x$ is an infon$\}$

and

$A = A_1 \cup A_2$

Suppose we have a situation $s = \{ \sigma, \tau , \pi\}$ and σ = Rating $(p_i, do_j, r_k, da_l, a_m)$ so that σ is our $a \in A$ and $\underline{s}$ is our $x \subseteq A$. Then:

$Ca = \{ p_i, do_j, r_k, da_l, a_m\}$

and since $x = \{ \sigma, \tau , \pi\}$ it follows that

$Ca \cap x = \varnothing$.

We may show by cases that in our application the Situation Theory is well founded for each type of object of A . A is the universe of metatheoretic objects in our application.

CONCLUSION

In our paper, we use the Logic of Information Flow to model the cooperative information filtering problem. The novelty of our approach is the use of a formal system to model this problem where we do not have the pretension of presenting a complete formalization, i.e., a syntax and semantic of a formal system. Instead, we use Logic of Information Flow to provide a first step towards a complete formalization.

To model the cooperative information filtering problem we need to specify sites and channels. To specify both we borrow the fundamental concepts of the Situation Theory such as it was presented in [3] and the discussions about crucial issues that this theory arises.

As we were modelling, we noted that it was a novel application for this purpose, either for the modelling of the user's profile - the sites, or for the modelling of the filters that explain the information flow from one site to another - the channels and constraints.

The approach described here is in agreement with Barwise's theory. As the modelling is independent of the type of filter or site implemented, we may particularize it, formalizing other systems. This tool may aid many other researchers to design their own systems, abstracting and concentrating their attention on the main problems. Although, the formalization is not trivial, it may be a powerful and useful tool.

REFERENCES

1) Aczel, P. Replacement Systems and Axiomatization of Situation Theory, in *Situation Theory and Its Application*, Cooper, Mukai & Perry (eds), 1990, CSLI, Stanford, CSLI Publications, 3-31.

2) Barwise, J. Situation, Facts and True Propositions, in *the Situation in Logic*. CSLI Lectures Notes, n. 17, Stanford, CSLI Publications, 1989.

3) Barwise, J. Constraints, Channels and Flow of Information., in *Situation Theory and Its Applications*. Vol. 3. Aczel, Israel, Katagiri and Peters (eds), 1993, CSLI, Stanford University, 3-27.

4) Barwise, J. & Etchmendy, J. Information, Infons, and Inference, in *Situation Theory and Its Application*, Cooper, Mukai & Perry (eds), 1990, CSLI, Stanford University, 33-78.

5) Barwise, J. & Perry, J. Situations and Attitudes, Cambridge, MA: MIT Press, 1983.

6) Barwise, J., Gabby, D. & Hartonas,, C. On the Logic of Information Flow, in *Bulletin of IGPL*, vol 3, n.1, 7-50, 1995.

7) Devlin, K., Logic and Information, Cambridge: Cambridge University Press, 1991.

8) Fernando, T. On the Logic of Situation Theory, in *Situation Theory and Its Application*, Cooper, Mukai & Perry (eds), 1990, CSLI, Stanford University, 97-116.

9) Goldberg, D. Nicholas, D. OKI, B.M. and Terry, D. Using Collaborative Filtering to Weave and Information Tapestry. *Commun. ACM 35*, December 1992 12, 61-70.

10) Hill, W. & Terveen, L. Using Frequency-of-mention in Public Conversations for Social Filtering, in *Proceeding of CSCW'96*, (Cambridge MA, 1996), ACM Press, 106 -112.

11) Resnick, P. Neophytos, I. Mitesh, S. Bergstrom, P. & Riedl, J. GroupLens: An Open Architecture for Collaborative Filtering of Netnews, in *Proceedings of CSCW'94* (Chapel Hill, 1994), Addison-Wesley, 175-186.

12) Resnick, P. & Varian H.R. (guest eds.) Recommender Systems. *Commun. ACM* 40, 3 March 1997, (entire issue).

13) Shardanand, U. & Maes, P. Social Information Filtering: Algorithms for Automating "Word of Mouth"., in *Proceedings of CHI'95* (Denver CO, May 1995), ACM Press, 210-217.

14) Westerståhl, D. Parametric Types and Propositions in First-Order Situation Theory, in *Situation Theory and Its Application*, Cooper, Mukai & Perry (eds), 1990, CSLI, Stanford University, 193-230.

Sueli Mendes completed her Ph.D. at Stanford University in 1974. For many years was a researcher and a professor at Universidade Federal do Rio de Janeiro, Brasil. Her interest areas are Logic, Natural Language Processing and Information Filtering/Recommender Systems. E-mail: smendes@soft.com.br

Claudia Motta completed her Ph.D. at Universidade Federal do Rio de Janeiro, Brasil. Her interest areas are Computer Supported Cooperative Work, Information Filtering /Recommender Systems and Intelligent Agents. E-mail: claudiam@nce.ufrj.br