

## Luis Mateus Rocha

Los Alamos National Laboratory, MS B256

Los Alamos, NM 87545, USA

E-Mail: [rocha@lanl.gov](mailto:rocha@lanl.gov) or [rocha@santafe.edu](mailto:rocha@santafe.edu)

WWW: <http://www.c3.lanl.gov/~rocha>

---

# Adaptive Webs for Heterarchies with Diverse Communities of Users<sup>1</sup>

---

Paper submitted to *Global Brain Workshop, Brussels 2001*. LAUR 005173.

### Abstract.

We discuss the concept of Heterarchy, understood in the social sciences as a model of complex, adaptive human organizations capable of evolving in rapidly changing environments. We observe that heterarchies tend to lead to control hierarchies, because as complex adaptive systems that make use of tags, emergent levels of behavior and control are expected to occur. Given that tags play a fundamental role in heterarchies, we propose that they can make good use of adaptive webs as a communication fabric to manage and co-evolve the knowledge traded with their communities of members and users.

A recommendation system named *TalkMine* is then presented to advance adaptive web technology for heterarchies. *TalkMine* leads different information resources to learn new and adapt existing keywords to the categories recognized by its communities of users. It uses distributed artificial intelligence algorithms and is currently being implemented for the research library of the Los Alamos National Laboratory (<http://arp.lanl.gov>). *TalkMine* is based on the integration of evidence from users and several distributed knowledge networks using Evidence Sets, an extension of fuzzy sets. The interest of users is further fine-tuned by a human-machine conversation algorithm used for uncertainty reduction. Finally, the collective behavior of users of the system (agents) is employed to adapt the knowledge bases of queried information resources. This adaptation allows information resources to respond well to the evolving expectations of users.

Finally, we further describe the necessary mechanisms to automatically identify communities of practice in heterarchies endowed with adaptive webs. This identification is essential to decide whether a given member belongs to a particular community, which facilitates the control of a heterarchy in an emergent control hierarchy.

**Keywords:** Heterarchies, Organizations, Adaptive Webs, Distributed Artificial Intelligence, Recommendation Systems, Information Retrieval, Web-related technologies, Collaborative Systems, Adaptive Systems, Human-machine Interaction, Communities of Agents, Knowledge Representation.

---

<sup>1</sup> This paper was written from an oral presentation delivered at a workshop on *Heterarchies: Distributed Intelligence and the Organization of Diversity*, organized by David Stark at the Santa Fe Institute, October 13-14, 2000.

---

## 1. Social Agents for Heterarchies or Control Hierarchies?

If we are to study and build artificial Socially Intelligent Agents that are able to connect and interface to humans, we must first understand the environments in which they will be most useful. The fabric of human organization has been witnessing unprecedented transformation with the ubiquitous presence of networked information systems (such as the World Wide Web). At the same time, and probably causally related, we also witness the rapid establishment of globalized commercial, social, political, and cultural exchanges. This has led to the appearance and success of very open, diverse, flexible, and adaptive human organizations (businesses and non-profit organizations alike), such as modern Biotech [Clark, 1999; Koput and Powell, 2000] and Advertising [Grabher, 2001] firms. Given this transformation in human organizations, we need to study and produce agents that can facilitate and mediate interaction, communication, and cooperation among the people that comprise them. Indeed, we need to build systems that can foster and enhance the diversity and adaptability of these modern organizations. So let us start by understanding better the latter, and then continue with the description of our efforts to build the former.

“Whereas hierarchies involve relations of dependence and markets involve relations of independence, heterarchies involve relations of interdependence.” [Stark, 1999, page 159]

Stark [1999] has proposed “*Heterarchy*” to characterize social organizations with an enhanced capacity for innovation and adaptability. As *Complex Adaptive Systems* (CAS) [Holland, 1995], these organizations do not merely adapt to present demands, but are capable of deploying resources to innovate and improve their response to future, unpredictable demands. They achieve this adaptability due to two fundamental features: lateral (rather than vertical) accountability and organizational heterogeneity. The first of these features is a result of networked control and communication, similarly to the neural organizations that produce distributed intelligence. Networked or lateral organizations are in direct contrast with the tree-like, vertical chains of control of traditional hierarchies. The second feature means that heterarchies require diversity of components and building blocks.

What allows diversity to work in concerted ways to produce adaptability and intelligent behavior in such organizations, is the existence of a system of communication which makes use of *tags* to identify, produce, and communicate building blocks, situations, goals and whatever other elements organizations negotiate. Clippinger [1999] has emphasized the power of tags in managing these organizations. Clark [1999] too proposes that managers should use the genetic system as a guiding metaphor for managing heterarchies with tags. He defends that, like DNA, tags should not be used to impose an explicit control on organizations, but rather “to create

conditions for flexible, self-organizing response and to modulate the unfolding of intrinsic dynamics by the use of diffuse influences and the judicious application of simple forces and nudges.” [Ibid, page 65].

In any case, tags serve as the fundamental instrument for recognizing diversity, selecting appropriate agents and building blocks, and maintaining complex chains of interdependence and communication in heterarchies. Interestingly, it is the introduction of tags in network organizations that ends up producing hierarchies in biological and social systems as Holland observes in the following quote:

“Well-established tag-based interactions provide a sound basis for filtering, specialization, and cooperation. This, in turn, leads to the emergence of meta-agents and organizations that persist even though their components are continually changing. Ultimately, tags are the mechanism behind hierarchical organization – the agent / meta-agent / meta-meta-agent / . . . organization so common in CAS.” [Holland, 1995, pp. 14-15]

Indeed, as tagging is introduced in self-organizing networks, we can more easily observe subsystems, clusters and complex patterns of interaction. This is related to the process of emergence of different levels of organization from self-organizing networks so fundamental to CAS. As some components aggregate to produce a coherent, novel behavior not observed in the individual components, we require distinct levels of organization to explain the network. [Rocha, 1996; Clark, 1996]. For instance, Stark [1999] observed that the concept of property for firms in post-socialist Hungary, has itself emergent properties. That is, new phenomena is observed above the level of constituent unit firms when the interdependence network among these is analyzed beyond the sum of owner portfolios. A Hungarian business network is a complex network of intersecting alliances with distinct strategies, which are distinguishable only at the level of the network.

But when networks with tags generate emergent properties, we observe a hierarchy of levels of description. Not necessarily a traditional tree-like or a Chinese box hierarchy, but the presence of distinct levels of phenomena where an organization at one level becomes a component of a larger organization at a higher level. Rosen [1969], has precisely defined hierarchical organization as a system that has more than one simultaneous activity, such that alternative modes of description are an absolute necessity [see also Cariani, 1989]. So where do complex adaptive organizations that rely on diverse components, networked (rather than vertical) control and communication, and tagging lead us to? What is it going to be, heterarchy or hierarchy?

An answer to this question can be drawn from Pattee’s [1973, 1976] work on hierarchies. In particular, his distinction between structural and control (functional) hierarchies, and the realization that what is relevant for

biological and social organization is not so much the structural arrangements of components, but the relation among levels [Umerez, 1995]. In other words, the hierarchical relations of interest in biological and social organizations, are not to be found in the pattern of interrelations of components, but in the emergence of different levels of organization and the nature of the interface between these levels [Pattee, 1973].

He distinguished control hierarchies from the structural hierarchies examined by Simon [1962, 1973], whose levels depend on the criteria of number, forces, and time scales, and where one component can be treated as representative of the collection at each level, which can thus be averaged out to a boundary condition when describing the level above. This is what Simon refers to as “near-decomposability”. In contrast, Pattee observed that in the control hierarchies of biological and social systems “the upper level exerts a specific, dynamic constraint on the details of the motion at lower level, so that the fast dynamics of the lower level cannot simply be averaged out. [...] This amounts to a feedback path between levels.”

Furthermore, the one-particle approximation possible in structural hierarchies, fails for control hierarchies because constrained components are atypical. Consider a control hierarchy of only two levels: a network of components (e.g. Hungarian firms) and an emergent level (e.g. Hungarian business networks) which exerts controls on the components of the lower level. Even though lower level components may start as typical, given the constraints of the higher level which are imposed on the collection, each component is eventually selected to perform atypical roles. In other words, the constraints of a higher level increase the diversity of components at the lower level.

Pattee [1973] also noted that the control value of certain molecules in the biological organization, such as DNA, “is not an inherent chemical property; it is a complex relation established by a collective hierarchical organization requiring the whole organism.” [Ibid, page 78]. DNA possesses a special control function in the biological organization only in the integrated collection of molecules that forms a cell. When contrasted with other molecules, it does not have any special characteristics. Its control role is observable only from the emergent properties of the whole cellular organization. Finally, with the following quote Pattee takes us back to tags in heterarchies:

“A control molecule is not a typical molecule even though it has a normal structure and follows normal laws. In the collection where it exerts some control it is not just a physical structure – it functions as a *message*, and therefore the significance of this message does not derive from its detailed structure but from the set of hierarchical constraints which we may compare with the integrated rules of a language. These rules do not lie in the structure of any element.” [Pattee, 1973, page 81].

The point of this exposition is to defend that the biological organization and the flexible, innovative, adaptable social organizations we want to manage, at their structural level should effectively be seen as heterarchies because of all the arguments presented by Stark [1999]. They need to be networks of interdependent components with lateral accountability and high redundancy. But if they are also CAS, then their emergent properties define a control hierarchy whose feedback constraints the components of the lower level heterarchy. In particular, as Pattee noted, components of CAS need to be flexible in taking different roles assigned by the needs of the whole network – as they move from typical to atypical components by the constraints of the control hierarchy. The nature of this constraint or control leads the diversity of components in the heterarchy to increase, namely because some of these components become messages or tags used by the upper level to produce the desired behavior in the heterarchy.

What is necessary in the heterarchy formulation is this grounding in an emergent control hierarchy. Indeed, while Clark [1999] rightly proposes that managers, like DNA, should not function as explicit specification controls, but rather as a system of modulating the dynamics of the heterarchy, nonetheless fails to notice that DNA can be seen as control only from the standpoint of the emergent level of the whole organism. This way, the role of a manager of a heterarchy arises from the constraints set up by the emergent level of the control hierarchies that controls the lower level heterarchy. Therefore, rather than following an absolute recipe, managers need to first and foremost recognize and identify control hierarchies in which their heterarchy participates, and seek context-specific strategies appropriate for the former. Indeed, the same heterarchical organization may participate in distinct control hierarchies, as much as an organism can be a member of different environmental self-organizing groupings (societies, ecological processes, etc.)

As Pattee and Holland propose, tags are the key mechanism used by control hierarchies to harness the heterarchies at lower levels. Indeed, tags are the fundamental control lever in organizations. “Tags act to define a community of interest or activity” [Clippinger, 1999, page 68], and communities and activities are precisely the expressions of higher levels of control hierarchies. It is therefore important to study how to discover and generate appropriate tags and communication systems to facilitate the control of heterarchies. Below I discuss adaptive webs, which are essential to enable diverse, adaptable social organizations, and the identification and automatic tagging of communities of users.

## 2. Adaptive Webs: The Communication Fabric of Evolving Heterarchies

### 2.1 Distributed Information Systems and Information Retrieval

*Distributed Information Systems* (DIS) are collections of electronic networked *information resources* (e.g. databases) in some kind of interaction with communities of users; examples of such systems are: the Internet, the World Wide Web, corporate intranets, databases, library information retrieval systems, etc. DIS serve large and diverse communities of users by providing access to a large set of heterogeneous electronic information resources. *Information Retrieval* (IR) refers to all the methods and processes for searching relevant information out of information systems (isolated or part of DIS) that contain extremely large numbers of documents. As the complexity and size of both user communities and information resources grows, the fundamental limitations of traditional information retrieval systems have become evident in modern DIS.

Traditional IR systems are based solely on *keywords* that index (semantically characterize) documents and a query language to retrieve documents from centralized databases according to these keywords – users need to know how to pull relevant information from passive databases. This setup leads to a number of flaws [Rocha and Bollen, 2000], which prevent traditional IR processes in DIS to achieve any kind of interesting coupling with users. The human-machine interaction observed in these systems is particularly rigid: Most cannot proactively push relevant information to its users about related topics that they may be unaware of, there is typically no mechanism to exchange knowledge, or crossover of relevant information among users and information resources, and there is no mechanism to recombine knowledge in different information resources to infer new linguistic categories of keywords used by evolving communities of users. In other words, traditional IR keeps DIS as static, passive, and isolated repositories of data; no interesting human-machine co-evolution of knowledge or learning is achieved.

### 2.2 Enabling Evolving DIS for Heterarchies

The limitations of traditional IR and DIS are even more dramatic when contrasted with biological distributed systems such as immune, neural, insect, and social networks. Biological networks function largely in a distributed manner, without recourse to central controllers, while achieving tremendous ability to respond in

concerted ways to different environmental necessities. In particular, they are typically endowed with the ability to elicit appropriate responses to specific demands, to transfer and process relevant information across the network, and to adapt to a changing environment by creating novel behaviors (often from recombination of existing ones). These abilities, necessary to establish adaptable heterarchies, are precisely what has been lacking in IR.

Biological networks effectively evolve in an open-ended manner; we are interested in endowing DIS with a similar open-ended capacity to evolve with their users – to achieve an open-ended semiosis with them [Rocha, 2000]. In biology, open-ended evolution originates from the existence of material building blocks that self-organize non-linearly [e.g. Kauffman, 1993] and are combined via a specification control, such as the genetic system [Rocha, 1998]. In contrast, computer systems were constructed precisely with rigid building blocks constrained in such a way as to allow minimum dynamic self-organization and maximum programmability, which results in no inherent evolvability [Conrad, 1990]. Therefore, to attain any evolvability in current digital computer systems, we need to program in some “softer” building blocks that can be used to realize the kind of dynamical richness we encounter in biological systems [Rocha and Bollen, 2000].

The ultimate goal of IR is to produce or recommend relevant information to users. It seems obvious that the foundation of any useful recommendation should be first and foremost based on the identification of users and subject matter. In this sense, the goal of recommendation systems can be seen as similar to that of most biological systems, or heterarchies: to recognize agents (users) and elicit appropriate responses from components of the distributed information network. Furthermore, the information network should learn and adapt to the community of agents (users) it interacts with – its environment.

The ability to use adaptive DIS, or *Adaptive Webs*, is increasingly important as social organizations become more and more heterarchical. With the flexibility inherent in heterarchies, particularly as they become components of more complex emergent structures at higher levels of control hierarchies, it is important to find intelligent ways to store the knowledge of an organization. Furthermore, because they lack a specific, vertical chain of command, a collaborative communication fabric is essential to maintain internal coherence in the horizontal organization by proactively recommending appropriate documents and components working on similar problems. Thus, the two main goals of adaptive webs are to endow heterarchies with *evolving knowledge repositories* and a *recommendation interface* among its components. In this sense, adaptive webs need to be seen as the communication fabric that permeates heterarchies and helps them to adapt and co-evolve with changing environments. To achieve this, adaptive webs must:

1. Adapt and evolve with the needs of the communities of users.
2. Represent the knowledge of the organization (its identity), even as individual elements of the organization abandon it.
3. Automatically identify and tag relevant, emergent communities of users thus preserving the diversity of components.

As the communication fabric of heterarchies, adaptive webs negotiate in tags which ultimately help to control the former (as discussed in section 1). In this paper, I discuss adaptive webs as technology capable of instantiating adaptive tag communication in heterarchies, as well as creating tags necessary to cope with a changing environment. The three requirements for adaptive webs are dealt with in the sections below. In sections 3 to 8 the *TalkMine* system is detailed. This is a recommendation system that provides points 1 and 2. Point 2 is further discussed in section 8. Finally section 9 describes how point 3 is implemented in an adaptive web implemented with *TalkMine*.

---

### 3. From IR to Adaptive Recommendation: From “Pull” to “Push”

As discussed above, IR systems are based on *keywords* that index documents. The recommendation systems we seek for heterarchies work with an increased domain, though we can still think of them as operating with a generalized keyword/document relation. Keywords should be thought of as *tags*; the tags on which the communication in heterarchies is cast. Documents should be thought of as actual documents (reports, web pages, books, and the like) as well as any other *components* of an organization, such as people, processes and strategies which can be tagged. Even though in the subsequent sections recommendation systems are described as operating on the keyword/document relation (to be more in sync with the IR literature), they should be thought of as operating on the generalized tag/component relation.

New approaches to IR have been proposed to improve its inflexible passive “pull” algorithms. *Active recommendation systems*, also known as *Active Collaborative Filtering* [Chislenko, 1998] or *Knowledge Self-Organization* [Johnson et al, 1998] are IR systems which rely on active computational environments that interact with and adapt to their users. They effectively push relevant information to users according to previous patterns of IR or individual user profiling.

Recommendation systems are typically based on human-machine interaction mediated by intelligent agents, or other decentralized components, and come in several varieties:

1. In *content-based* recommendation, user profiles are created based on the system's keywords. Documents are recommended to users according to the similarity of their profiles and the similarity of keywords constructed from a semantic distance function obtained from the associations between keywords and documents. Two documents are close when they are classified by many of the same keywords. This is the case of systems such as *InfoFinder* [Krulwich and Burke, 1996], *NewsWeeder* [Lang, 1995], and many systems developed for the routing task at the TREC Conferences [Harman, 1994].
2. In *collaborative* recommendation no description of the semantics or content of documents is involved, rather recommendations are issued according to a comparison of the profiles of several users that tend to access the same documents. The comparison depends on a distance function between user profiles, defined not by keywords, but on the sets of actual documents retrieved. Two user profiles are close when their users have retrieved many of the same documents. This is the case of systems such as *GroupLens* [Resnick et al, 1994; Kostan et al, 1997], *Bellcore Video Recommender* [Hill et al, 1995], *Ringo* [Shardanad and Maes, 1995]. When user feedback is allowed, this type of recommendation is known as *Information Filtering* [Good et al, 1999]. For a description of the collaborative recommendation framework see Herlocker et al [1999].
3. In *structural* recommendation, data-mining techniques are employed on the relations among documents and keywords, to discover related documents or documents of particular importance (authorities) in a given information resource. A large portion of work in this area, is concerned with the analysis of the graph structure of Web Hyperlinks (regardless of document keywords), e.g. work pursued under the *CLEVER* Project [Kleinberg, 1998; Chakrabarti et al, 1999], or other graph-theoretic approaches such as Watts' [1999] Small World graphs. A second large area of research is concerned with the semantic relations between documents and keywords, which are analyzed with algebraic techniques such as Singular Value Decomposition, known in IR as Latent Semantic Indexing (LSI) [Berry et al, 1994; Kannan and Vempala, 1999]. Documents are recommended to users according to the way they are associated with other documents and/or keywords: the semantic structure of information resources.

4. In *collective* recommendation, the behavior of communities of users is integrated, and utilized to adapt the structure (the pattern of associations) of information resources. This kind of system tracks the paths users follow in the structure of information resources as they retrieve documents. The more certain sets of documents tend to be retrieved together in paths followed by different users, the closer they become in the structure of the information resource. This type of algorithm employs the distributed behavior of a collection of users to adapt DIS, resulting in systems that learn the interests of their communities of users much in the same way as social insects discover paths based on the pheromone trails left behind by other insects in their colony [Rocha and Bollen, 2000], thus, in time, recommending more and more appropriate documents. This is the case of Adaptive Hypertext systems [Brusilovsky et al, 1998; Bollen and Heylighen, 1998; Eklund, 1998], Knowledge Self-Organization [Johnson et al, 1998; Heylighen, 1999], as well as the work on the collective discovery of linguistic categories [Rocha, 1997a, 1999] detailed below.

Content-based systems depend on single user profiles, and thus cannot effectively recommend documents about previously unrequested content to a specific user. That is, these systems cannot compare and recommend related documents characterized by keywords not previously collected into a given user's profile. Conversely, pure collaborative systems, match only the profiles of users that (to a great extent) have requested exactly the same documents; for instance, different book editions or movie review web sites from different news organizations may be considered distinct documents.

The shortcoming of structural approaches is that they assume that the existing, often static, structure of an information resource contains all the relevant knowledge to be discovered. However, it is often the case that such structure is very poorly designed. On the web in particular, the hypertext links are often not created between important documents, due perhaps to the hurried way in which web sites are created. Indeed, the Web is often more a repository of isolated documents, than a good example of a hypertext fabric. The same applies to the keyword/document relations necessary for LSI.

Collective approaches have the important advantage of adapting to the collective behavior of users, even as it develops in time. This way, a poor initial structure can improve, by creating, strengthening or weakening associations among documents or between documents and keywords. This is ideal for heterarchies. Furthermore, collective recommendation systems can operate without storing individual profiles, thus offering

a more private platform for recommendation. Indeed, recommendations are issued according to the adapted structure of the information resources, not according to user profiles. Users can be seen as anonymous social agents. Furthermore, as we shall discuss later, the adapted information resources allow us to capture the knowledge traded by a community of agents. Nonetheless, a disadvantage of collective approaches is that they implement a positive feedback with their communities of users, possibly leading to an excessive adaptation to the interests of a majority of users, thus reducing the diversity of knowledge by recommending only the most retrieved documents in a given area: e.g. the “best of” lists found at Web sites such as *Amazon.com* – this is the so-called “curse of averages”.

It is clear that good recommendation systems require aspects of all approaches to avoid the shortcomings of each individual one. This is the case, for instance, of *Fab* [Balabanović and Shoham, 1997] and *Amalthea* [Moukas and Maes, 1998], which are both content and collaborative recommendation systems. This way they can discover similar users who have not simply retrieved many of the same exact documents, but documents characterized by many of the same keywords. Furthermore, keywords from documents that users have not actually retrieved, may be added to their profiles because they belong to the profiles of other similar users.

Still, neither *Fab* nor *Amalthea* (nor similar systems) adapt the structure of their information resources with collective user behavior, nor do they use the data-mining techniques of structural algorithms to characterize the knowledge those store. In this sense, they cannot capture the evolving nature of the knowledge of communities of users. In other words, even though they are able to characterize the interests of individual users (both with documents and keywords), the structure of information resources (e.g. Web hyperlink structure or document/keyword matrix) remains unchanged. Furthermore, they rely on individual user profiles, and there is also not an explicit means to discover the knowledge categories that particular communities of users employ. Next I describe the Active Recommendation Project [Rocha and Bollen, 2000] which is building a hybrid Collective/Structural/Content recommendation system designed precisely to tackle these issues. Namely, to adapt information resources to their evolving communities of users, to characterize the knowledge stored in these information resources, and to preserve diversity while not accumulating private user profiles.

---

## 4. The Active Recommendation Project

The *Active Recommendation Project*<sup>2</sup> (ARP), part of the Library Without Walls Project, at the Research Library of the Los Alamos National Laboratory (LANL) is engaged in research and development of recommendation systems for digital libraries. The *information resources* available to ARP are large databases with academic articles. These databases contain bibliographic, citation, and sometimes abstract information about academic articles. Typical databases are *SciSearch*<sup>®</sup> and *Biosis*<sup>®</sup>; the first contains articles from scientific journals from several fields collected by ISI (Institute for Scientific Indexing), while the second contains more biologically oriented publications. We do not manipulate directly the records stored in these information resources, rather, we created a repository of XML (about 3 million) records which point us to documents stored in these databases [Rocha and Bollen [2000]].

These information resources are components of the LANL organization which facilitate the dissemination of scientific knowledge among communities of researchers. LANL is far from being a heterarchy in a managerial sense, but its research community is largely a horizontal, heterogeneous organization with many flexible research units without fixed boundaries. Thus, LANL's digital library is interested in functioning as an adaptive web as described in section 2<sup>3</sup>, so that its research community can be innovative and respond to challenges in an evolving scientific and technological world – in other words, so that it can be more of an heterarchy. The two main goals of LANL's digital library are the same as those of adaptive webs at large: to be an *evolving knowledge repository* and to provide a *recommendation interface* to enhance collaboration and discovery in its research community.

To build an adaptive web we need to design recommendation systems endowed with:

1. A means to recognize the knowledge of *users* .
2. A means to characterize the knowledge stored in *information resources*.
3. A 2-way means to exchange knowledge between users and information resources: a *conversation* process. As information resources become more and more complex, we cannot expect a simple 1-way query (“pull”) to work well. Instead, we need a means to integrate the interests of the user with the knowledge specific to each information resource via an interactive recommendation process (“push”).

---

<sup>2</sup>More information, results, and testbed available at <http://www.c3.lanl.gov/~rocha/lww>.

<sup>3</sup> See the *Library Without Walls* Project web page.

4. **Adaptation** mechanisms. We also want DIS to adapt to their community of users, as well as to exchange and re-combine knowledge leading to evolvability and creativity.

Below I describe efforts to include these design requirements for recommendation systems using Soft Computing technology. I also discuss how a useful and more natural knowledge management of DIS is achieved with these soft computing designs.

#### 4.1 Characterizing the Knowledge stored in an Information Resource

We have compiled relational information between records<sup>4</sup> and keywords and among records: the *semantics* and the *structure* respectively. The semantics is formalized as a very sparse **Keyword-Record Matrix** *A*. The structure is formalized as the very sparse **Citation Matrix** *C*, which is a record-record matrix [details in Rocha and Bollen, 2000]. From these matrices, we have calculated additional matrices holding measures of closeness between records and between keywords: the **Inwards Structural Proximity Matrix** or *co-citation* [Small, 1973], the **Outwards Structural Proximity Matrix** or *bibliographic coupling* [Kessler, 1963], the **Record Semantic Proximity Matrix** (for any two records it is defined by the number of keywords that qualify both, divided by the number of keywords that qualify either one), and the **Keyword Semantic Proximity Matrix** (for two keywords, it is the number of records they both qualify, over the number of records either one qualifies).

These matrices holding measures of closeness, formally, are proximity relations [Klir and Yuan, 1995; Miyamoto, 1990] because they are reflexive and symmetric fuzzy relations. Their transitive closures are known as similarity relations [Ibid]. The collection of this relational information, all the proximity relations as well as *A* and *C*, is an expression of the particular knowledge an information resource conveys to its community of users. Notice that distinct information resources typically share a very large set of keywords and records. However, these are organized differently in each resource, leading to different collections of relational information. Indeed, each resource is tailored to a particular community of users, with a distinct history of utilization and deployment of information by its authors and users. For instance, the same keywords will be related differently for distinct resources. Therefore, we refer to the relational information of each information resource as a *Knowledge Context*. We do not mean to imply that information resources possess cognitive abilities. Rather, we note that the way records are organized in information resources is an expression of the knowledge traded by its

---

<sup>4</sup> Records contain bibliographical information about published documents. Records can be thought of as unique pointers to documents, thus, for the purposes of this article, the two terms are interchangeable.

community of users. Records and keywords are only tokens of the knowledge that is ultimately expressed in the brains of users. A knowledge context simply mirrors some of the collective knowledge relations and distinctions shared by a community of users.

In [Rocha and Bollen, 2000] we have discussed how these proximity relations are used in ARP. However, the ARP recommendation system described in this article (*TalkMine*) requires only the Keyword Semantic Proximity (*KSP*) matrix, obtained from  $A$  by the following equation:

$$ksp(k_i, k_j) = \frac{\sum_{k=1}^m (a_{i,k} \wedge a_{j,k})}{\sum_{k=1}^m (a_{i,k} \vee a_{j,k})} = \frac{N_{\cap}(k_i, k_j)}{N_{\cup}(k_i, k_j)} = \frac{N_{\cap}(k_i, k_j)}{N(k_i) + N(k_j) - N_{\cap}(k_i, k_j)} \quad (1)$$

The semantic proximity between two keywords,  $k_i$  and  $k_j$ , depends on the sets of records indexed by either keyword, and the intersection of these sets.  $N(k_i)$  is the number of records keyword  $k_i$  indexes, and  $N_{\cap}(k_i, k_j)$  the number of records both keywords index. This last quantity is the number of elements in the intersection of the sets of records that each keyword indexes. Thus, two keywords are near if they tend to index many of the same records. Table I presents the values of *KSP* for the 10 most common keywords in the ARP repository.

**Table I:** Keyword Semantic Proximity for 10 most frequent keywords

	cell	studi	system	express	protein	model	activ	human	rat	patient
cell	1.000	0.022	0.019	0.158	0.084	0.017	0.085	0.114	0.068	0.032
studi	0.022	1.000	0.029	0.013	0.017	0.028	0.020	0.020	0.020	0.037
system	0.019	0.029	1.000	0.020	0.017	0.046	0.022	0.014	0.021	0.014
express	0.158	0.013	0.020	1.000	0.126	0.011	0.071	0.103	0.078	0.020
protein	0.084	0.017	0.017	0.126	1.000	0.013	0.070	0.061	0.041	0.014
model	0.017	0.028	0.046	0.011	0.013	1.000	0.016	0.016	0.026	0.005
activ	0.085	0.020	0.022	0.071	0.070	0.016	1.000	0.058	0.053	0.021
human	0.114	0.020	0.014	0.103	0.061	0.016	0.058	1.000	0.029	0.021
rat	0.068	0.020	0.021	0.078	0.041	0.026	0.053	0.029	1.000	0.008
patient	0.032	0.037	0.014	0.020	0.014	0.005	0.021	0.021	0.008	1.000

From the inverse of *KSP* we obtain a distance function between keywords:

$$d(k_i, k_j) = \frac{1}{KSP(k_i, k_j)} - 1 \quad (2)$$

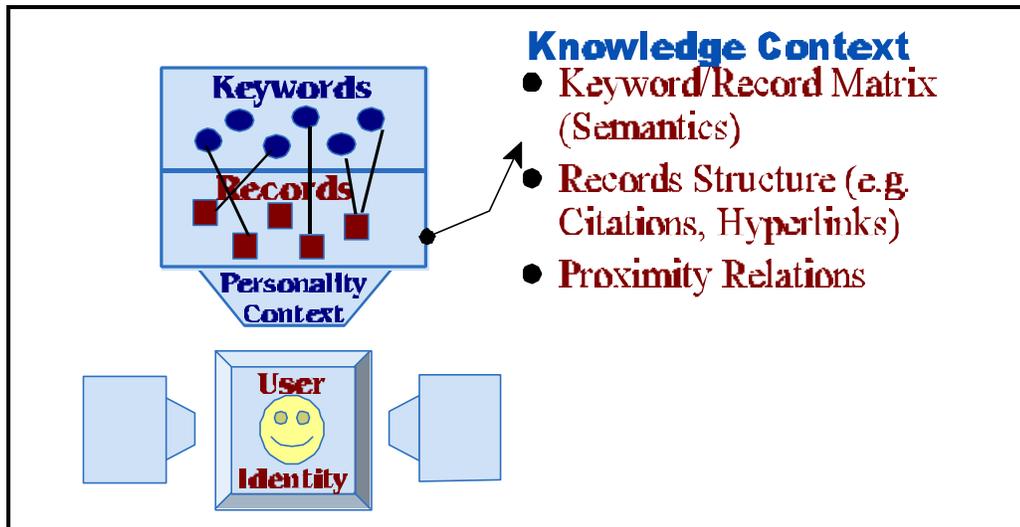
$d$  is a distance function because it is a nonnegative, symmetric real-valued function such that  $d(k, k) = 0$ . It is not an Euclidean metric because it may violate the triangle inequality:  $d(k_1, k_2) \leq d(k_1, k_3) + d(k_3, k_2)$  for some keyword  $k_3$ . This means that the shortest distance between two keywords may not be the direct link but rather an indirect pathway. Such measures of distance are referred to as semi-metrics [Galvin and Shore, 1991].

## 4.2 Characterizing Users

Users interact with information resources by retrieving records. We use their retrieval behavior to adapt the respective knowledge contexts of these resources (stored in the proximity relations). But before discussing this interaction, we need to characterize and define the capabilities of users: our agents. The following capabilities are implemented in enhanced “browsers” distributed to users.

1. ***Present interests*** described by a set of keywords  $\{k_1, \dots, k_p\}$ .
2. ***History of Information Retrieval (IR)***. This history is also organized as a knowledge context as described in 2.1, containing pointers to the records the user has previously accessed, the keywords associated with them, as well as the structure of this set of records. This way, we treat users themselves as information resources with their own specific knowledge contexts defined by their own proximity information.
3. ***Communication Protocol***. Users need a 2-way means to communicate with other information resources in order to retrieve relevant information, and to send signals leading to adaptation in all parties involved in the exchange.

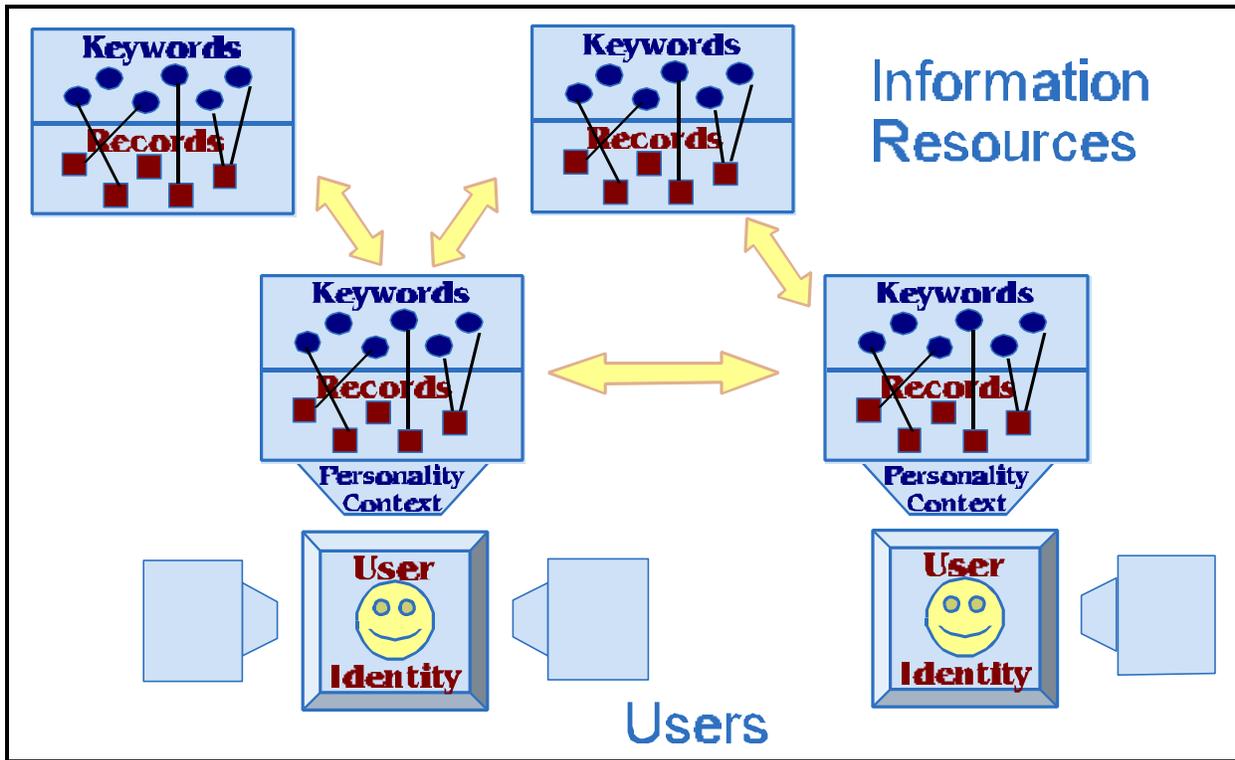
Regarding point 2, the history of IR, notice that the same user may query information resources with very distinct sets of interests. For example, one day a user may search databases as a biologist looking for scientific articles, and the next as a sports fan looking for game scores. Therefore, each enhanced browser allows users to define different “personalities”, each one with its distinct history of IR defined by



**Figure 1:** Each user can store different personalities in enhanced browsers. Each personality is stored as a knowledge context created from previous history of IR. The actual identity of the user can remain private.

independent knowledge contexts (with distinct proximity data), see figure 1.

Because the user history of IR is stored in personal browsers, information resources do not store user profiles. Furthermore, all the collective behavior algorithms used in ARP do not require the identity of users. When users communicate (3) with information resources, what needs to be exchanged is their present interests or query (1), and the relevant proximity data from their own knowledge context (2). In other words, users make a query, and then share the relevant knowledge they have accumulated about their query, their “world-view” or context, from a particular personality, without trading their identity. Next, the recommendation algorithms integrate the user’s knowledge context with those of the queried information resources (possibly other users), resulting in appropriate recommendations. Indeed, the algorithms we use define a communication protocol between knowledge contexts, which can be very large databases, web sites, or other users. Thus, the overall architecture of the recommendation systems we use in ARP is highly distributed between information resources and all the users and their browsing personalities (see figure 2).



**Figure 2:** The algorithms we use in ARP define a distributed architecture based on communication between knowledge contexts from information resources and users alike.

The collective behavior of all users is also aggregated to adapt the knowledge contexts of all intervening information resources and users alike. This open-ended learning process [Rocha, 2000] is enabled by the *TalkMine* recommendation system described below.

---

## 5. Categories and Distributed Memory

### 5.1 A Model of Categorization from Distributed Artificial Intelligence

*TalkMine* is both a content-based and collaborative recommendation system based on a model of linguistic categories [Rocha, 1999], which are created from conversation between users and information resources and used to re-combine knowledge as well as adapt it to users. The model of categorization used by *TalkMine* is

described in detail in [Rocha, 1997a, 1999, 2000]. Basically, as also suggested by Clark [1993], categories are seen as representations of highly transient, context-dependent knowledge arrangements, and not as model of information storage in the brain. In this sense, in human cognition, categories are seen as linguistic constructs used to store temporary associations built up from the integration of knowledge from several neural sub-networks. The categorization process, driven by language and conversation, serves to bridge together several distributed neural networks, associating tokens of knowledge that would not otherwise be associated in the individual networks. Thus, categorization is the chief mechanism to achieve knowledge recombination in distributed networks leading to the production of new knowledge [Rocha, 1999, 2000].

*TalkMine* applies such a model of categorization of distributed neural networks driven by language and conversation to DIS and recommendation systems. Instead of neural networks, knowledge is stored in information resources, from which we construct the knowledge contexts with respective proximity relations described in section 4. *TalkMine* is used as a conversation protocol to categorize the interests of users according to the knowledge stored in information resources, thus producing appropriate recommendations and adaptation signals.

## **5.2 Distributed Memory is Stored in Knowledge Contexts**

A knowledge context of an information resource (section 4.1) is not a connectionist structure in a strong sense since keywords and records are not distributed as they can be identified in specific nodes of the network [van Gelder, 1991]. However, the same keyword indexes many records, the same record is indexed by many keywords, and the same record is typically engaged in a citation (or hyperlink) relation with many other records. Losing or adding a few records or keywords does not significantly change the derived semantic and structural proximity relations (section 4) of a large network. In this sense, the knowledge conveyed by such proximity relations is distributed over the entire network of records and keywords in a highly redundant manner, as required of sparse distributed memory models [Kanerva, 1988]. Furthermore, Clark [1993] proposed that connectionist memory devices work by producing metrics that relate the knowledge they store. As discussed in section 4, the distance functions obtained from proximity relations are semi-metrics, which follow all of Clark's requirements [Rocha, 2000]. Therefore, we can regard a knowledge context effectively as a distributed memory bank. Below we discuss how such distributed knowledge adapts to communities of users (the environment) with Hebbian type learning.

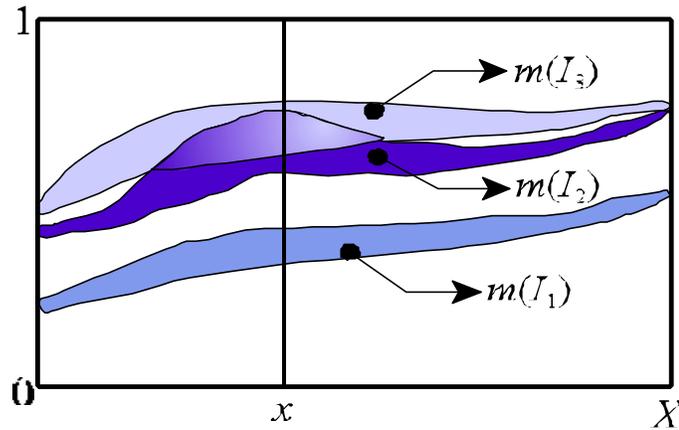
In the *TalkMine* system we use the KSP relation (eq. 1) from knowledge contexts. It conveys the knowledge stored in an information resource in terms of a measure of proximity among keywords. This proximity relation is unique to each information resource, reflecting the semantic relationships of the records stored in the latter, which in turn echo the knowledge of its community of users and authors. *TalkMine* is a content-based recommendation system because it uses a keywords proximity relation. Next we describe how it is also collaborative by integrating the behavior of users. A related structural algorithm, also being developed in ARP, is described in [Rocha and Bollen, 2000].

---

## 6. Evidence Sets: Capturing the Linguistic “And/Or” in Users’ Queries

### 6.1 Evidence Sets Model Categories

*TalkMine* uses a set structure named *evidence set* [Rocha 1994, 1997a, 1997b, 1999], an extension of a fuzzy set [Zadeh, 1965], to model of linguistic categories. Evidence sets are set structures which provide interval degrees of membership, weighted by a probability constraint. They are defined by two complementary



**Figure 3:** Example of an Evidence Set.

dimensions: membership and belief. The first represents an interval fuzzy degree of membership, and the second a subjective degree of belief on that membership (see figure 3). More details in Rocha [1999, 2001]

Each interval of membership with its correspondent evidential weight, represents the degree of importance of a particular element  $x$  of  $X$  in category  $A$  *according* to a particular *perspective*. Thus, the membership of each element  $x$  of an evidence set  $A$  is defined by distinct intervals representing different, possibly conflicting, perspectives. This way, categories are modeled not only as sets of elements with a membership degree (or

prototypicality see Rocha[1999]), but as sets of elements which may possess different interval membership degrees for different contexts or perspectives on the category.

The basic set operations of complementation, intersection, and union have been defined and establish a belief-constrained approximate reasoning theory of which fuzzy approximate reasoning and traditional set operations are special cases [Rocha 1997a, 1999]. Intersection (Union) is based on the minimum (maximum) operator for the limits of each of the intervals of membership of an evidence set. For the purposes of this article, the details of these operations are not required, please consult [Rocha, 1999] for more details.

## 6.2 The Uncertainty Content of Evidence Sets

Evidence possess three distinct types of uncertainty [Rocha, 1997b, 2001]: fuzziness, nonspecificity, and conflict. Membership of an element  $x$  in an evidence set is defined as a set of intervals constrained by a probability restriction; this means that the membership is fuzzy, nonspecific, and *conflicting*, since the element is a member of the set with several degrees (fuzziness) that vary in each interval (nonspecificity) with some probability (conflict).

To capture the uncertainty content of evidence sets, the uncertainty measures of Klir [1993] were extended from finite to infinite domains [Rocha, 1997b]. The total uncertainty,  $U$ , of an evidence set  $A$  was defined by:  $U(A) = (IF(A), IN(A), IS(A))$ . The three indices of uncertainty, which vary between 1 and 0,  $IF$  (*fuzziness*),  $IN$  (*nonspecificity*), and  $IS$  (*conflict*) were introduced in [Rocha, 1997a, 1997b], where it was also proven that  $IN$  and  $IS$  possess good axiomatic properties wanted of information measures. For the purposes of this article, all we need to know is that these measures vary in the unit interval, for full details see [Rocha, 1997b].

## 6.3 Obtaining an Evidence Set from Fuzzy Sets: The Linguistic “And/Or”

Fundamental to the *TalkMine* algorithm is the integration of information from different sources into an evidence set, representing the category of topics (described by keywords) a user is interested at a particular time. In particular, as described below, these sources of information contribute information as fuzzy sets. A procedure for integrating several fuzzy sets into an evidence set has been developed in Rocha [1997, 2001]. This procedure combines several pieces of evidence defined by fuzzy sets, by modeling an ambiguous linguistic “and/or”. In common language, often “and” is used as an unspecified “and/or”. In other words, what we mean

by the statement “I am interested in x and y”, can actually be seen as an unspecified combination of “x and y” with “x or y”. This is particularly relevant for recommendation systems where it is precisely this kind of statement from users that we wish to respond to. For details of this procedure please refer to Rocha [2001]. For the purposes of this article, it suffices to understand that a procedure exists to combine evidence (in the form of fuzzy sets) from different sources into an evidence set or linguistic category.

---

## 7. *TalkMine*: Integrating Several Sources of Knowledge via Conversation

### 7.1 Inferring User Interest

The act of recommending appropriate documents to a particular user needs to be based on the integration of information from the user (with her history of retrieval) and from the several information resources being queried. With *TalkMine* in particular, we want to retrieve relevant documents from several information resources with different keyword indexing. Thus, the keywords the user employs in her search, need to be “decoded” into appropriate keywords for each information resource. Indeed, the goal of *TalkMine* is to project the user interests into the distinct knowledge contexts of each information resource, creating a representation of these interests that can capture the perspective of each one of these contexts.

Evidence Sets (section 6) were precisely defined to model categories (knowledge representations) which can capture different perspectives. As described in section 4.2, the present interests of each user are described by a set of keywords  $\{k_1, \dots, k_p\}$ . Using these keywords and the keyword distance function (eq. 2) of the several knowledge contexts involved (one from the user and one from each information resource being queried), the interests of the user, “seen” from the perspectives of the several information resources, can be inferred as an evidence category using the evidence set procedure mentioned in 6.3.

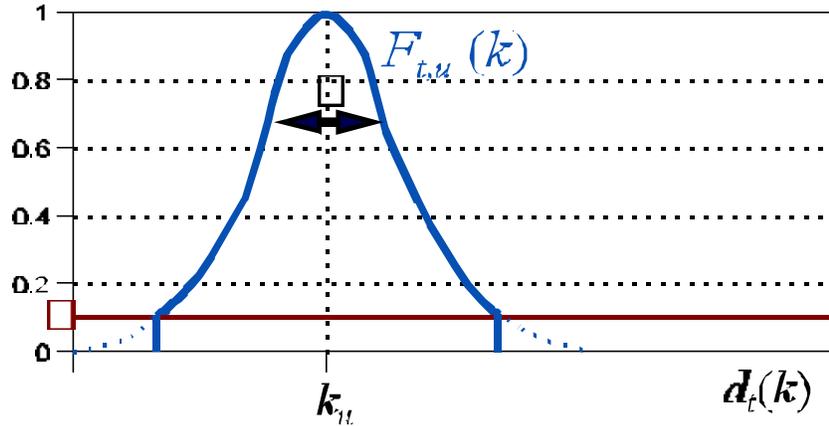
Let us assume that  $r$  information resources  $R_i$  are involved in addition to the user herself. The set of keywords contained in all the participating information resources is denoted by  $\mathbf{K}$ . As described in section 4, each information resource is characterized as a knowledge context containing a KSP relation among keywords (eq. 1) from which a distance function  $d$  is obtained (eq. 2).  $d_0$  is the distance function of the knowledge context of the user, while  $d_1 \dots d_r$  are the distance functions from the knowledge contexts of each of the information resources.

#### 7.1.1 Spreading Interest Fuzzy Sets

For each information resource  $R_t$  and each keyword  $k_u$  in the user's present interests  $\{k_1, \dots, k_p\}$ , a *spreading interest fuzzy set*  $F_{t,u}$  is calculated using  $d_t$ :

$$F_{t,u}(k) = \max \left[ e^{\left( -a \cdot d_t(k, k_u) \right)^2}, e \right], \forall k \in R_t, t = 1 \dots r, u = 1 \dots p \quad (3)$$

This fuzzy set contains the keywords of  $R_t$  which are closer than  $k_u$ , according to an exponential function of  $d_t$ .  $F_{t,u}$  spreads the interest of the user in  $k_u$  to keywords of  $R_t$  that are near according to  $d_t$ . The parameter  $a$  controls the spread of the exponential function.  $F_{t,u}$  represents the set of keywords of  $R_t$  which



**Figure 4:** The exponential membership function of  $F_{t,u}(k)$  spreads the interest of a user on keyword  $k_u$  to close keywords according to distance function  $d_t(k)$  for each information resource  $R_t$ .

are near or very related to keyword  $k_u$ . Because the knowledge context of each  $R_t$  contains a different  $d_t$ , each  $F_{t,u}$  will also be a different fuzzy set for the same  $k_u$ , possibly even containing keywords that do not exist in other information resources. There exist a total of  $n = r \cdot p$  spreading interest fuzzy sets  $F_{t,u}$ . Figure 4 depicts a generic  $F_{t,u}$ .

### 7.1.2 Combining the Perspectives of Different Knowledge Contexts on the User Interest

To combine the  $n$  fuzzy sets  $F_i$  the evidence combination [Rocha, 2001] mentioned in 6.3 is employed. This results in an evidence set  $ES_0(k)$  defined on  $\mathbf{K}$ , which represents the interests of the user inferred from spreading the initial interest set of keywords in the knowledge contexts of the intervening information resources. The inferring process combines each  $F_{t,u}$  with the “and/or” linguistic expression entailed by such combination. Each

$F_{t,u}$  contains the keywords related to keyword  $k_u$  in the knowledge context of information resource  $R_t$ , that is, the perspective of  $R_t$  on  $k_u$ . Thus,  $ES_0(k)$  contains the “and/or” combination of all the perspectives on each keyword  $k_u \in \{k_1, \dots, k_p\}$  from each knowledge context associated with all information resources  $R_t$ .

As an example, without loss of generality, consider that the initial interest of an user contains one single keyword  $k_1$ , and that the user is querying two distinct information resources  $R_1$  and  $R_2$ . Two spreading interest fuzzy sets,  $F_1$  and  $F_2$ , are generated using  $d_1$  and  $d_2$  respectively. The evidence set  $ES_0(k)$  obtained from the “and/or” combination contains the keywords related to  $k_1$  in  $R_1$  “and/or” the keywords related to  $k_1$  in  $R_2$ .  $F_1$  is the perspective of  $R_1$  on  $k_1$  and  $F_2$  the perspective of  $R_2$  on  $k_1$ .

## 7.2 Reducing the Uncertainty of User Interests via Conversation

$ES_0(k)$  obtained from the “and/or” combination is a first cut at detecting the interests of a user in a set of information resources. But we can compute a more accurate interest set of keywords using an interactive algorithm, conversation process, between the user and the information resources being queried. Such conversation is an uncertainty reducing process based on Nakamura and Iwai’s [1987] IR system, and extended to Evidence Sets by Rocha [1999, 2001].

The final  $ES(k)$  obtained with this algorithm is a much less uncertain representation of user interests as projected on the knowledge contexts of the information resources queried, than the initial evidence set. The conversation algorithm lets the user reduce the uncertainty from the all the perspectives initially available. The initial evidence set  $ES_0(k)$  includes all associated keywords in several information resources. The conversation algorithm allows the user and her knowledge context to select only the relevant ones. Thus, the final  $ES(k)$  can be seen as a low-uncertainty linguistic category containing those perspectives on the user’s initial interest (obtained from the participating information resources) which are relevant to the user and her knowledge context. [Rocha, 1999, 2001].

Notice that this category is not stored in any location in the intervening knowledge contexts. It is temporarily constructed by integration of knowledge from several information resources and the interests of the user expressed in the interactive conversational process. Such a category is therefore a temporary container of knowledge integrated from and relevant for the user and the collection of information resources. Thus, this algorithm implements many of the, temporary, “on the hoof” [Clark, 1993] category constructions as discussed in Rocha[2000].

### 7.3 Recommending Documents

After construction of the final  $ES(k)$ , *TalkMine* must return to the user documents relevant to this category. Notice that every document  $n_i$  defines a crisp subset whose elements are all the keywords  $k \in \mathbf{K}$  which index  $n_i$  in all the constituent information resources. The similarity between this crisp subset and  $ES(k)$  is a measure of the relevance of the document to the interests of the user as described by  $ES(k)$ . This similarity is defined by different ways of calculating the subsethood [Kosko, 1993] of one set in the other. Details of the actual operations used are presented in Rocha[1999]. High values of these similarity measures will result on the system recommending only those documents highly related to the learned category.

### 7.4 Adapting Knowledge Contexts

From the many  $ES(k)$  obtained from the set of users of information resources, we collect the information used to adapt the KSP and semantic distance of the respective knowledge contexts. The scheme used to implement this adaptation is very simple: the more certain keywords are associated with each other, by often being simultaneously included with a high degree of membership in the final  $ES(k)$ , the more the semantic distance between them is reduced. Conversely, if certain keywords are not frequently associated with one another, the distance between them is increased. An easy way to achieve this is to have the values of  $N(k_i)$ ,  $N(k_j)$  and  $N_{\cap}(k_i, k_j)$  as defined in eq. 1, adaptively altered for each of the constituent  $r$  information resources  $R_r$ . After  $ES(k)$  is constructed and approximated by a fuzzy set  $A(x)$ , these values are changed according to:

$$N^t(k_i) = N^t(k_i) + w \cdot A(k_i), t = 1 \dots r, k_i \in R_0 \cup R_1 \cup \dots \cup R_r \quad (3)$$

and

$$N_{\cap}^t(k_i, k_j) = N_{\cap}^t(k_i, k_j) + w \cdot \min[A(k_i), A(k_j)], k = 1 \dots r, k_i, k_j \in R_0 \cup R_1 \cup \dots \cup R_r \quad (4)$$

Where  $w$  is the weight ascribed to the individual contribution of each user. The adaptation entailed by (3) and (4) leads the semantic distance of the knowledge contexts involved, to increasingly match the expectations of the community of users with whom they interact. Furthermore, when keywords with high membership in  $ES(k)$  are not present in one of the information resources queried, they are added to it with document counts given by equations (3) and (4). If the simultaneous association of the same keywords keeps occurring, then an

information resource that did not previously contain a certain keyword, will have its presence progressively strengthened, even though such keyword does not index any documents stored in this information resource.

---

## 8. Collective Evolution and Representation of Knowledge

*TalkMine* models the construction of linguistic categories. Such “on the hoof” construction of categories triggered by interaction with users, allows several unrelated information resources to be searched simultaneously, temporarily generating categories that are not really stored in any location. The short-term categories bridge together a number of possibly highly unrelated contexts, which in turn creates new associations in the individual information resources that would never occur within their own limited context.

Consider the following example. Two distinct information resources (databases) are searched using *TalkMine*. One database contains the documents (books, articles, etc) of an institution devoted to the study of computational complex adaptive systems (e.g. the library of the *Santa Fe Institute*), and the other the documents of a Philosophy of Biology department. I am interested in the keywords GENETICS and NATURAL SELECTION. If I were to conduct this search a number of times, due to my own interests, the learned category obtained would certainly contain other keywords such as ADAPTIVE COMPUTATION, GENETIC ALGORITHMS, etc. Let me assume that the keyword GENETIC ALGORITHMS does not initially exist in the Philosophy of Biology library. After I conduct this search a number of times, the keyword GENETIC ALGORITHMS is created in this library, even though it does not contain any documents about this topic. However, with my continuing to perform this search over and over again, the keyword GENETIC ALGORITHMS becomes highly associated with GENETICS and NATURAL SELECTION, introducing a new perspective of these keywords. From this point on, users of the Philosophy of Biology library, by entering the keyword GENETIC ALGORITHMS would have their own data retrieval system point them to other information resources such as the library of the *Santa Fe Institute* or/and output documents ranging from “The Origin of Species” to treatises on Neo-Darwinism – at which point they would probably bar me from using their networked database!

Given a large number of interacting knowledge contexts from information resources and users (see Figure 2), *TalkMine* is able to create new categories that are not stored in any one location, changing and adapting such knowledge contexts in an open-ended fashion. Open-endedness does not mean that *TalkMine* is able to discern all knowledge negotiated by its user environment, but that it is able to recombine all the semantic information (KSP and  $d$  described in section 4) of the intervening knowledge contexts in an essentially open-ended manner, as expected of CAS. The categories constructed by *TalkMine* function as a system of social, collective linguistic

recombination of distributed memory banks, capable of transferring knowledge across different contexts and thus creating new knowledge. In this way, *TalkMine* can adapt to an evolving environment and generate new knowledge given a sufficiently diverse set of information resources and users. Readers are encouraged to track the development of this system at <http://arp.lanl.gov>.

*TalkMine* is a collective recommendation algorithm because it uses the behavior of its users to adapt the knowledge stored in information resources. Each time a user queries several information resources, the category constructed by *TalkMine* is used to adapt those (section 7). In this sense, the knowledge contexts (section 4) of the intervening information resources becomes itself a representation of the knowledge of the user community. Indeed, knowledge contexts are distributed memory banks abiding by Clark's [1993] definition of connectionist devices [Rocha, 1999, 2000]. The adapted proximity associations between the tags/keywords instantiate an associative knowledge structure that stores the way its communities of users relate these tags in a distributed manner: an adaptive web. That is, the knowledge shared by distinct communities of users exists superposed in the proximity weights of each knowledge context. We can regard the adapted proximity data of knowledge contexts as a *shared knowledge structure*, in the sense of Richards et al [1998], for the entire collection of its users. We [Bollen and Rocha, 2000] have started to study the graph theoretical characteristics of these shared knowledge structures at LANL. Below the efforts to identify particular communities of users that negotiate with subsets of the entire share knowledge structure of an information resource.

---

## 9. Identification and Tagging of Communities of Users

As mentioned in section 8, as we let *TalkMine* adapt a knowledge context to become a proper shared knowledge structure of the entire collection of users of the respective information resource, we risk the curse of averages discussed in section 3, thus diminishing the diversity of knowledge stored. It becomes clear that heterarchies endowed with adaptive webs need to instantiate a community check mechanism, in which different weight is given to different users as they are used to adapt subsets of the knowledge structure. Clearly, when a keyword or document is not highly associated with a user's own knowledge context, then he should not be granted much weight in adapting the relevant associations with *TalkMine*. To instantiate this community check, we need first to identify the communities that co-exist in an information resource and then relate individual users to these.

## 9.1 Identifying Communities<sup>5</sup>

We have conducted a small experiment at LANL to demonstrate the viability of identifying user communities in DIS. The digital library at LANL maintains extensive web logs to keep track of users' retrieval patterns. A proximity matrix for journal titles had previously been generated based on the co-occurrence of journal titles in user retrieval paths in the February 1999 Research Library web logs (for a more detailed description of this technique, see [Bollen, Vandesompel, and Rocha, 1999]). A hierarchical cluster analysis was performed on this

**Table II:** Three of the clusters identified from research library web logs

Cluster 1	Cluster 2	Cluster 3
<b>Information Sciences</b>	<b>Molecular Biology</b>	<b>Nonlinear Dynamics</b>
Inform. Proc. Letters	Analyt. Chim	Prev. Med.
Inform. Sciences	Tetrahedr.	A Chaos
	J. Mol. Biol.	Siam J. Comp
	Tetrahedron	Quant. Res.
	Tetrahedron L.	Siam Rev.
		Arch. Env. C.

matrix, revealing a number of persistent journal clusters. Three clusters were selected for further analysis based on their size and content see table II.

A user community was derived by determining a set of users that had frequently downloaded articles published in the journals in the journal clusters. For each of those communities, a list of the 20 most frequently downloaded articles and their associated keywords (via matrix A in section 4) was compiled. From the KSP relation (eq. 2 in section 4), semantic proximity values were calculated for all pairs of keywords, resulting in a keyword graph for each user cluster. Such a graph represents the shared knowledge structure of the community of users associated with each cluster. Figure 7 shows a subgraph for the community of users in cluster 3.

---

<sup>5</sup> The research in this subsection was conducted with Johan Bollen at LANL. More details in [Bollen and Rocha, 2000].



for extracting and inferring relations of expected contextual usage of words in documents (or smaller passages of discourse). It is used basically to reduce the dimensionality of the original matrix by collapsing related keywords into a new dimension (a singular vector). Each singular vector is a linear combination of the original keywords, and is supposed to represent a particular sense of a group of keywords<sup>6</sup>. The importance of each singular vector is given by the singular values the decomposition produces. LSA is based on selecting the vectors associated with the highest values, while discarding lower ones. The set of singular vectors with higher values approximates the original matrix, but with a smaller number of dimensions. This dimensional collapse, however, takes into account more than direct associations: two keywords being associated with the same document. It considers all higher order relations, that is, indirect chains of associations. The lower-dimensionality LSA approximation of the original matrix contains the essential semantic relationships between keywords for given a set of documents.

To obtain a large keyword-document matrix associated with each cluster, we start from the shared knowledge graphs obtained in section 9.1, and expand them via the KSP proximity (eq.2) for the entire information resource. This way, we obtain a sufficiently large matrix of keywords and indexed documents which contains the set keywords and documents the community of users trades. We then perform the LSA on this matrix, producing an LSA space containing the relevant semantic associations represented as singular vectors. In such a space, a keyword is a vector with specific singular value contributions. Likewise, a document becomes a vector defined by the centroid of the keywords that index it. Thus, we gain the ability to discern if a given keyword or document is close to another via the cosine of the angle of their vectors in the LSA space.

Furthermore, given the history of IR stored in the knowledge contexts of users (section 4.2), we can define a user as the set of documents he has retrieved. In the LSA space, a user can then be defined also as the centroid vector of these documents. Finally, the community of users identified in 9.1, again defines a set of vectors in the LSA Space which can be approximated by a centroid vector. This way, any new user that enters the system can be compared to existing communities. The cosine of the angle between the user vector and the community vector, defines the closeness of the user to a community. This way, we obtain a mechanism to decide if a user can be tagged as belonging to any community associated with an information resource. If the user is tagged as belonging to a community, his retrieval behavior can be used to adapt the proximity information of the keywords and documents close to the community vector.

---

<sup>6</sup> For instance, there are 3 main contexts for the word Java: the coffee bean, the computer language, and the island. With LSA, 3 dimensions could be created to accommodate these contexts. One could be a linear combination of Java, with other words such as Sun Microsystems, C, Computer, etc.

---

## 10. Adaptive Webs for Socially Intelligent Heterarchies

*TalkMine* is an approach to produce adaptive webs for heterarchies whose knowledge bases can evolve with the expectations of its members. It establishes a different kind of human-machine interaction in IR, as the machine side rather than passively expecting the user to pull information, effectively pushes relevant knowledge. This pushing is done in the conversation algorithm of *TalkMine*, where the user, or her browser automatically, selects the most relevant subsets of this knowledge. Because the knowledge of communities is represented in adapting information resources, and the interests of individuals are integrated through conversation leading to the construction of linguistic categories and adaptation, *TalkMine* achieves a more natural, biological-like, knowledge management of DIS, capable of coping with the evolving knowledge of user communities.

With *TalkMine*, the information systems of a heterarchy become adaptive webs which represent the knowledge of the organization (its identity) as proximity associations between tags and its components. Notice that this knowledge is preserved even as individual elements of the organization (its agents) abandon it. Thus, the adaptive web functions as a distributed learning device for the social organization. That is, knowledge stored adapts according to the behavior of its social agents. Finally, as the information systems of the heterarchy adapt to its users, we gain the ability to identify emergent communities of practice, and tag elements of the heterarchy appropriately.

By endowing heterarchies with the capabilities of an adaptive web, we expect them to be able to cope with their environments in a more responsive manner. On the one hand, an adaptive web preserves the knowledge of a heterarchy in its memory, and on the other, it allows this knowledge to evolve to the expectations of new agents. This balance of stability and innovation should be advantageous for the success of a heterarchy in a changing environment. Furthermore, the ability to automatically identify and tag members and users of heterarchies, facilitates the control of a heterarchy by the emergent control hierarchies it is a member of as discussed in section 1.

---

## References

Balabanović, M. and Y. Shoham [1997]. "Content-based, collaborative recommendation." *Communications of the ACM*. March 1997, Vol. 40, No.3, pp. 66-72.

- Berry, M.W., S.T. Dumais, and G.W. O'Brien [1995]. "Using linear algebra for intelligent information retrieval." *SIAM Review*. Vol. 37, no. 4, pp. 573-595.
- Bollen, J. and F. Heylighen [1998]. "A system to restructure hypertext networks into valid user models." *The New Review of Hypermedia and Multimedia*. Vol. 4.
- Bollen, J., H. Vandesompel, L.M. Rocha [1999]. "Mining associative relations from website logs and their application to context-dependent retrieval using Spreading Activation." In: *Workshop on Organizing Web Space (WOWS), ACM Digital Libraries 99, August 1999, Berkeley, California*. In press.
- Bollen, J., L.M. Rocha [2000]. "An Adaptive Systems Approach to the Implementation and Evaluation of Digital Library Recommendation Systems." In: *Proceedings of the European Conference on Digital Libraries 2000*. Lectures Notes on Artificial Intelligence, Springer-Verlag, In press.
- Brusilovsky, P., A. Kobsa, and J. Vassileva (Eds.) [1998]. *Adaptive Hypertext and Hypermedia Systems*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Cariani, Peter [1989]. *On The Design of Devices With Emergent Semantic Functions*. PhD Dissertation. State University of New York at Binghamton.
- Chakrabarti, S. et al [1999]. "Mining the Web's link structure." *Computer*. Vol. 32, No.8, pp. 60-67.
- Chislenko, Alexander [1998]. "Collaborative information filtering and semantic transports." In: . . WWW publication: <http://www.lucifer.com/~sasha/articles/ACF.html>.
- Clark, A. [1993]. *Associative Engines: Connectionism, Concepts, and Representational Change*. MIT Press.
- Clark, A. [1996]. "Happy couplings: emergence and explanatory interlock." In: *The Philosophy of Artificial Life*. M. Boden (ed.). Oxford University Press, pp. 262-281.
- Clark, A. [1999]. "Leadership and influence: the manager as coach, nanny, and artificial DNA." In: *The Biology of Business: Decoding the Natural Laws of Enterprise*. J.H. Clippinger III (Ed.). Jossey-Bass Publishers, pp.47-66.
- Clippinger, J.H. [1999]. "Tags: the power labels in shaping markets and organizations." In: *The Biology of Business: Decoding the Natural Laws of Enterprise*. J.H. Clippinger III (Ed.). Jossey-Bass Publishers, pp.67-88.
- Conrad, Michael [1990]. "The geometry of evolutions." *BioSystems*. Vol. 24, pp.61-81.
- Eklund, J. [1998]. "The value of adaptivity in hypermedia learning environments: a short review of empirical evidence." In: *Proceedings of the 2nd Workshop on Adaptive Hypertext and Hypermedia (Hypertext 98)*. P. Brusilovsky and P. de Bra (Eds.). pages 13-21, Pittsburgh, USA, June 1998.
- Galvin, F. and S.D. Shore [1991]. "Distance functions and topologies." *The American Mathematical Monthly*. Vol. 98, No. 7, pp. 620-623.

- Good, N. et al [1999]. "Combining collaborative filtering with personal agents for better recommendations." In: *Proceeding of the National Conference on Artificial Intelligence*. . Orlando, Florida, July 1999. AAAI, pp. 439-446.
- Grabher, G. [2001]. "Spaces of creativity: Heterarchies in British advertising". *Environment and Planning A*. Vol. 33. In Press.
- Harman, D. [1994]. "Overview of the 3rd Text Retrieval Conference (TREC-3)." In: *Proceedings of the 3rd Text Retrieval Conference*. . Gaithersburg, Md, November 1994..
- Herlocker, J.L. [1999]. "Algorithmic framework for performing collaborative filtering." In: *Proceedings of the 22nd International Conference on Research and Development in Information Retrieval*. . Berkeley, California, August 1999. ACM, p.. 230-237.
- Heylighen, Francis [1999]. "Collective Intelligence and its Implementation on the Web: Algorithms to Develop a Collective Mental Map." *Computational & Mathematical Organization Theory*. Vol. 5, no. 3, pp. 253-280.
- Hill, W. et al [1995]. "Recommending and evaluating choices in a virtual community of use." In: *Conference on Human Factors in Computing Systems (CHI'95)*. Denver, May, 1995.
- Holland, J.H. [1995]. *Hidden Order: How Adaptation Builds Complexity*. Addison-Wesley.
- Johnson, N., S. Rasmussen, C. Joslyn, L. Rocha, S. Smith, and M. Kantor [1998]. "Symbiotic intelligence: self-organizing knowlede on distributed networks, driven by human interaction." In: *Proceedings of the 6th International Conference on Artificial Life*. C. Adami, R. K. Belew, H. Kitano, C. E. Taylor (Eds.). MIT Press, pp. 403-407.
- Kanerva, P. [1988]. *Sparse Distributed Memory*. MIT Press.
- Kannan, R. and S. Vempala [1999]. "Real-time clustering and ranking of documents on the web." In: . . Unpublished Manuscript.
- Kauffman, S. [1993]. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford university Press.
- Kessler, M.M. [1963]. "Bibliographic coupling between scietific papers." *American Documentation*. Vol. 14, pp. 10-25.
- Kleinberg, J.M. [1998]. "Authoritative sources in a hyperlinked environment." In: *Proc. of the the 9th ACM-SIAM Symposium on Discrete Algorithms*. . pp. 668-677.
- Klir, G.J. and B. Yuan [1995]. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice Hall.
- Klir, George, J. [1993]. "Developments in uncertainty-based information." In: *Advances in Computers*. M. Yovits (Ed.). Vol. 36, pp. 255-332.

- Koput, K.W. and W.W. Powell [2000]. "Science and strategy: Organizational evolution in a knowledge-intensive field". Presented at workshop on *Heterarchies: Distributed Intelligence and the Organization of Diversity*, organized by David Stark at the Santa Fe Institute, October 13-14, 2000
- Kosko, B. [1992]. *Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence*. Prentice-Hall.
- Kostan, J.A. et al [1997]. "GroupLens: applying collaborative filtering to usenet news." *Communications of the ACM*. V. 40, No. 3, pp. 77-87.
- Krulwich, B. and C. Burkey [1996]. "Learning user information interests through extraction of semantically significant phrases." In: *Proceedings of the AAAI Spring Symposium on Machine Learning in Information Access*. . Stanford, California, March 1996.
- Landauer, T.K., P.W. Foltz, and D. Laham [1998]. "Introduction to Latent Semantic Analysis." *Discourse Processes*. Vol. 25, pp. 259-284.
- Lang, K. [1995]. "Learning to filter news." In: *Proceedings of the 12th International Conference on Machine Learning*. . Tahoe City, California, 1995.
- Miyamoto, S. [1990]. *Fuzzy Sets in Information Retrieval and Cluster Analysis*. Kluwer.
- Moukas, A. and P. Maes [1998]. "Amalthea: an evolving multi-agent information filtering and discovery systems for the WWW." *Autonomous agents and multi-agent systems*. Vol. 1, pp. 59-88.
- Nakamura, K. and S. Iwai [1982]. "A representation of analogical inference by fuzzy sets and its application to information retrieval systems." In: *Fuzzy Information and Decision Processes*. M.M. Gupta and E. Sanchez (Eds.). North-Holland, pp. 373-386.
- Pattee, Howard H. [1973]. "The physical basis and origin of control." In: *Hierarchy Theory: The Challenge of Complex Systems*. H.H. Pattee (Ed.). George Braziller, pp.71-108.
- Pattee, Howard H. [1976]. "The role of instabilities in the evolution of control hierarchies." In: *Power and Control*. T.R. Burns and W. Buckley (Eds.). Sage, pp. 171-184.
- Resnick, P. et al [1994]. "GroupLens: An open architecture for collaborative filtering of netnews." In: *Proceedings of the ACM Conference on Computer-Supported Cooperative Work*. . Chapel Hill, North Carolina, 1994.
- Richards, D., B.D. McKay, and W.A. Richards [1998]. "Collective choice and mutual knowledge structures." *Advances in Complex Systems*. Vol. 1, pp. 221-236.
- Rocha, Luis M. and Johan Bollen [2000]. "Biologically motivated distributed designs for adaptive knowledge management." In: *Design Principles for the Immune System and Other Distributed Autonomous Systems*. Cohen I. And L. Segel (Eds.). Santa Fe Institute Series in the Sciences of Complexity. Oxford University Press. In Press.

- Rocha, Luis M. [1994]. "Cognitive categorization revisited: extending interval valued fuzzy sets as simulation tools concept combination." *Proc. of the 1994 Int. Conference of NAFIPS/IFIS/NASA*. IEEE Press, pp. 400-404.
- Rocha, Luis M. [1996]. "Eigenbehavior and symbols." *Systems Research*. Vol. 13, No. 3, pp. 371-384.
- Rocha, Luis M. [1997a]. *Evidence Sets and Contextual Genetic Algorithms: Exploring Uncertainty, Context and Embodiment in Cognitive and biological Systems*. PhD. Dissertation. State University of New York at Binghamton. UMI Microform 9734528.
- Rocha, Luis M. [1997b]. "Relative uncertainty and evidence sets: a constructivist framework." *International Journal of General Systems*. Vol. 26, No. 1-2, pp. 35-61.
- Rocha, Luis M. [1998]. "Selected self-organization and the Semiotics of Evolutionary Systems." In: *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-Organization*. S. Salthe, G. Van de Vijver, and M. Delpos (eds.). Kluwer Academic Publishers, pp. 341-358..
- Rocha, Luis M. [1999]. "Evidence sets: modeling subjective categories." *International Journal of General Systems*. Vol. 27, pp. 457-494.
- Rocha, Luis M. [2000]. "Adaptive Recommendation and Open-Ended Semiosis ." *International Journal of Human-Computer Studies*. (In Press).
- Rocha, Luis M. [2001]. "TalkMine: a Soft Computing Approach to Adaptive Knowledge Recommendation". In: Loia, V. and S. Sessa (Eds.) .*Soft Computing Agents: New Trends for Designing Autonomous Systems*. Studies in Fuzziness and Soft Computing. Physica-Verlag, Springer. In Press.
- Rosen, R. [1969]. "Hierarchical organization in automata-theoretic models of biological systems." In: *Hierarchical structures*. LL. Whyte, A. Wilson, and D. Wilson (Eds.). Elsevier, pp. 179.
- Shafer, G. [1976]. *A Mathematical Theory of Evidence*. Princeton University Press.
- Shardanad, U. and P. Maes [1995]. "Social information filtering: Algorithms for automating 'word of mouth'." In: *Conference on Human Factors in Computing Systems (CHI'95)*. . Denver, May, 1995.
- Simon, H.H. [1962]. "The architecture of complexity." *Proc. Am. Philos. Soc.* Vol. 106, pp. 467.
- Simon, H.H. [1973]. "The organization of complex systems." In: *Hierarchy Theory: The Challenge of Complex Systems*. H.H. Pattee (Ed.). George Braziller, pp.1-27.
- Small, H. [1973]. "Co-citation in the scientific literature: a new measure of the relationship between documents." *Journal of the American Society for Information Science*. Vol. 42, pp. 676-684.
- Stark, D. [1999]. "Heterarchy: Distributing Authority and Organizing Diversity." In: *The Biology of Business: Decoding the Natural Laws of Enterprise*. J.H. Clippinger III (Ed.). Jossey-Bass Publishers, pp.153-179.
- Turksen, B. [1986]. "Interval valued fuzzy sets based on normal forms." *Fuzzy Sets and Systems*. Vol. 20, pp. 191-210.

- Turksen, I.B. [1996]. "Non-specificity and interval-valued fuzzy sets." *Fuzzy Sets and Systems*. Vol. 80, pp. 87-100.
- Umerez, J. and A. Moreno [1995]. "Origin of life as the first MST- Control hierarchies and interlevel relation." *World Futures: The Journal of General Evolution*. Vol. 45, pp. 139-154.
- van Gelder, Tim [1991]. "What is the 'D' in 'PDP': a survey of the concept of distribution." In: *Philosophy and Connectionist Theory*. W. Ramsey et al. Lawrence Erlbaum.
- Watts, D. [1999]. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton University Press.
- Yager, R.R. [1979]. "On the measure of fuzziness and negation. Part I: membership in the unit interval." *Int. J. of General Systems*. Vol. 5, pp. 221-229.
- Yager, R.R. [1980]. "On the measure of fuzziness and negation: Part II: lattices." *Information and Control*. Vol. 44, pp. 236-260.
- Zadeh, Lofti A. [1965]. "Fuzzy Sets." *Information and Control*. Vol. 8, pp. 338-353.