Algorithms for the self-organisation of distributed, multi-user networks. Possible application to the future World Wide Web

Johan Bollen & Francis Heylighen*

Center "Leo Apostel", Free University of Brussels,
Pleinlaan 2, B-1050 Brussels, Belgium
email: jbollen@vnet3.vub.ac.be, fheyligh@vnet3.vub.ac.be

Abstract

The World Wide Web has a number of striking similarities with other learning networks (natural or artificial). Its structure is that of a distributed network of nodes and links. It also evolves and adapts by being continuously updated and expanded by its contributors and users. This paper describes our attempts to devise a number of algorithms that can make distributed hypertext networks such as the World Wide Web self-organise according to their users' knowledge. A number of experiments were conducted in which experimental networks of English nouns were being browsed via the Internet by several thousands of participants. These experimental networks evolved into a stable state which more or less represented the participants shared knowledge structure and associations.

1 Introduction

The World Wide Web (WWW), as a global network of linked nodes, clearly belongs to a large class of networks adhering to a distributed principle of knowledge representation [Hinton, 1981]. It consists of a huge number of nodes (containing texts, pictures, movies, sounds) connected through hyperlinks to form 'hypernetworks' that can represent complex ideas and concepts above the level of the individual node.

It is this network-like character that has inspired a number of authors to suggest "neural" metaphors for the WWW functioning and structure. [Mayer-Kress, 1994] discussed the possibility of a global brain, capable of information and knowledge processing, spontaneously emerging from the enormous amount of nodes and their continuously evolving connections. Others suggested to study the WWW as an adaptive neural network, where the list of 'Bookmarks' (stored, favorite connections) controls human activity patterns and thereby acts as a possible adaptation mechanism for the network.

In spite of these similarities, the WWW lacks some important functional attributes typical for biological and artificial neural networks. First, neural networks are normally not intended to merely store information, but to control and guide goal-directed behaviour. The WWW, however, does not perform any tasks except information storage. Second, most neural networks are equipped with mechanisms to adapt the knowledge and models they contain. This phenomenon lies at the heart of an error-correcting feedback loop, characterizing biological as well as artificial neural networks [Mc. Clelland & Rumelhart, 1986]: 'knowledge behaviour effect perception knowledge adjustment'. The WWW does not have any such error-correcting mechanisms. It evolves, but does not adapt.

One might argue that it is not the WWW's goal to simulate brains or neural networks, but to provide a reliable and user-friendly access to stored knowledge. But it is questionable whether the present WWW—and the hypermedia paradigm in general [Nielsen, 1990]—succeeds in this [Jonassen, 1989; 1993]. The WWW's content is presently expanding at an enormous pace, but the quality of its structure does not seem to improve. This should not surprise us, as the only mechanism for network restructuring at present is the contributions of individual web-designers, each adding their own, often poorly designed, sub-networks to the WWW. The WWW, being not more than the sum of its parts, can achieve no better quality of structure than that of these sub-networks. This causes the WWW to be, in general, very poorly organized, which in its turn seriously hampers efficient and user-friendly retrieval of information [Hamond, 1993]. With an ever-expanding amount of information being added to the WWW, this problem can only be expected to worsen within the present set-up.

We believe the only solution to these practical and fundamental problems is to move beyond metaphors and implement the necessary conditions to make the WWW really function in a more "brain-like" manner [Heylighen & Bollen, 1996]. As a first step in that direction, we tried to design a mechanism for the self-organization of a hypertext network. Our goal was to develop algorithms that would allow the WWW to autonomously change its structure and organise the knowledge it contains, by "learning" the ideas and knowledge of its human users as manifested in their browsing behaviour, thus producing a

---

more ergonomic and user-friendly network.

2   The Experiment

To study the scientific and practical issues involved, we set up an experimental network of 150 nodes linked by weighted connections. The connection strengths were set to small, random values to initialize the web. The ten strongest connections were made visible to the user. Network structure was then adjusted through a number of learning rules operating on the connection weights. These learning rules were based on the frequency with which users traversed (configurations of) connections. The network was made publicly accessible to the global community of WWW users who were invited to browse the network.

When entering the network, users were presented with a node representing a word (e.g. "Cat"). Below the title word, ten other words (e.g. "car", "dog", "mouse", "house", etc.) were listed, each linking to the respective nodes representing that word. The users were asked to select from that list the word that has the strongest association with the title word. Once they had selected a good association for the word "Cat", say "dog", they would be presented with the "Dog" node, again with a list of possible associations (e.g. "garden", "barking", "water", etc.) to select from. Thus, they could wander through the network, each time selecting new associated nodes, until they got tired and quit the experiment.

2.1   Nodes

A list of the 150 most frequent English nouns was derived from the LOB corpus [Johansson, 1989], to be used as nodes. The number 150 was chosen because it is large enough to allow a rich set of associations (the matrix of all possible connections has 150 x 150 = 22500 elements), yet is still manageable for analysis. The decision to use the most frequent nouns was inspired by the following considerations. First, word frequency is a selection criterion unbiased by possible preferences of the experimenter. Second, frequently used words are most likely to be well-understood and to have a rich set of associations with other words. A weakness of this set-up is that potentially strong associations (e.g. "cat" -> "mouse") could never be made if one of the words was not sufficiently frequent to be included in the list.

2.2   Implementation

A HyperCard stack was used to implement the self-organising network. The stack kept track of connections strengths, implemented the learning algorithms and handled the requests from our MacHTTP server application, which communicated with the rest of the World-Wide Web. Each node in the network was represented by a "card" (record) in the stack. A number of fields on each of the cards kept track of each specific node's status and its connections to other nodes.

Each time a subject contacted our server and requested to take part in the experiment, he/she was assigned a random starting position in the network of nouns. The stack then returned an HTML page with the name of the present node on top, followed by an ordered list of the 10 highest ranking associated words, each hyperlinked to the node representing that word. If the subjects did not find a sufficiently good association in the list, they could select a "More words..." link at the bottom of the page, which would call up a new list with the 10 next ranking nouns. By repeatedly selecting this link, they could in principle go through all 149 possibilities in the order of decreasing connection weight (although this never seemed to happen).

Subjects could travel the entire network at will by repeatedly selecting the "best" association from the list of proposed connections. If a user selected a link, the stack would immediately adjust the connection strengths according to the learning algorithms and sort the connections according to their new strengths, so that the next time a subject would visit the same node, the order of the links might have changed. Thus, links which previously were not strong enough to appear among the 10 strongest ones, might make the transition to this list of links which is directly visible to the user, and thus become "actual". Similarly, links that would lose strength relative to the other ones, might disappear from the "actual" list, and thus become more difficult to select.

2.3   Subjects

Our potential pool of experimental subjects was the entire Internet community. In a letter sent to a large number of newsgroups and mailing lists, people were invited to contact our server and browse the network for a while. Two experiments were conducted that each involved over 600 participants, resulting in a total of some 1200 participants. A limitation of this use of the Internet on a voluntary basis was that we had no control over the variables determining the population, such as age, gender or cultural background.

2.4   Learning algorithms.

We used three rules which were all inspired by the Hebbian principle of learning: the link between nodes of the network that have been activated within the same interval of time is reinforced [Hull, 1952; Thorndike, 1911]. This mechanism is entirely associative in nature and claims to achieve global optimisation of network structure through the adjustment of local connections and is as such in accord with the associative nature of the WWW and the absence of centralised control. In line with the principles of evolutionary epistemology [Campbell, 1974], our learning algorithms were also based on the principles of variation and selection, which are assumed to guide the evolution of knowledge [Heylighen, 1993]. Of the three learning rules, two produce variation by introducing new candidate links to the list of 10 "actual" links, the third one produces selection by rewarding or punishing already actual connections.

Frequency
The Frequency learning rule is based on the assumption that subsequent nodes in a user's path have a high level of 'relatedness' and their connection should therefore be strenghtened. The rule adds a small "bonus" *Fb* (set to the value 1) to the strength S of the connection between two nodes A and B which are subsequently visited by the

user: $(A \rightarrow B)$ $S(A, B)+Fb, Fb=1$.

Frequency applies selection to the network by selectively encouraging the use of already established connections, thereby indirectly inhibiting the use of unpopular links. (rarely used links would diminish in rank, move down on the ordered list of links, and therefore become more difficult to select by the user).

Transitivity

If we consider the associative relations between nodes in a continuous path through a hypertext network then we could say that they are essentially all related to each other via intermediate associations, but the relatedness decreases with increasing distance between two nodes in a path. Thus, if a node B in a path is preceded by A and followed by C, we might assume that A and C are indirectly related. The transitivity rule implements this observation by strenghtening the connection from node A to node C with a bonus Tb, which is smaller than the frequency bonus for a direct connection Fb: $(A \rightarrow B)$ & $(B \rightarrow C) \Rightarrow S(A, C) + Tb, Tb=0.5$

Transitivity opens up an unlimited realm of new links. Indeed, one or several increases in strength of A   C may be sufficient to make a previously "potential" link actual (move it to the first 10). The user can now directly select A   C, and from there perhaps C   D. This increases the strength of the potential link A   D, which may in turn become actual, providing a starting point for an eventual further link A   E, and so on. Eventually, an indefinitely extended path may thus be replaced by a single link A   Z. Under the assumption that browsing behaviour is goal-directed, the transitivity rule bridges intermediate nodes and reduces the number of links that have to be followed to reach the destination.

Sometimes the new connections by transitivity are spurious (e.g. Cat Mouse Cheese   Cat Cheese?), but this will be corrected by the Frequency learning rule rewarding only the worthwhile links.

Symmetry

This rule also introduces variation to the network's development. It is based on the assumption that if A is related to B, then B is also related to A. For every traversed connection between a node A and a node B, symmetry reinforces the connection between the node B and A by a small addition *Sb* to its connection strength: $(A \rightarrow B)$ $S(B, A) + Sb, Sb=0.3$.

Symmetry too introduces new connections to the network by reinforcing connections that have not explicitly been chosen, but it is more limited in that respect than transitivity since only a single symmetric link can be generated for every existing link. Frequency will afterwards determine whether these links will further develop or not. The symmetry rule was only used in the second experiment.

It must be noted that the symmetry and transitivity rules have a synergetic effect that cannot be produced by a single rule. For example, consider two links A_1 B, A_2 B. The fact that A_1 and A_2 point to the same node seems to indicate that A_1 and A_2 have something in common, i.e. are related in some way. However, none of the rules will directly generate a link between A_1 and A_2.

Yet, the repeated selection of the link A_2 B may actualize the link B A_2 by symmetry. The repeated selection of the already existing link A_1 B followed by this new link B A_2 can then actualize the link A_1 A_2 through transitivity. Similar scenarios can be conceived for different orientations or different combinations of the links.

Given the choice of Fb as unity, the values of Tb and Sb are largely arbitrary and are not considered critical for the functioning of our algorithms. We have chose them such that they seem to provide a good balance between the amount of variation and selection.

3   Results

Response to our request for participation was surprisingly massive. Both experiments attracted over 600 participants. Although there were no upper or lower limits imposed on the length of the path travelled through the network, on average subjects ended their participation in the experiment after about 10 jumps. Few people used the possibility to see an additional list of 10 links for a given word: on average only 1 in 4 of the subjects would request more words for any given list. This means that, for example, the 5th list of ten words would be consulted only by about 1 in 1000 (4^5) users, a negligible number.

3.1   Network Development

Network development was surprisingly fast and efficient. After only 2500 link selections (out of 22500 potential links) both experimental networks had achieved a fairly well-organised structure in which most nodes had been connected to large clusters of related words. This was in particular true for the second experiment, where the addition of the symmetry rule practically doubled the introduction of new links in the beginning stage. After the initial phase in which a rapid consolidation of connections and clusters of nodes occurred, development seemed to slow down considerably afterwards, until nothing much was changing anymore. This is due at least in part to the way in which the connections strengths are calculated, since the addition of a constant bonus will make less difference when all connections have already gathered high total strengths than in the initial phase where any bonus may be sufficient to make a link move to the highest ranks.

MIND				
0	600	1200	2400	4200
table	thought	thought	thought	thought
order	idea	idea	idea	idea
figure	research	research	knowledge	knowledge
party	problem	change	development	view
question	need	need	change	education
school	light	knowledge	theory	theory
act	development	problem	research	development
history	change	development	need	research
fact	view	example	education	change
wife	law	life	view	problem

Table 1: self-organization of the list of 10 strongest links

from the word "Mind", in different stages: initial random linking pattern, after 600, 1200, 2400 and 4200 steps. (a step corresponds to a user selecting one link on one of the 150 nodes in the network)

A typical example of how connections are gradually introduced and rewarded until their strength reaches an equilibrium value, may illustrate how fast and efficient the self-organising process was: Table 1 provides an overview of the connections that were formed as the node 'Mind' became connected to related nodes. The position of these associated words shifted upwards in the list until they reached a position that best seemed to reflect their relative strength.

At the end of each experiment, after some 6000 selections, the most frequented nodes had gathered a list of 10 strongest links that quite well reflected their direct semantic environment, with words that were near synonyms of the node name at the top of the list (see Table 1). However, this positive result was much less strong in the less frequented nodes, because of what we termed the "attractor effect". Nodes that had many incoming links, by accident, or because they were associated with many other words in the list, would tend to attract more users. This would result in increasing strength of their incoming paths, and their replacement by even stronger direct links. Especially in the first experiment, almost all paths would end up in a cluster of semantically related, strongly cross-linked nodes, forming an approximate attractor for the network. Although the random assignment of starting nodes meant that all nodes would be consulted on first entry with the same average frequency, the subsequent moves would very quickly end up in the attractor cluster. As a result, nodes outside the attractor would get little chance to learn and thus remain poorly connected.

In our second experiment, the introduction of the symmetry rule attenuated this effect, since strong links leading into an attractor would necessarily produce weaker, inverse links leading out of the attractor. This gave nodes outside the attractor the chance to develop some links of their own, generating new local attracting clusters, weakly connected to other clusters. The overall learning seemed more efficient in the sense that less time was needed to develop good associations, and the result was more balanced, in the sense that the differences in frequentation between nodes were less strong.

3.2 Cluster Analysis

Network structure not only depends on the dynamics of the self-organization but on the content of the domain. Although we selected the words to be a priori independent, by not restricting them to any particular domain, the network's evolution discovered a number of strongly related semantic "families". A k-clique cluster analysis of the matrix of connections revealed a number of stable and separable clusters corresponding to highly general categories. The following 9 clusters of associated words, each denoted by an intuitive label for the underlying conceptual category, were found in the second experiment's final structure:

"Time":	age, time, century, day, evening, moment, period, week, year
"Space":	place, area, point, stage
"Movement":	action, change, movement, road, car
"Control":	authority, control, power, influence
"Cognition":	knowledge, fact, idea, thought, interest, book, course, development, doubt, education, example, experience, language, mind, name, word, problem, question, reason, research, result, school, side, situation, story, theory, training, use, voice
"Intimacy":	love, family, house, peace, father, friend, girl, hand, body, face, head, figure, heart, church, kind, mother, woman, music, bed, wife
"Vitality":	boy, man, life, health
"Society":	society, state, town, commonwealth
"Office":	building, office, work, room

Although the learning algorithms only work on links and not on groups of nodes, it is remarkable how well the resulting clusters fit in with intuitive categories. With rare exceptions (e.g. "side" in the "Cognition" cluster), all of these words seem to be located in the right class. This again seems to confirm that the set-up achieves its aim of absorbing the common semantics of a heterogeneous group of users. The 'cognition' cluster makes up 33% of all words over all clusters, indicating its central position and importance in the network. It should be noted that this prominence may be due to the specific selection of texts in the LOB corpus, which may have been biased towards more "intellectual" activities.

3.3 The Role of Positive Feedback

The above described experimental setup, and in fact the entire concept of a self-organising network, involves a positive feedback loop. As could be expected, subjects are more likely to select the items they read first in the list proposed to them. They are even more likely to select items from the first available list, without calling up additional 10 word lists. Therefore, connections which rise in the rank ordering because they are selected, would get a significantly higher probability of being selected on a following occasion. Thus, reinforcement of a link tends to produce further reinforcement. Such a positive feedback might make the network's development highly sensitive to its initial conditions. A 'Mattheus' effect might continuously reinforce initial random connection values and thus severely distort the network's development.

Apart from the less deleterious "attractor" effect mentioned earlier, such a distortion did not seem to occur. This is probably due to the continuous generation of novel links by transitivity and symmetry, which can quickly displace unjustly reinforced existing links. The relatively weak sensitivity for initial conditions can be seen by comparing the results of the two experiments. The experiments were conducted independently within an interval of 3 months, with a different random initialization of connection weights, a different group of subjects, and the addition of the symmetry rule in the second exper-

iment. Yet both delivered quite similar final network structures. Cluster analysis showed an overlap of 76% over all clusters in the two experiments. This result, however, cannot be considered conclusive evidence of developmental stability since quantitative comparison between associative networks remains a difficult task, and our two experiments can certainly not be considered adequate samples.

The beneficial effect of this feedback loop was that it considerably speeded up network development. Feedback can strengthen new connections introduced by Symmetry and Transitivity in a fast loop of continuing rewards, administered by the Frequency learning rule. Any worthwhile link could as such rapidly achieve a high position, and pull transitively or symmetrically related links up in its wake.

3.4 Variety and Selection.

As said, our learning rules were constructed to implement variation and selection. The Symmetry and Transitivy rules add variety by introducing new, not necessarily useful connections, that could later be selected upon by the Frequency rule. Our data seemed to confirm the functionality of this set-up. A temporal analysis of the final states of the network generated in our second experiment revealed that in at least 6 of the 20 best scoring connections Transitivity and Symmetry bonuses preceded a loop of Frequency rewards.
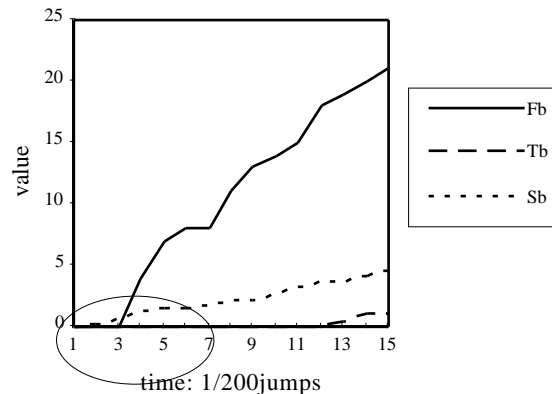
Knowledge-Research
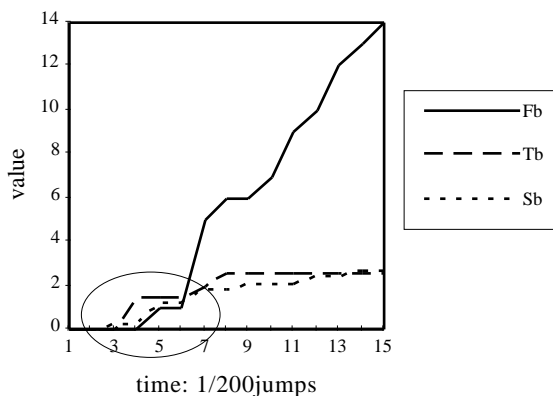
Graph 2 Cumulative weight over all network states for the connection Knowledge->Research split up in Symmetry (Sb), Transitivity (Tb) and Frequency (Fb) bonusses.

This indicates that at least 6 out of 20 high scoring connections were introduced to the network's structure by transitivity and symmetry rather than the human browser's active selection. Due to the rather large interval of 200 jumps between subsequent measurements of network state, apparently 'simultaneous' onsets of rewards from all three learning rules could not be resolved into separately measured events, so we expect this number to be even higher in reality. An example of this pattern is shown in Graph 1 and Graph 2. These data also support the acceleration through feed-back as hypothesized in the previous section.

Mind-Thought

Graph 1: Cumulative weight over all network states for the connection Mint-Thought, split up in Symmetry (Sb), Transitivity (Tb) and Frequency (Fb) bonusses.

4 Application to the World-Wide Web

Although many further refinements and tests of the experimental set-up need to be made, these experiments show that a set of simple learning rules may be sufficient to let an associative network self-organize until it more or less reflects the intuitive semantics of its users. It would be very useful to extend such a mechanism to the World-Wide Web at large. This, however, requires several modifications.

First, the present experiment is run by a single application on a single server. This facilitates the exchange of information between the different nodes. An implementation on the web would require a change to the HTTP protocol that controls the communication between servers and browser applications. The protocol already includes a "Referer" line, which tells the server from which previous node B the present node C was linked to. To implement the symmetry rule the server may use this information in order to reinforce a potential link from C to B. The transitivity rule requires a more substantial change, since it requires the introduction of the additional address of a "Previous Referer", i.e. the node A which referred to B, which in turn referred to C. The server getting a request for C should then be able to signal this to the server that originally provided A, so that it could strengthen a potential link from A to C.

The creation of new links is not obvious in the present HTML protocol for hypertext, since links are embedded in the text of a document, and thus are difficult to change. This obstacle is overcome in the Hyper-G system, which proposes an enhanced version of the World-Wide Web, though it appears similar to casual users. In Hyper-G, links are separate from the document and under the control of the server. This makes it easy to add a list of "learned" links, e.g. at the bottom of a document, which would provide the user with further possibilities for navigation, in addition to the (unchanging) links provided by the author of the document. Initially, this list would be empty, but as soon as selections are made, tran-

sitivity and symmetry will start adding potentially related documents. The server should only show the highest ranking links to the user, and keep the others in memory until they gather sufficient strength (or until a user requests to see "more links").

Another difference with the experimental set-up is that Web documents typically contain texts or graphics and not just a list of linked nodes. Thus, users, when judging whether a document is interesting, will need more than the name or title provided by the link. A link that seems promising may well lead to a document with an irrelevant or useless content. Therefore, the learning mechanism should need some kind of "quality control". The most obvious, but least user-friendly, way to implement this, would be to let the user evaluate the usefulness of the document before transmitting the bonus. A less obtrusive, but perhaps less reliable, method would be to only transmit a bonus after the user has stayed for a sufficiently long time with the same document, or selected a further link in this document. Link selections immediately followed by backtracking to a previous document should not be strenghtened. Perhaps the user himself or herself may decide to "reward" particularly good nodes, and "punish" particularly bad ones, whereas a background mechanism would make an estimate of usefulness in all other cases.

Such enhancements require a more sophisticated exchange of messages between browser and server than provided by the present HTTP, but may lead to a much higher quality Web, without making additional demands on users and authors. The criticism that selecting links on the basis of frequent use will merely strengthen "conventional wisdom" or the "lowest common denominator", stifling creative, original or unpopular associations, is misdirected. First, the links made initially by the author of the document are not touched by the mechanism, and thus there is no restriction on introducing non-conformist associations. Second, the popularity of a link only matters relative to the other links "learned" by the same document. Specialized or controversial documents will normally only be consulted by a limited public interested in the specific subject of these documents. These users are likely to continue their search for more information about this little known domain, thus creating more specialized links, and not to suddenly jump to mainstream, popular subjects. As a group, they are likely to know more about the subject than the individual author, and are therefore likely to select relevant links unknown to that author.

5   Conclusion

This paper has attempted to demonstrate that simple, locally operating learning rules can make an initially random hypertext network self-organise into a meaningfully structured network, by using the information implicit in the users' browsing behaviour. Although the experiments we performed need to be repeated and refined, the preliminary results seem quite convincing.

The practical advantages of these such algorithms in terms of efficient hyper-structure are obvious. Hypermedia networks for storage and retrieval of information are becoming increasingly important tools, but the paradigm's complete dependence on human network design is a serious limitation to the further development of this medium. This structural problem can only worsen with an increasing amount of information being added.

We believe that these algorithms can be further extended to solve many fundamental and practical problems of the present World-Wide Web. With a number of minor technical adjustments to the protocols and servers, these learning rules could transform the WWW as we know it into a truely adapting and active associative network. It would be able to absorb the implicit knowledge of its users and discover new relations between pieces of information. Such a world-wide associative network can be seen as the first step towards a "global brain". [Heylighen *et al.*, 1996] and [Mayer-Kress *et al.* 1994] discuss this 'super-brain's enormous potential for humanity, but even the simple addition of associative might be sufficient to spectacularly increase the power of the Web.

References

Campbell D.T., Evolutionary epistemology, in: *The philosophy of Karl Popper*, ed. Schillp P.A., (Open Court Publish., La Salle, III.), p. 413-463, 1974.

Hinton, Geoffrey en James Anderson ed. Parallel Models of Associative Memory. Hillsdale, N.J., 1981.

Hamond, N.: Learning with Hypertext: Problems, Principles and Prospects, In: *Hypertext: a Psychological Perspective*, p. 51, ed. C. McKnight, A. Dillon and J. Richardson, Ellis Horwood, 1993

Heylighen, F.: Selection criteria for the Evolution of Knowledge, *Proc. 13th Int. Congress on Cybernetics*, Namur, 1995.

Heylighen F. & Bollen J. : The World Wide Web as a Super-Brain: from metaphor to model, Cybernetics and Systems '96, World Science- Singapore (this volume), 1996.

Johansson S. & Hofland, K.: *Frequency analysis of English vocabulary and grammar: based on the LOB corpus*, Oxford : Clarendon Press, 1989. - 2 v.

Jonassen D. H.: Effects of Semantically Structured Hypertext Knowledge Bases on Users' Knowledge Structures. in: *Hypertext: A Psychological Perspective*, p. 153, ed. C. McKnight, A. Dillon and J. Richardson, Ellis Horwood, 1993

Jonassen, D. H. : Semantic Network Elicitation: Tools for Structuring Hypertext in "*HYPERTEXT II: State of the Art*", p. 142-152, 1989

Klimesch, W.: *The Structure of Long Term Memory*, Lawrence Erlbaum, 1994.

Mayer-Kress, G. & Barczys, C.: The Global Brain as an Emergent Structure from the Worldwide Computing Network, and its Implications for Modelling, *The Information Society*, Vol 11 No 1, 1994.

Mc. Clelland, J. & Rumelhart, D. E.: *Parallel Distributed Processing.*Volume 1: Foundations, Bradford Books, MIT Press, 1986.

Nielsen, J.: *Hypertext andhypermedia*. Academic Press, 1990.