

# Research Report

## Handling Complexity With Self-Organizing Fractal Semantic Networks

Jürgen Klenk\*, Gerd Binnig\*, Günter Schmidt‡

\*IBM Research  
Zurich Research Laboratory  
8803 Rüschlikon  
Switzerland

‡Definiens AG  
Rindermarkt 7  
80331 Munich  
Germany

### LIMITED DISTRIBUTION NOTICE

This report has been submitted for publication outside of IBM and will probably be copyrighted if accepted for publication. It has been issued as a Research Report for early dissemination of its contents. In view of the transfer of copyright to the outside publisher, its distribution outside of IBM prior to publication should be limited to peer communications and specific requests. After outside publication, requests should be filled only by reprints or legally obtained copies (e.g., payment of royalties).



Research  
Almaden • Austin • Beijing • Delhi • Haifa • T.J. Watson • Tokyo • Zurich

# Handling Complexity With Self-Organizing Fractal Semantic Networks

Jürgen Klenk\*, Gerd Binnig\*, Günter Schmidt‡

\*IBM Research, Zurich Research Laboratory, 8803 Rüschlikon, Switzerland

‡ Definiens AG, Rindermarkt 7, 80331 Munich, Germany

*Abstract* - This paper is concerned with aspects of complexity related to the multitude of information available today. We introduce the *Self-Organizing Fractal Semantic Network*, a model to handle complexity by taking into account aspects of knowledge and thinking as well as self-organization. Its basic building blocks and processes are described in detail, and *Classification* and *Segmentation* are identified as the fundamental processes for driving a self-organizing network on a local scale. An example is given to illustrate the basic concept, and further research directions are highlighted.

## INFORMATION AND SELF-ORGANIZING NETWORKS

In today's globally networked world people are facing dramatically increasing complexity. Part of this complexity is due to the rapidly growing amount of information that is becoming available online. Because knowing more usually leads to better decisions than knowing less (even if the "less" seems clearer and more definite), complex problems require a complex analysis [1], i.e. being aware of the multitude of relevant information available for the decision process, and being able to select the right pieces of information from this multitude. Despite the fact that dealing with complexity is one of the strengths of the human mind, because of the current inflation of information there is a growing need for tools that intelligently assist people in handling complexity.

Current research and state-of-the-art tools for handling complexity are concentrated around topics such as searching and indexing, data mining, classification and clustering, summarization and information extraction, and case-based reasoning, among others. The underlying mathematical models come from disciplines such as statistics, logic, and optimization theory. Depending on the complexity these methods achieve their tasks with varying degree of success. In particular, search on the extremely complex Internet is still rather unsatisfactory in many cases due to both low precision and low recall. This is not too surprising, as most tools do not exploit the possibly most important aspect of the human mind when dealing with complexity, the network-like structure of knowledge and thinking.

Network-based approaches [2] such as Neural Networks [3] and Semantic Networks [4,5] do concentrate on this aspect. The theory of Neural Networks even attempts to re-invent the evolution of the brain (in a rather short period of time). From a scientific point-of-view this is certainly an exciting approach. However, we believe that Semantic Networks provide a more powerful approach for the development of tools that can deal with the exploding complexity of today's world, as they jump into this evolution on a higher semantic level. In this case the evolving network does not have to re-invent human knowledge and thinking to the extent that these can be built right into the network.

While the network-like structure of knowledge and thinking is one important aspect when dealing with complexity, complexity theory (as originally developed in the biological sciences as a means of understanding how organic entities and communities form) also focuses on the aspect of self-organization. The problem of self-organization itself is extremely complex because the number of causal factors and the degree of interdependence between them is too great to permit the use of mechanistic prescriptive models to solve it. Instead, self-organization of a complex system is driven on a small (local) scale by a collection of (typically) relatively simple processes, the causal factors. The global outcome very much depends on these local processes and the initial state of the system. Despite the fact that individual local processes are deterministic the global outcome usually cannot be predicted because the system exhibits chaotic behavior due to its non-linearity caused by the interdependence of its local processes [6].

When combining the ideas and theories about self-organization with the ideas and theories about knowledge and thinking, that is, when dealing with a Self-Organizing Semantic Network, one can model complex problems such as cognition and learning, which relate to the internal state of an entity, as well as the behavior of such entities in communities or collectives, which relate to the external state of the entities. However, in modeling this external state one really deals with Self-Organizing Semantic Networks on two different scales, the internal Self-Organizing Semantic Network of the community or collective, and the internal Self-Organizing Semantic Networks of the entities that make up the collective. This idea can be extended quite naturally to more than two scales. In fact, since most Semantic Networks typically have a hierarchical structure, all Self-Organizing Semantic Networks actually consist of a number of smaller Self-Organizing Semantic Networks across different scales.

If one wants to locally drive the self-organization of the network with a few simple processes that are the same on all scales (such as in cellular automata [7]), then the network must look the same everywhere, so that these processes can be applied at every location in the network. In other words, the Self-Organizing

Semantic Network must be self-similar across all scales by being constructed out of a few fundamental building blocks and construction principles. Mathematical objects that are self-similar across hierarchical scales are called fractals, which is why we call the model described in this paper a Self-Organizing Fractal Semantic Network.

## DEFINITION OF A SELF-ORGANIZING FRACTAL SEMANTIC NETWORK

We begin with the basic definitions of semantic and hierarchical networks.

**Definition 1:** *A Semantic Network is a directed or non-directed graph with the additional property that its nodes and links carry semantic information. Nodes and links of a Semantic Network are called Semantic Units. The semantic information provides a particular meaning for a Semantic Unit. Frequently, semantic information is encoded in natural language, i.e. specific names or processes are used to label Semantic Units.*

**Definition 2:** *A hierarchical network is a directed or non-directed graph with the additional property that some of its links carry scaling information. This means that one of the nodes connected to a scaling link is at a higher level of hierarchy than the other node.*

It should be noted that this definition does not yield an absolute value for the level of hierarchy, i.e. it does not assign to every node an integer that corresponds to its level of hierarchy. Instead it gives a relative definition for the level of hierarchy. While this approach is more general, it can cause conflicts because one can have cycles (loops) of hierarchical links. This conflict exists only on a global scale and can be resolved if one considers only local neighborhoods of the whole network at any one time. Local neighborhoods are introduced by the notion of a topology.

**Definition 3:** *A topological network is a directed or non-directed graph with the additional property that for every node and every link one or several local neighborhoods are defined. The local neighborhoods of a node or a link are sets consisting of this node or link and other nodes or links of the network.*

In many cases topological networks are obtained by assigning a weight (a value between 0 and 1) to every link of the network. The negative logarithm of this weight then yields a metric or distance function on the network. The local neighborhoods are defined with the help of some threshold mechanism applied to this metric. Every metric determines a topology, but the converse is not true; therefore, the above definition is more general than the method of assigning weights.

In the special case of Semantic Networks it is often required that a Semantic Unit be used both as a node and as a link. For example, a Semantic Unit labeled *friendship* can be viewed both as a node ("*friendship* is in particular a *relation*", here it is a node connected hierarchically to the node *relation*) and as a link ("*friendship* between two *people*", here it is a link between the two nodes representing the two *people*). This gives rise to the following definition.

**Definition 4:** *A higher-order network is a directed or non-directed graph in which links can at the same time be nodes. This means that a link can connect two nodes, one node and one link, or two links.*

In the next definition we capture what we mean by a fractal network. We will not give an exact mathematical definition (for ideas on how to do this see [8,9]) but rather a working definition, which will suffice for the scope of this paper. In particular, our formulation will allow us more easily to define the Self-Organizing Fractal Semantic Network and to understand the example given in the last section. But this definition will not allow us to explore in more detail the fractal structure of the network. In particular, we will not be able to give a definition for the fractal dimension, which, according to separate studies, seems to be related to the more subjective quantity of complexity. This topic will be covered in a forthcoming paper.

**Definition 5:** *A fractal network is a hierarchical network with the additional property that all of its nodes and links are derived from a small set of basic building blocks. In this sense a fractal network exhibits a self-similar structure because it looks the same everywhere on all levels of hierarchy.*

The following definition deals with the processes that perform the self-organization of the network.

**Definition 6:** *A (locally) self-organizing network is a directed or non-directed graph with the additional property that at least some of its nodes and links are connected to one or several (local) processes out of a set of processes. A process is an algorithm that performs a transformation of the network, while a local process is an algorithm that performs a transformation of the network only in the local neighborhood of the node or link it is connected to. The (local) processes are triggered by the state of the node or link they are connected to, which in turn is a function of the entire network.*

For practical purposes the state of a node or link is often only a function of its local neighborhood. For example, when dealing with Semantic Networks a Semantic Unit is often connected to attributes, and the values of these attributes determine the state of the Semantic Unit.

It is conceivable that the processes themselves form a hierarchical network. This is a topic under current investigation. The fundamental processes will be covered in Sections 4 and 5.

After making all of the above definitions we are now in the position to define a Self-Organizing Fractal Semantic Network, our fundamental model that we use to study complexity problems.

**Definition 7:** *A (Locally) Self-Organizing Fractal Semantic Network is a hierarchical, topological, higher-order, fractal, (locally) self-organizing Semantic Network.*

The following diagram shows a Self-Organizing Fractal Semantic Network. Nodes are depicted as spheres, links as cylinders, and processes as Janus heads (see next section for details). The hierarchical structure is clearly visible from the fact that nodes contain internal networks. Note too that some links are nodes, as their cylinders have spheres in their center that connect them to other nodes or links.

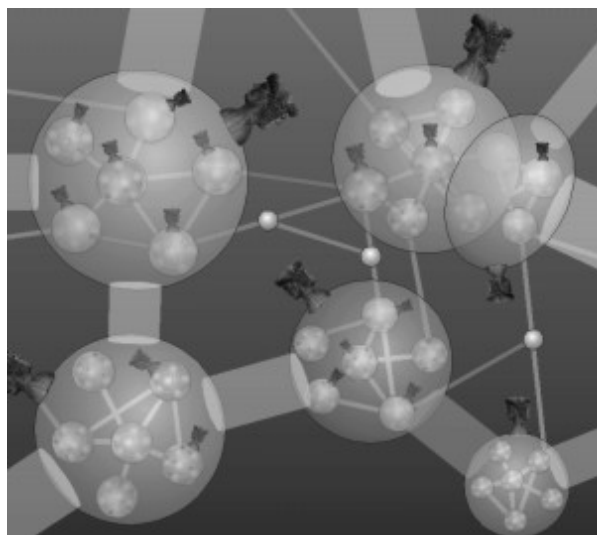


Figure 1: A sample Self-Organizing Fractal Semantic Network

## THE BASIC BUILDING BLOCKS

In the following sections we will define a set of basic building blocks and basic processes, which will be used to construct Self-Organizing Fractal Semantic Networks. We will also look at mechanisms or functions that determine the states of Semantic Units.

As specified in Definition 1, nodes and links of the network are called Semantic Units. All Semantic Units are subdivided into concepts and instances as is usual [10]. We further subdivide nodes into Information Units, Attribute Units, and Process Units or Janus Units. In our model we use the metaphor of a Janus [11] (pl. Jani) for a process. In Roman mythology, Janus is a two-faced god looking in opposite directions, i.e. taking into account a multitude of perspectives. This is exactly what a process in a fractal network does, if one interprets the two directions as the internal, lower-scale and the external, higher-scale parts of the network with respect to the Semantic Unit that the process is connected to. Information Units are general elements that can represent concepts or instances, and they are identified by specific names. Attribute Units are identified by specific names and values, which can be set, retrieved, or computed.

All links of the network are either scaling or non-scaling. Standard inheritance principles [10] are defined across all scaling links, making use of the network's topology or neighborhood concept. Links are further subdivided into Comparison Units, Interaction Units, Description Units, and Controller Units. Non-scaling Comparison Units allow us to describe the degree of similarity or dissimilarity of two Semantic Units, while scaling Comparison Units allow us to describe how close one Semantic Unit comes to being an instance of another Semantic Unit, or how close one Semantic Unit comes to being a special case of another Semantic Unit. Non-scaling Interaction Units allow us to describe the type of interaction of two Semantic Units, while scaling Interaction Units allow us to describe the role one Semantic Unit plays as part of another Semantic Unit. Description Units connect Semantic Units to their Attribute Units, which describe the Semantic Units in more detail. Finally, Controller Units connect Semantic Units to their Janus Units, which in turn control and act upon the Semantic Units' local neighborhoods.

The following diagram shows how the basic building blocks are used to construct a network. Note that each building block labeled "Semantic Unit" can be replaced with any basic building block. Information Units do not appear in this diagram, as there is no restriction on their use. In practice most of the building blocks labeled "Semantic Unit" are Information Units.

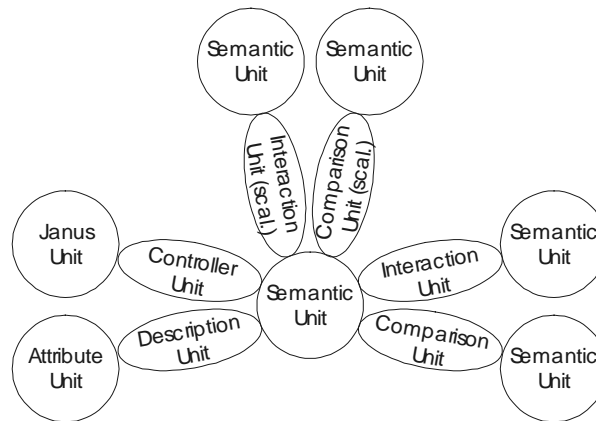


Figure 2: The basic building blocks

In the Self-Organizing Fractal Semantic Network each Semantic Unit has a specific meaning, which it receives indirectly through its multitude of links to other Semantic Units. Thus, in our model meaning is defined (and is definable) not in an absolute way, but only in a relative way, in accordance with Minsky's philosophy [12].

The state of a Semantic Unit is a (possibly complex) function of all Semantic Units of its local neighborhood. A simple example for the state of a Semantic Unit is the state of *usefulness* of the Information Unit

representing a *soccer ball*. The *soccer ball's* local neighborhood with respect to the state of its *usefulness* may consist of the Attribute Unit *size* and the Attribute Unit *pressure*, together with their respective values. The state of the *soccer ball's usefulness* can then be computed from the current values of its two Attribute Units, preferably in terms of fuzzy set theory.

A more complex example is the *mood* state of a *human being*. Here clearly the Attribute Units describing the *human being*, such as *height* and *weight*, among others, are not sufficient to describe his *mood* state. Psychologists know that the whole context within which the *human being* lives affects his *mood* state. There are many theories about what factors should and should not be taken into account when determining this context, and what kind of influence each of these factors has on the *mood* state. In our model it is precisely this context that makes up a *human being's* local neighborhood with respect to his *mood* state, and it is precisely the kind of influence each of these factors has that determines the *mood* state function.

## THE BASIC PROCESSES

This section will cover the basic processes that perform the self-organization task of the network on a local scale. To motivate the choices for the basic processes in our Self-Organizing Fractal Semantic Network, we continue with our examples from the last section.

It is important to note that there are Janus Units that influence the *usefulness* of the *soccer ball*, or the *mood* state of the *human being*. Semantic Units that are in the local neighborhood of the *soccer ball* or the *human being* typically control and trigger these Janus Units. If the *soccer ball* becomes *useless* because it goes flat (caused by a Janus Unit), it may get thrown away (caused by another Janus Unit) and thus removed from the network, and a new *soccer ball* may be purchased (caused by yet another Janus Unit) and thus a new Information Unit representing the new *soccer ball* is created. As for the *human being*, he may identify himself as the typical representative of a certain *group*, and may consequently join this *group* to improve his *mood* state, or he may found a new *interest group* and invite others to join it. To this extent he may even acquire new skills or knowledge.

These examples illustrate the basic processes required in our model. We have Janus Units that are able to create new Semantic Units, to modify or destroy existing Semantic Units (and even themselves), to create links between Semantic Units, to classify or identify Semantic Units as other Semantic Units, and to create new groups or segments in the network. Other Janus Units must be able to set, retrieve, or compute values of Attribute Units and to determine the states of Semantic Units. Finally, there are Janus Units that are able to perform a learning task in the form of knowledge acquisition or restructuring.

As can be seen from the Janus Units described above, the processes carried out by Janus Units can range from generic to specific, meaning that they use generic properties of the basic building blocks that make up the self-similar structure of the network, or very specific properties of the local neighborhood of a certain Semantic Unit, respectively. Therefore, some Janus Units can be connected to any Semantic Unit, while others require the presence of particular Semantic Units in the local neighborhood of the Semantic Unit they are connected to. Clearly the Janus Units that simply create, modify, or destroy Semantic Units (including the ones that create links, as this is a special case of creating Semantic Units) perform very generic tasks. Therefore, they can be connected to any Semantic Unit. However, they are usually not triggered directly but rather invoked as parts of more complex processes such as Classification or Segmentation.

The Janus Units that perform the evaluation of Attribute Units' values typically have a set of mathematical tools at hand from areas such as fuzzy set theory, statistics, geometry, topology, and algebra, among others. Therefore, these Janus Units are more specific as they can be applied to only Attribute Units whose values satisfy certain type constraints. Finally, the Janus Units that determine the states of Semantic Units are even more specific, as their processes might only be applicable to one or a small group of Semantic Units.

In the next section we will discuss the processes of Classification and Segmentation in more detail, as they are the fundamental drivers of the self-organization of the network.

## **CLASSIFICATION AND SEGMENTATION**

The process of Classification stands for the common task of comparing one Semantic Unit to others. The goal here is to find comparable Semantic Units in the sense that they are alike, can perform similar tasks, have similar goals, are more general or more specific, are constituents or groups, or are in similar states, among other things. Psychology tells us that this is a very important task as human beings are constantly in search of their identities by comparing themselves to and (even more importantly) differentiating themselves from others. In our model the process of Classification is performed through extensive local neighborhood analyses. This means that the degree of similarity of two Semantic Units is determined by the degree of similarity of their local neighborhoods with respect to the above comparison factors. As with determining the status of a Semantic Unit, when comparing Semantic Units it is not enough to simply take into account the values of the Attribute Units of these Semantic Units. Instead, the topology of the network, i.e. the entire local neighborhood structures of the Semantic Units, must be considered. Therefore, the process of Classification deals with the more general task of finding similar structures and not just similar values of Attribute Units. Because of the self-similar structure of the network, this Classification process can be implemented in a generic way, thus allowing the Janus Unit representing the Classification process to be used throughout the entire network.

While Classification focuses on finding similar structures among Semantic Units, Segmentation focuses on grouping Semantic Units according to similarities found during Classification. There are two main types of groupings that our model deals with, corresponding to the two types of scaling links we defined, the scaling Comparison Units and the scaling Interaction Units. These two types of groupings correspond to again well-known results from psychological studies, people's desire to categorize and organize information and knowledge and people's desire to form working groups to better achieve a common goal [13,14]. While the categorization and organization of knowledge predominantly uses comparing and contrasting mechanisms, working groups are formed according to skills and common goals. Here it is often the case that diversity is more important than similarity, because working groups are often more successful if they consist of group members with the right mix and variety of skills.

From a process point-of-view, the results of Classification processes determine and trigger Segmentation processes. In particular, the Classification results determine which new links from Semantic Units to other Semantic Units representing categories or groups should be created, and trigger the appropriate Segmentation processes that create these links. The Classification results also determine which new categories or groups should be created or formed if a number of Semantic Units have been classified as being similar or having a common goal and thus should be united/joined/combined/integrated in such new categories or groups. The triggered Segmentation processes then create these new categories or groups and also create all links to members of these categories or groups. Again, because of the self-similar structure of the network, this Segmentation process can be implemented in a generic way, thus allowing the Janus Unit representing the Segmentation process to be used throughout the entire network.

Finally the creation of new Semantic Units during the Segmentation process triggers a new Classification process, this time based on the new network structure and topology. This sequence of Classification and Segmentation, which continuously determines and changes the neighborhood structure and thus influences the states of the Semantic Units, is the main driver of the self-organization of the network.



## AN EXAMPLE

In this section we illustrate how a Self-Organizing Fractal Semantic Network can be used to model a complex economical and ecological society. Our model is derived from a real-world situation in a rural area in southern Germany and is based on the following assumptions:

1. A number of farmers are running small agricultural firms. These firms are influenced by standard ecological and economical factors.
2. Agricultural goods are traded on a market following standard market rules.
3. There is a dominant (almost monopolistic) market player, a consortium that purchases almost all agricultural products and tries to dictate prices. The consortium performs certain refinement processes such as making dairy products from milk or running a slaughterhouse.
4. Farmers have an internal knowledge about producing and selling agricultural goods. Some of the farmers also have a good understanding of the economic situation in marketplaces.
5. The farmers as well as the consortium are connected to the marketplace. Each has a limited influence on the marketplace, depending on their relative size and importance.
6. There are consumers of varying size, such as supermarket chains, butchers, stores, and individuals, which purchase the refined agricultural products.
7. The goals of the farmers are to maximize their revenue and at the same time to minimize their workload. Their states are determined by the degree of success or failure to achieve their goals.
8. The goals of the consortium are to maximize its revenue by minimizing the prices for agricultural products, while at the same time ensuring sufficient supply of goods.

When the network evolved out of a given initial state, because of the dominance of the consortium the prices for agricultural goods dropped significantly, causing the farmers' revenues to decrease substantially, despite a similar or even higher workload. As a consequence the farmers' states changed from satisfactory to unsatisfactory, triggering their Classification Jani, which attempted a re-classification of the farmers within the network, based on the new state of the network.

The result of the Classification process was that most farmers were in a similar state of dissatisfaction, so that in the subsequent Segmentation process a new working group was created within which the farmers organized themselves and shared their knowledge. It was then this newly created working group that took over the connection to the marketplace from its individual members. Because of its greater importance, according to the implemented economic rules it could balance the pressure on the prices exercised by the consortium. The prices went up again, giving the individual farmers a higher revenue and thus greater satisfaction.

One could imagine how this scenario might continue, given that the players have enough information and knowledge at hand to adapt their strategies to the evolution of the network. However, the assumptions made were too simplistic to expect this model to reach a final stable state. Studies based on more refined models are currently under investigation.

## CONCLUSIONS

We have shown that aspects of complexity can be modeled with Self-Organizing Fractal Semantic Networks, where generic processes drive the self-organization of the network on a local scale. Necessary requirements of this model are the existence of a topology or neighborhood structure and the self-similarity of the network on all scales of hierarchy. The two processes of Classification and Segmentation are fundamental in driving the self-organization of the network, and it appears that cognition and learning can be derived from them. However, more research is necessary to precisely determine the nature of this relation.

Despite the fact that we have given some guidelines on how to extend the concept of fractals from geometry to topological hierarchical networks (Definition 5), the question of how to do this in strict mathematical

terms is still open. If an answer could be obtained, then it would be interesting to study how the fractal dimension in this network is related to the subjective term of complexity, because this may shed some light on driving mechanisms for learning.

Finally we would like to mention that we have applied the concept of a Self-Organizing Fractal Semantic Network to the problems of Natural Language Understanding and Image Recognition. In these cases texts or images were transformed into initial Input Networks. Structuring and connecting these Input Networks to World Knowledge Networks with the help of the Classification and Segmentation methods described above then accomplished the task of understanding these texts or images.

## ACKNOWLEDGEMENTS

Two of us (Gerd Binnig and Günter Schmidt) would like to thank the Deutsche Bundesstiftung Umwelt for supporting the simulation project mentioned in Section 6.

## REFERENCES

- [1] Davenport, Thomas H. and Prusak, Larry, 1997 "Working Knowledge", Harvard Business School Press, December 1997.
- [2] Quillian, M. R., "Word Concepts: A Theory and Simulation of Some Basic Semantic Capabilities", *Behavioral Science* **12** (1967), 410-430.
- [3] Müller, B., Reinhardt, J., and Strickland, M. T., "Neural Networks: An Introduction", 2<sup>nd</sup> edition, Springer, Berlin, 1995.
- [4] Minsky, Marvin L., ed., "Semantic Information Processing", 6<sup>th</sup> printing, MIT Press, Cambridge, MA, 1988.
- [5] Reimer, U., "Einführung in die Wissensrepräsentation" (Introduction to Knowledge Representation), B. G. Teubner, Stuttgart, 1991.
- [6] Badii, Remo and Politi, Antonio, "Complexity", Cambridge Nonlinear Science Series 6, Cambridge University Press, 1997.
- [7] Gutowitz, H. A., ed., "Cellular automata: theory and experiment", MIT Press, Cambridge, MA, 1991.
- [8] Edgar, G. A., "Measure, topology, and fractal geometry", Springer, New York, 1990.
- [9] Mandelbrot, B. B., "The Fractal Geometry of Nature", Freeman, San Francisco, 1982.
- [10] Cohen, G., "Memory in the Real World", Hove: Lawrence Erlbaum, 1989.
- [11] The Encyclopedia Mythica, an encyclopedia on mythology, folklore, and legend. <http://www.pantheon.org/mythica/articles/j/janus.html>.
- [12] Minsky, Marvin L., "The Society of Mind", Simon & Schuster, March 1988.
- [13] Wenger, Etienne, "Communities of Practice: Learning, Meaning, and Identity", Cambridge University Press, 1998.
- [14] Prusak, Larry and Lesser, Eric, "Communities of Practice, Social Capital, and Organizational Knowledge", *Knowledge Connections*, Vol. 2, No. 1, February 2000.