Tricot, A., & Coste, J.-P. (1995). Evaluating complex learner-computer interaction : What criteria for what task? In *EARLI'95 Conference*, symposium « Analysing learner-computer interaction : lessons from empirical studies ». Nijmegen, Netherlands, August 26-31

Evaluating complex learner-hypermedia interaction: What criteria for what tasks?

André Tricot (1) and Jean-Paul Coste (2)

(1) CREPCO-CNRS, University of Provence, 29 avenue Robert Schuman, 13 621 Aix en Provence Cedex, France (e-mail: raison@romarin.univ-aix.fr).
(2) Equipe Hermès, University of Provence, 3 place Victor Hugo, 13 331 Marseille Cedex, France.

Abstract

In this paper we address the issue of criteria definition when analysing learner-hypermedia interaction (LHI). So far studies on hypermedia usability have used simple information search tasks where the subject's goals correspond to a small subset of nodes and for which simple dependent measures can be used. However, hypermedia-based learning tasks often involve more complex interactions, for which the usual criteria are no longer relevant. We suggest that, like any human behavior, description and evaluation of LCI should refer to some psychologically relevant model, *i.e.* a model of the task performed by the learner. To illustrate that point, we describe an experiment in which university students were asked to use a very large hypermedia database in order to perform a complex learning task. We propose several methods to characterize learner-hypermedia interactions in situation, and we present some outcomes of our analyses. We conclude that, in order to understand the potential of hypermedia for learning, comprehensive activity models related to different learning tasks or objectives are needed. However, such models are not yet available at the present time.

Introduction

In this paper, we address the issue of criteria definition when analysing learner-hypermedia interaction (LHI).

Hypermedia system may be used for a wide range of tasks, which may be defined as a goal to achieve through a series of actions within a certain environment (in this case a computer system).

We have categorised learning tasks as, more generaly, “searching and integrating knowledge” tasks, in function of two crossed criteria (Tricot, 1994, here table 1): (a) goal implementation in the document: it can be single (goal is located on one node) or multiple (goal is distributed on several nodes); in this second case, the different relevant nodes should be directly linked or not; (b) user’s goal representation, (which may be more or less precise).

The combination of these two dimensions results in four typical information usage tasks.

		Goal representation	
		precise	*fuzzy*
Goal	*single*	locating	exploring
implementation	*multiple*	searching	aggregating

Table 1. *Four typical information usage tasks*

- "Locating" corresponds to cases where the user has to deal with an explicit request about a unique piece of information.
- "Exploring": The user doesn't have an explicit request but he or she looks for a relevant (and unique) piece of information.
- "Searching": The user has an explicit query which corresponds to a set of units in the hypertext. The set of relevant units may be grouped or distributed.
- "Aggregating": The user doesn't have a precise query (or cannot represent it) but he or she thinks that he or she may find several relevant units in the hypertext.

For this study, we proposed a particular “aggregating” task, where the subjects are updating their knowledge through a consultation (they previously learned the knowledge presented in the hypertext). This type of task may increase the "functionality" of knowledge by showing the user how previous knowledge can be inserted in a new context (see Spiro Feltovitch, Jacobson & Coulson’s (1991) concept of criss-crossing). We would try to define criteria and an experimental protocol adapated to this task, which are absolutely different from those of “locating” task. So far, most empirical studies on hypermedia usage (*e.g.* Foss, 1989; Gray, 1990; Gray, Barber & Shasha, 1991; Leventhal, Mynatt Teasley, Instone, Schertler Rohlman & Farhat, 1993; Mohageg, 1992; Wright & Lickorish, 1990, 1994) have used information search tasks in which:

- The subject has to answer a small number of questions, with little or no relation between them.

- The number of relevant nodes is rather small: 1 or 2 relevant nodes per question, sometimes 5 or 6 for questions labelled "jugement" or "synthesis".

- The systems are themselves very simple: A few dozen of nodes at most.

In these studies, criteria used to evaluate subject's performance often are simple quantitative measures:

- <u>Recall</u>: Did the subject open the relevant node(s)?

- <u>Precision</u>: Did the subject not open the irrelevant node(s)?

- <u>Economy</u>: Did the subject use the shortest path to reach a target node? Did the subject not return several times to a given node (looping)?

(It should be noticed that those criteria are directly drawn from valuation of automatic information retrieval systems, rather than on models of human performance).

However, real-life learning tasks involve more complex interactions between the learner and the system. The learner's goal may be more ambitious than a simple information retrieval (e.g., learn a complex set of related concepts in a given content area), and the goal-node correspondance may be less thoroughly defined. In these siutations, the simple quantitative criteria above listed are no longer relevant. Consequently, an attempt must be made to define relevant methods to characterize the learner's activity.

In order to explore this issue, we designed an experiment in which university students were asked to use a large hypermedia CAL system in order to perform a complex learning task. Following is a brief description of this study.

<u>Learning in Correl...</u>

Within a piagetian metaphor, we should say that learning in our experiment with Correl... concerns more "accommodation" than "assimilation". In other terms, we would like to involve contextualisation of previously acquired knowledge, mediation in knowledge representation processes and synthesis working out in a specific physics topic (theories and phenomenon in wave propagation). We (Coste, 1991, 1993) and others (see for example Redish, Wilson & McDaniel, 1992) outlined the great interest of hypermedia for this aspect of teaching physics.

Knowledge in Correl... is mainly presented as static and dynamic pictures (videos), designed by simulation. The goal of thoses pictures is to favour the way throught reality representation and concept representation. The "hyper-strucutre" of the software made us able to present different representation modes and knowledge access.

Description or organisation of knowledge transfer in terms of task or didactive objectives is most of the time based on the analytic presupposition that a complex objective could be described as a set of subgoals, more and more specified. The result is a range of elementay (or supposed to be) sub-boals. This "task rationnal model", designed by the task designer, in terms of goals and means, is very probably unlike the subject's task representation, in terms of goals and/or means. In other words, the use of "space problem" (Newell & Simon, 1972) or "task rational model" (Anderson, 1990) notions implicate two risks: The risk to correspond to the designer representation and not to all possible relevent descriptions; the risk to not correspond to the student problem representation (see, for exemple Bastien, Pélissier & Tête, 1990; Pélissier & Tête, 1994).

Moreover, in some teaching or evaluating situations, the difficulty is, properly, to draw up this task representation. The subject doesn't know what he is searching. Then what hypothesis are made? What control and regulation mechanisms are operating? What analogies induce the subject's constructions? However, the fundamental role of task representation in subject activy management is well known (Chatillon, 1988; Chatillon & Baldy, 1994; Devichi, 1994), in reference with goal representation and means to be used (rules application, etc.). Activity management and task representation will evolve according to a continuous process, both co-regulated by means of information taking on situation, task representation linking current goal and final goal. Working out goal representation becomes a major focus when we point out:

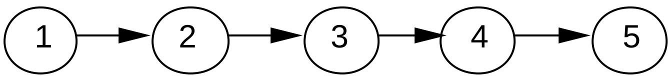
- the activity of working out a problematic from a situation rather than problem solving;

- rules and processing method definition rather than question answering and rules application;

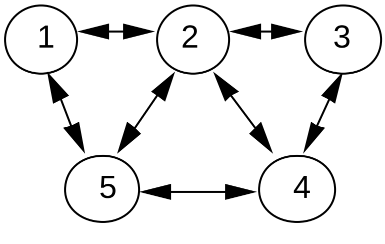
- the operation of modelling a situation rather than applicating theory to a situation.

The kind of learning involved in our experiment follows that way:

- we don't define a (chronological) range of (elementary) sub-tasks for recomposing a (complex) global task;



- but rather propose a global task to achieve, according to different specifications (with no order constraint), trying to help subject to draw up global task representation by linking local tasks, each of them adding a piece of meaning in drawing up global task representation.



Following that way, we hope that subject, at the end, doesn't say " I've known how to answer questions" but "I've understood THE question". That he or she doesn't say "I've known how to solve the problem" but "I've understood what means the problem".

"Synthesis" questions were asked, after a self-regulated activy of the subject on the tutor. The answer of each of those local questions was very much determinated by the comprehension of the global problematic, this global comprehension being itself determinated by linking the different local levels of questions. That is what we try to illustrate about the general question of modelling a physical system and drawing up a representation space of its states.

This study is an observation and not an experiment. We will nor test hypothesis, control factors, neither describe cognitive processes. We have just wanted to illustrate for what type of learning Correl... has been drawn, and to present what kind of task the subject has to achieve.

Method

The content area studied in this experiment was physics and the specific topic was wave propagation. A major instructional problem is to help university students to build up different forms of cognitive representations for a given phenomenon (e.g., shift from a physical to a vectorial or to a mathematical representation). An hypermedia-CAL database was designed in order to provide student multiple representations (including dynamic simulations) of wave propagation phenomena.

Material

The three kind of representations (mathematical, physical, vectorial) of any wave propagation phenomena were linked in the Correl... software. This database was implemented as a set of 61 HyperCard stacks (8 Mo) and included more than 1300 nodes:

- 3 orientation stacks (index, tables of contents (fig.1), menus, tables of relations (fig.2) ...),
- 3 explanation stacks (theoremes, pictures),
- 1 exercices stack (fig.3),
- 1 notebook stack,
- 53 animation or "video" stacks (fig.4).

An animation stack presents a continuing run of the different system states; an "help" button give a comment / explanation; an animation stack could be linked to other animation stacks: for example, the vibration vertical mode of one kind of system is linked to the vibration horizontal mode of the same kind of system.

A 12 button menu bar provides different data access ways: accessing to orientation stacks or moving locally: going on, going back, first card, local links... Some orientation or explantation cards get many other buttons (semantics ones).

At the begining of the session, the subjects were given a booklet containing background information (mainly equations), a user manual for the hypermedia system and a set of questions to be answered.

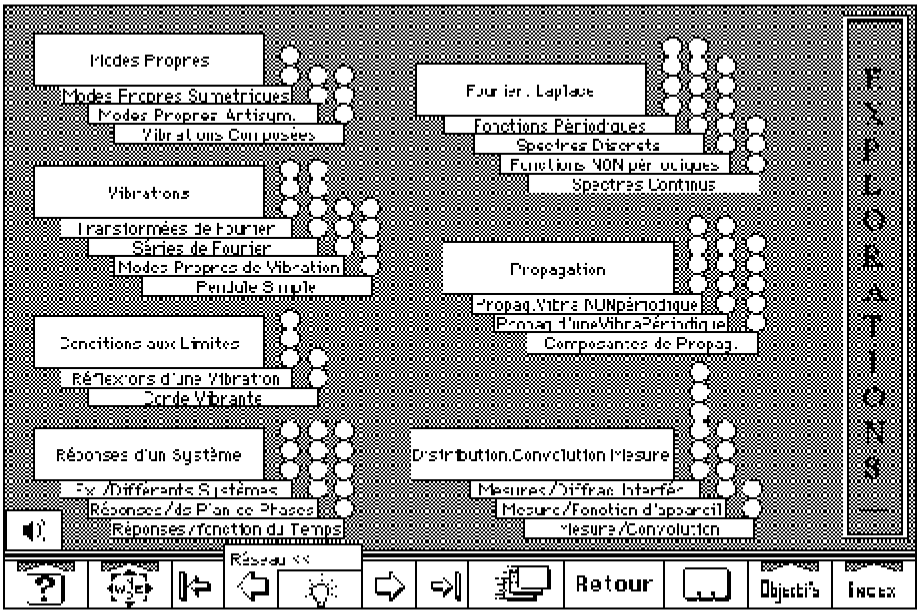


Figure 1. *One example of table of content in Correl...*

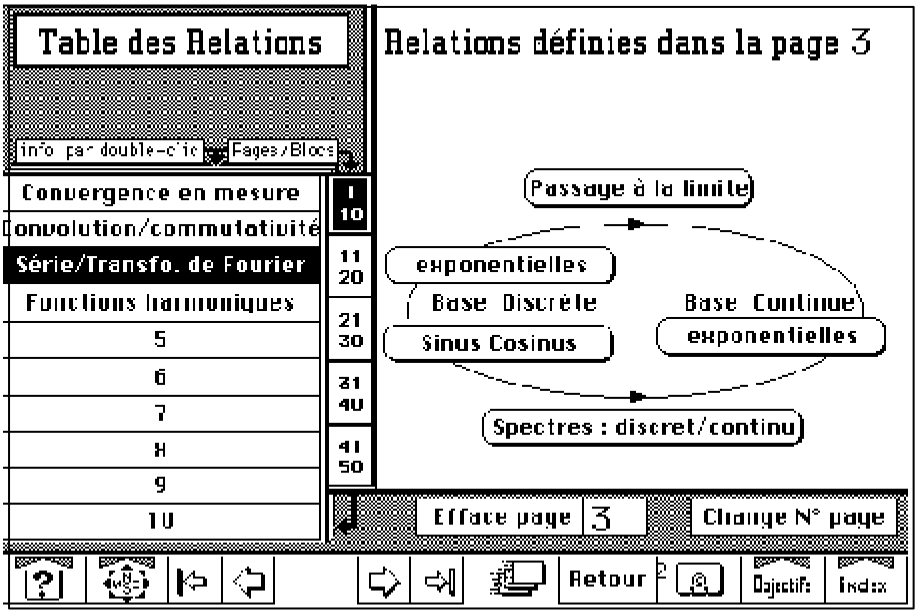


Figure 2. *One example of table of relations in Correl...*

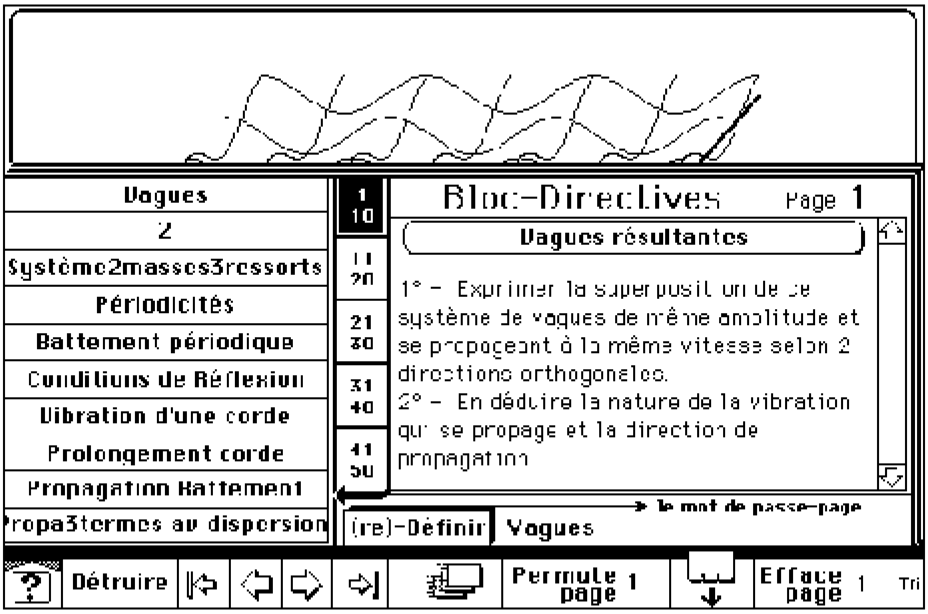


Figure 4. *One example of exercice inCorrel...*

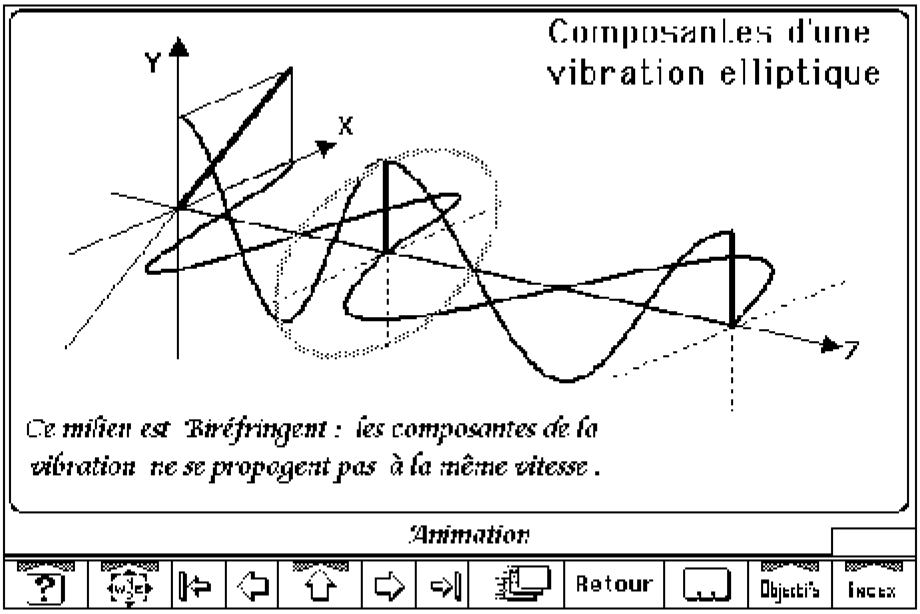


Figure 5. *One example of card in an animation stack*

<u>Subjects</u>

The subjects were senior students in engineering. The experiment took place as a 6-hours lab session during which the subjects were grouped by pairs and asked to study wave propagation problems.

Task

The task is an "aggregating" one (see table 1). The global task that subjects have to achieve is to model a few physic's situations, on wave propagation. More precisely, the core of this modeling task is to build up a representation space of the different states of a given physical system, fitted with a vectorial space structure. The building up of this space and its components depend on the model built: a basis of this space represents a direction of the physical space, or a vibration own mode, or a particularly periodic vibration polarised in a given direction, or a basis function of the decomposition of Fourier's series...

However, whatever the system and the formalism are, subjects have to build up the vectorial space representation. In the task we designed, this level of comprehension is necessary. It is needed to understand links between space representation features and links between their corresponding in physical space. The knowledge of the states vectorial space by one of its bases (base vector or base function) allows subjects to know all the possible solutions, then to know the particularly solution corresponding to the muzzle conditions imposed to the system.

Concretely, the task is to complete a four page report, with formula, pictures or comments.

Questions concern the following systems and phenomena:

. transversal vibration of one system (2 mass - 3 springs)	-page 1-
. vibration and propagation of a pinched rope	-page 1-
. propagations of different kind: terms' combinations propagating in same speed or not	-page 2-
. limit conditions and periodical prolongations of a closed system	-page 3-
. synthesis between different formalisms: discrete *vs* continuous environment, vibrations *vs* propagations	-page 4-

The reading of this four pages report should not allow the subjects to understand immediatly links (between locals questions) that constitute the representation problematic (*i.e.* the global task). However, we designed those four pages to support the representation drawing up process. Like this, page 1, the writing of each own mode has only one analogue, in the writing of rope vibration, *i.e.* the basis functions [$\sin(K_n z+a_n)$] that generate solutions' space;

but, synthesis asked page 4 induce subjects to bring them together: Homology between decompositions makes clear the common structure between the two representation spaces. Expression of progressive and regressive waves, asked page 4 must help comprehension of page 1 question on propagation along the rope. Thinking about the links between the different phenomena must help subjects to represent the modelling process (building up the vectorial space representation), and this approach must ease the comprehension of expressions asked page 1 concerning some recompositions: linear combination of symmetrical (or anti-symmetrical) vibration modes, horizontal or vertical modes. The page 3 question about periodical prolongation of a system defined by its limits conditions can be understood only by the physical meaning of the decomposition basis terms: each of them should be interpreted as system vibration own mode (page 1, page 4).

Task description we give makes clear that subject cognitive activity can not be described as a classical problem solving activity. Subjects have to represent theirselves the comprehension objective through asked questions, to inscribe these questions in a common problematic, to coordonnate the set of models and phenomena in a representation structure: All that constitutes an “epistemological” complex objective that is difficult to describe objectively. We can just imagine cognitive strategies that should more or less follow our previous description, with no obliged ways nor chronological order.

For the purpose of the experiment a subset a 25 questions was used (Q = 25); for each of these questions, some nodes (between 1 and 12) in the system were considered directly relevant. Some nodes are relevant for many questions. For the 25 questions set, 32 nodes spread across the system were considered directly relevant (n = 32, that is 1 "orientation" card, 10 “explanation” cards, 1 complete “explanation” stack (12 cards), 20 “animation” (video) stacks). Answers could be either explicit in the system or inferable from the system information. Questions and relations between questions were so complex that the subjects could not apply a sequential global representation of the dynamic states of a physical space and local representations associated with each question.

Each answer is quoted 1 (false), 2 (partial) or 3 (right).

Results

General results

On average, subjects used the system during 5 hours 03 min. They opened 521 nodes (1 node = 1 “explanation” or “orientation” card or 1 “animation” stack or the 12 card relevant explanation stack). Orientation nodes (*i.e.* menus, indexes, tables of contents) represented 35,8% of the opened nodes (s' = 13,6), 34,5% (s' = 17) for the “explanation” nodes, 25,7% (s' = 9,9) for the “animation” nodes and 4% for the access to the system. Note that 2 pairs only open 164 and

167 nodes, that is contrasting with other pairs. Subjects don't used the notebook nor exercices stacks.

Answers and selected nodes

We decided not to consider any search strategy as more or less relevant in absolute terms. Instead, we studied the relationships between search patterns and learning outcomes (*i.e.*, students' ability to answer the questions). Overall, the students managed to provide acceptable answers in about two thirds of the cases, which is a first indication that they managed to get information out of the system. Following are the main observations concerning interaction protocols.

First, we observed that a great number of different routes were used among questions and students. Overall, the routes were not the most "economic" ones: only 36% of the selected nodes were directly relevant. The relevant nodes were selected 5 times on average. There was a negative relation between the total number of relevant nodes opened or the "paths economic rate" and subjects' performance (it should be quite considered as inverse ($r \approx -0,39$; $p < 0,1$)): Subjects who provided correct answers opened less relevant nodes than other subjects. This last result, that should mean that system does not help subjects to achieve the task, leads us to more finely analyze relationships between answers and selected nodes. We found that the production of a correct answer required the subject to open relevant nodes several times; however opening relevant nodes was no guarantee of a correct answer:

We located q question / item pairs[1] (q = 139): for each question Qi, we located ni correponding relevant items, that is for each question there are ni question / item pair.

$q = ni + nj + ...$

For each subjects pair, each question / item pair can be:

answer	*right*	*partial*	*ialse*
item owned by an opened node	O.R.	O.P.	O.F.
item not owned by an opened node	NO.R.	NO.P.	NO.F.

We gave the value c to the pairs, following that way:

for O.R. and NO.F. , c = 1

for NO.R. and O.F. , c = 0

for O.P. and NO.P. , c = 0,5

We computed the rate (T) "observed corresponding" item / answer where:

0 indicate inverse corresponding

0,5 no-corresponding

[1] An item is a "data" on a node: Figure, sentence, equation. One node can own many "items".

1 perfect corresponding

This rate indicates the system role (positive, absent, negative) in the subjects' answers.

$$T = \frac{\sum c_i}{Nq}$$

T = 0,64: it is significatively different of 0,5 ($t = 11,83$; $p < 0,001$) but nearer 0,5 than 1. For 7 questions T < 0,5; those questions are also less succeeded (note between 1,55 and 2): opening some nodes that we believed relevant could have involved false answers. For 4 animation stacks T < 0,5; but there is no regularity of the opening number: 1 hasn't been much seen (average opening = 0,86) by only 9 pairs, 1 hasn't been much seen (average opening = 2,23) by 17 pairs, 1 has been fairly seen (average opening = 5,91) by 19 pairs, 1 has been much seen (average opening = 9,45) by 18 pairs.

We computed the average number of relevant selection (*RS*) for the 32 task-relevant nodes. *RS* was defined as the average number of selections of a relevant node for the subset of subjects who provided correct answers, minus the same measure for the subset of subjects who did not answer correctly.

$$RS = \frac{\frac{\sum oc_i}{N} - \frac{\sum oi_j}{N}}{n}$$

The average *RS* was 3,17. Negative *RS* values are corresponding to cases where $T < 0,5$. *RS* tended to be higher for subjects who opened many orientation nodes.

<u>Paths description</u>

We describe a path as a sequences series, and each sequence as an opening series of same or alternate type of card. For example, opening two orientation cards is an orientation sequence of lenght 2 ($\lambda = 2$); opening alternatively orientation / explanation / orientation / explanation cards is an orientation / explanation of lenght 4 ($\lambda = 4$).

Average lenght of sequences = 4,6 (see details in table 2).

	average number of sequences	*average l*
Orientation	18,4	3,6
Explanation	9,8	20,6
Animation	30,3	5,1
Orientation / Explanation	15	1,31
Orientation / Animation	44,7	2,6
Animation / Explanation	6,2	1,2

Table 2. *Number and length of the different kind of paths sequences*

This result hide a great diversity paths. For example, 6 pairs conduced 10 sequences Animation / Explanation, whereas all others conduced less than 6 of these sequences, 11 of them conducing less than 2. The number of Animation sequences is between 0 and 43.

In figure 6, we see that 3 pairs are breaking free from the 19 others, concerning the kind of path.

Two of these pairs conduced sequences which average lenghts were upper than 10. We have already noticed that two pairs only opened 164 and 167 nodes.

It is usual to distinguish the following paths: “depth first search” and “breadth first search”. “Depth first search explores new nodes as far as possible and backtracks when a node does not lead to new nodes. Bread first

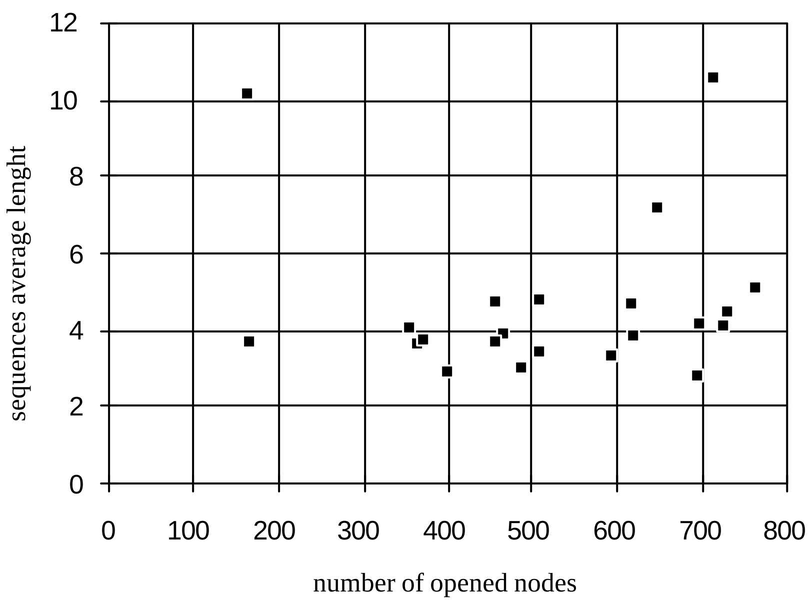


Figure 6. *Relationship between opened nodes and sequences lenght*

search explores all the nodes reachable from the initial node first, then all the nodes reachable from these nodes, and so on" (Lai et Manber, 1991, p.126). We rather like to give some shade and to distinguish (Tricot, 1993):

Orientation: user runs through orientation stacks (= orientation sequence with $\lambda > 4$ in our system); orientation should be usefull to get one's bearing, and to process local relations between nodes.

Topic path: user runs within one explanation stack (= explanation sequence with $\lambda > 4$)); topic path is quite like linear reading of contents (in a book for example).

Focusing: user jumps from one stack to another stack which concerns the same topic (= Animation and Animation / Explanation sequences); focusing corresponds to process contents as scanning, exploring.

Problematic re-boot: user jumps from one stack to another stack throught orientation cards (= Animation / Orientation or Explanation / Orientation sequences followed or preceded by a "Topic path" or a "Focusing").

Surface navigation: user jumps from one stack to another stack throught orientation cards opening only one card each. (Animation / Orientation or Explanation / Orientation sequences followed or preceded by Orientation, Animation or Explanation short sequences); surface navigation which is a kind of "sweeping contents", should allow subjects to draw up partial synthesis.

In our experiment "surface navigation" paths represent a third of all paths (see results in table 3). This type of move might have an orientation function, a function that may not be fulfilled by the so-called "orientation stacks" included in the system (*i.e.* series of cards which contain information about the system organisation).

Some pairs, using little "orientation paths", use very often "surface navigation" paths (more than 40% and sometimes even more thant 50%).

Navigation styles?

It is possible to classify pairs depending weither they openned more orientation, explanation or animation nodes; it is also possible to classify pairs according to they use preferentialy one or other type of path (see Levelt, 1982, for example). Another problem is to know if the definition of "navigation strategy" is relevant or not in hypermedia environments.

Type of path	*openned cards*
Orientation	17,0%
Topic path	29,3%
Focusing	9,6%
Problematic re-boot	11,9%
Surface navigation	32,1%

Table 3. *Openned cards and type of path*

a) We classified pairs according to the ratio of orientation, explanation or animation nodes they openned. We choose the 1 thrid threshold, so that we get 3 pairs "orientation", 5 pairs "explanation", 6 pairs "orientation / explanation", 6 pairs "animation / orientation", 1 pair "animation / explanation" and 1 pair "animation / explanation / orientation". Separately, the factors "explanation" and "orientation" have a positive effect on the number of openned cards (respectively $F(1,18) = 6{,}9$; $p < 0{,}02$ and $F(1,18) = 5{,}7$; $p < 0{,}03$); in other words, using this type of card is more costly.

For the rate RS, there are big contrasts (cf. Table 4), but factors have not significant effects.

"Animation" factor has an effect on performance ($F(1,18) = 4{,}7$; $p < 0{,}05$) and "orientation" factor has (nearly!) an effect on economy ($F(1,18) = 4{,}0$; $p \sim 0{,}06$).

b) We also classify pairs according to the fact that they use one or other type of path, but we haven't obtained signifivative results.

N > 30%	*# of pairs*	*RS rate*
orientation	3	4,05
explanation	5	2,24
orientation / explanation	6	3,29
animation / orientation	6	1,76
animation / explanation	1	2,20
animation / orientation / explanation	1	4,70

Table 4. *Number of relevant opened nodes corresponding to right answers according to the fact that pairs openned more orientation, explanation or animation nodes.*

Discussion

The objectives of this study were to evaluate the hypermedia-CAL application and to analyse the navigation of students confronted with a complex learning task. We believe that this type of situation (students learning through interaction with a complex information system) will become more and more frequent at least in self-education systems and libraries.

We observed that most subjects went through the task succesfully. However, there was no strict corespondance between opening relevant nodes and correct answers. In fact, looking at relevant information seems to be a necessary but unsufficient condition. Also we found that opening relevant nodes several times was associated with correct answers. Thus, it seems that redundancy in the learner's routes (*i.e.*, "looping") can have a positive effect, although it is sometimes interpreted as a negative symptom.

Our experiment also showed that interaction protocol can be interpreted only on light of a model of the task. There have been many references to task models in hypermedia litterature but frequently researchers use "ideal" task models, *i.e.* model of the task as it would be completed by a perfectly efficient system. Ideal task models are relevant only in cognitively simple situations where the subject can reach a minimal level of efficiency (speed, accuracy, performance, etc.).

More generally it is important to consider both a formal model of a task, which predicts the most efficient way to perform the task, and a model of the activity, which takes into account the subjective complexity of the task and the constraints of the human information processing system. However, in the case of learner-hypermedia interactions, this type of cognitive model remains to be developed.

Bibliography

Anderson, J.R. (1990). *The adaptative character of thought*. Hillsdale, NJ: Lawrence Erlbaum.

Bastien, C., Pélissier, A., & Tête, A. (1990). A experiment in learning logical negation by 6- and 7-year-old childre. *CPC: European Bulletin of Cognitive Psychology, 10*, 45-63.

Chatillon J.-F. (1988). *La régulation représentative des actes complexes. Hypothèses et expériences*. Thèse de doctorat de l'Université de Provence, Aix en Provence.

Chatillon, J.-F., & Baldy, R. (1994). Performance motrice et développement moteur, les liens au developpement cognitif. *Enfance*, *2*, 1994, 299-319.

Coste, J.-P. (1991). Gestion de stratégie d'accès à l'information. *2° journées EIAO de Cachan,* Ecole Normale Supérieure de Cachan, 24-25 Janvier 1991.

Coste, J.-P. (1993). Stratégies hypertextuelles et métaphores de stratégies. In G. Baron, J. Baudé & B. de La Passardière (Eds.), *Hypermédias et Apprentissages 2*, Actes des 2° journées scientifiques, Lille (pp. 153-168). Paris: Presses de l'INRP.

Devichi, C. (1994). *Représentation de la tâche et performance dans des tâches de sériation de longueur*. Thèse de Doctorat de l'Université de Montpellier, Montpellier.

Foss, C.L. (1989). Detecting lost users: Empirical studies on browsing hypertext. *Rapport de Recherche INRIA n° 972,* Sophia Antipolis.

Gray, S.H. (1990). Using protocol analyses and drawing to study mental model construction during hypertext navigation. *International Journal of Human-Computer Interaction, 2* (4), 359-377.

Gray, S.H., Barber, C.B., & Shasha, D. (1991). Information search with dynamic text vs paper text: an empirical comparison. *International Journal of Man-Machine Studies, 35*, 575-586.

Lai, P., & Manber, U. (1991). Flying through hypertext. *Hypertext'91 Proceedings*, San Antonio (pp. 123-132). New York, NY: ACM Press.

Levelt, W.J.M. (1982). Cognitives styles in the use of spatial direction terms. In R.J. Jarvella & W. Klein (Eds.), *Speech, place and action* (pp. 251-267). Chichester: Wiley.

Leventhal, L.M., Mynatt Teasley, B., Instone, K., Schertler Rohlman, D., & Farhat, J. (1993). Sleuthing in HyperHolmes: an evaluation of using hypertext vs a book to answer questions. *Behaviour & Information Technology, 12* (3), 149-164.

Mohageg, M.H. (1992). The influence of hypertext linking structures on the efficiency of information retrieval. *Human Factors*, *34* (3), 351-367.

Newell, A., & Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Pélissier, A., & Tête, A. (1994). Analysis of the logical characteristics of a complementation task. *International Journal of Psychology, 29* (2), 249-258.

Redish, E.F., Wilson, J.M., & McDaniel, C. (1992). The CUPLE project: a Hyper- and multimedia approach to restructuring physics education. In E. Barrett (Ed.), *Sociomedia. Multimedia, Hypermedia and the social construction of knowledge* (pp. 219-256). Cambridge, MA: MIT Press.

Spiro, R.J., Feltovitch, P.J., Jacobson M.J. & Coulson, R.J. (1991). Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology*, 31(5), 24-33.

Tricot, A. (1993). Ergonomie cognitive des systèmes hypermédia. *Actes du Colloque de prospective "Recherches pour l'Ergonomie"*, CNRS - PIR Cognisciences, Toulouse (pp. 115-122).

Tricot, A. (1994). Recherche d'information dans des documents non-linéaires et récupération volontaire en mémoire. In *Actes du 1° Colloque "Jeunes chercheurs en Sciences Cognitives"*, La Motte d'Aveillans, 23-25 mars. (pp. 129-138).

Wright, P., & Lickorish, A. (1990). An empirical comparison of two navigation systems for two hypertexts. In R. McAleese & C. Green (Eds.), *Hypertext: State of the Art* (pp. 84-93). Oxford: Intellect Ltd.

Wright, P., & Lickorish, A. (1994). Menus and memory load: navigation strategies in interactive search tasks. *International Journal of Human-Computer Studies, 40*, 965-1008.