

Vers une Documentation Centrée Lecteur : un modèle général de lecture sur mesure

Azzeddine CHIKH

Enseignant Chercheur à l'Institut National d'Informatique

BP. 68 M. Oued Smar 16270 ALGER

E-mail: Chikh azzeddine@hotmail.com

RESUME :

cet article s'inscrit dans le thème des hypermédias adaptatifs et plus particulièrement dans les hyperdocuments adaptatifs. Dans notre conception, un modèle général de lecture sur mesure est composé d'un modèle d'hyperdocument, d'un modèle de lecteur et d'un processus d'adaptation. Le modèle d'hyperdocument regroupe deux espaces, un espace de documents et un espace de connaissances du domaine. Le modèle de lecteur est composé du profil du lecteur, ses connaissances, ses objectifs, ses préférences et l'historique de ses lectures. Le processus d'adaptation transforme le modèle d'hyperdocument en fonction du modèle de lecteur. Ce travail ne représente pas une solution complète au problème de lecture sur mesure, mais il tente de bien le définir et de lui associer un modèle général.

Mots clés : modèle de lecteur, modèle d'hyperdocument, problème de lecture sur mesure, modèle général de lecture sur mesure.

Abstract: this paper is registered in the adaptive hypermedia theme and specially in the adaptive hyperdocuments. In our conception, a general tailor-made reading model, is composed of hyperdocument's model, reader's model and adaptation process. The hyperdocument's model regroups two spaces, a document space and a domain knowledge space. The reader's model is composed of his profile, his knowledge, his goals, his preferences and his reading historic. The reading process transforms the hyperdocument's model according to reader's model. These work no represents a complete solution to tailor-made reading problem, but it tents to well define it and associate him a general model.

Key words: reader's model, hyperdocument's model, tailor-made reading problem, general tailor-made reading model.

1.Introduction

Lire est une opération consistant à construire une unité de sens à partir de la source consultée. Cette opération est devenue très complexe avec l'apparition du document électronique et plus particulièrement de l'hyperdocument. En effet, on assiste aujourd'hui, grâce au multimédia, à Internet, etc., à une croissance exponentielle du nombre et du volume des hyperdocuments, de la variété des informations qu'ils véhiculent. De même, le nombre et la variété des lecteurs et de leurs intérêts augmentent sans cesse. Malheureusement, les hyperdocuments actuels, bien qu'offrant des techniques de navigation intéressantes (backtacks, fisheyes, guided tours, filters, etc.), ne sont pas réellement adaptés aux lecteurs. Ils sont souvent rédigés selon une logique de l'auteur ou du domaine qui n'est pas forcément compatible avec celles des lecteurs. Ils proposent ainsi, soit des mécanismes de lecture complètement libres (problème de désorientation et de surcharge cognitive) soit des mécanismes de lecture directifs (approche très restrictive), alors qu'il serait intéressant de pouvoir les adapter en fonction des caractéristiques du lecteur (son profil, ses intérêts, ses motivations, ses connaissances, etc.).

Les stratégies de lecture n'ont pas suscité beaucoup de travaux de recherche, comparativement aux processus de compréhension en lecture [1]. Goldman et Saul [2] identifient un nombre de stratégies de lecture (once through, review, regress, etc.). Ils ont trouvé que les lecteurs utilisent une variété de ces

approches pour lire un document. Il s'agit en fait d'une stratégie flexible qui est déterminée par un ensemble de facteurs qui incluent: les connaissances du lecteur sur le domaine, les objectifs du lecteur, etc. Cependant le lecteur doit tout de même maintenir l'objectif prédominant de naviguer à travers le document d'une manière cohérente.

L'originalité de ce travail est de définir les problèmes de lecture sur mesure, d'identifier leurs concepts et les modèles nécessaires à sa compréhension et à sa résolution. Après cette introduction, la seconde partie de cet article présente quatre travaux qui visent à adapter les hyperdocuments aux usagers en utilisant des approches différentes. Ensuite, la troisième partie évoque les approches de modélisation qui concernent le lecteur et l'hyperdocument. Les deux modèles d'hyperdocument et de lecteur, le problème de lecture sur mesure ainsi que le modèle général associé sont présentés dans la quatrième partie. Enfin, la conclusion évoque la nécessité de compléter cette étude par une définition formelle des différents modèles proposés.

2. Autres travaux

Nous présentons maintenant quatre travaux qui s'intéressent à l'adaptation des hyperdocuments aux usagers.

Le premier travail se base sur le contrôle des actions de l'utilisateur (Monitoring user actions), [3,4]. L'analyse automatique d'une séquence d'actions permet au système de chercher des courts-circuits possibles à l'intérieur du chemin de recherche. Cette analyse est basée sur des règles dérivées à partir de données collectées précédemment. Cependant le système doit s'assurer que l'ordre des actions est intentionné. La solution consistant à poser des questions aux utilisateurs n'est pas intéressante. En effet, il y a un risque que les réponses ne soient pas correctes ou que l'utilisateur soit surchargé.

Le second travail est la modélisation de haut niveau (high level modeling) [5,6,7,8,9] qui associe à l'hyperdocument un réseau sémantique contenant les différents concepts du domaine et leurs relations. Il présente les avantages suivants:

- La sémantique d'une relation entre deux nœuds peut être dérivée à partir du réseau sémantique. Ainsi les types de lien peuvent être déterminés et fournis avec exactitude.
- Les liens sont générés dynamiquement. Les documents peuvent ainsi être modifiés au besoin, par exemple à la demande du lecteur.
- Le profil de l'utilisateur et ses choix peuvent être pris directement en compte.

Le troisième travail est la spécification et formalisation d'un modèle de document réactif pour expertexte [10]. Il s'inscrit dans le contexte des documents actifs et propose une approche pour une meilleure adéquation à l'utilisateur afin de réduire, voire de supprimer les risques de désorientation et de surcharge cognitive. Cette démarche fait appel à des techniques issues des différents domaines : bases de données, intelligence artificielle et paradigme objets. Elle vise à contrôler le mécanisme de navigation par l'intermédiaire d'un mécanisme d'aide non intrusive. Il s'agit de développer un mécanisme de réorganisation dynamique des informations aussi bien au niveau de la structure du réseau qu'au niveau du contenu des nœuds afin de présenter une information adaptée à l'utilisateur. La contribution principale de ce travail consiste donc à définir non plus en terme d'outils mais en terme de fonctionnalités, le concept d'expertexte puis à le modéliser. Ainsi le modèle d'expertexte est fondé sur l'intégration d'une composante Hypertexte qui fournit une interface conviviale et la base d'informations, et d'une composante experte construite sur un modèle de système tutoriel intelligent.

Le quatrième travail est la navigation hypertextuelle et capitalisation des connaissances de spécialité [11]. Il s'agit d'une méthode d'aide à la construction de parcours de lecture, constituant une unité de sens, en fonction des spécifications de l'utilisateur. Cette modélisation comprend trois aspects principaux: les techniques annotatrices (qui témoignent du parcours de lecture et lui donnent sens), les modèles

terminologiques (qui témoignent du type de lecture illustré par les types et modes d'annotations) et les modèles de sémantique hypertextuelle (qui témoignent de l'objectif de la lecture illustré par son parcours). L'utilisateur est ainsi assisté par des fonctionnalités d'annotation qui lui permettent d'écrire sa lecture, des fonctionnalités de synthèse des annotations qui lui permettent de voir sa lecture et des fonctionnalités d'extraction de structure qui lui permette de naviguer à travers sa lecture.

3. Approches de modélisation

Dans le but de contribuer à cerner le problème de la lecture sur mesure, les approches de modélisation que nous proposons visent à modéliser le lecteur d'une part l'hyperdocument d'autre part.

3.1. Modélisation du lecteur

La modélisation du lecteur (figure 1) se trouve au centre de quatre approches : l'interaction homme-machine (IHM), les sciences cognitives (SC), les environnements d'apprentissage assisté par ordinateur (EIAO) et les systèmes de recherche d'information (SRI). En effet, le lecteur d'un hyperdocument est devenu l'utilisateur d'une application logicielle. De ce point de vue, la lecture a évolué vers l'interaction homme – machine (IHM) [12]. Le domaine de la communication homme –machine permet d'obtenir des systèmes faciles à utiliser [13]

Les sciences cognitives aident les auteurs à créer des documents faciles à explorer par les lecteurs [13]. Cette aide devient plus importante quand on considère des documents actifs, comme ceux utilisés par exemple dans la conception des artefacts [12].

Dans les environnements interactifs d'apprentissage assisté par ordinateur (EIAO) [14], l'apprenant construit ses connaissances par sa propre activité et l'enseignant guide cette activité sur un plan essentiellement méthodologique. En d'autres termes, il s'agit de dépasser l'approche tutorielle héritée des méthodes d'enseignement conventionnelles pour laisser une part plus importante aux démarches de type constructiviste au travers des systèmes du type environnement d'apprentissage. Ces environnements éducatifs se basent sur la supervision souple introduite par les tuteurs intelligents mariés aux potentiels des hypertextes, en terme d'interfaces et de navigation.

Enfin dans les systèmes de recherche de l'information (SRI), [15] on considère que l'utilisateur sait avec précision ce qu'il cherche. Ce dernier formule alors des requêtes exprimant un besoin en informations. Dans le cas où la base documentaire est organisée en hypertexte, le système tente d'extraire un nœud, un ensemble de nœuds ou une partie du réseau susceptible de fournir une réponse à la question posée. Deux mécanismes de recherche sont couramment employés, la recherche par contenu et la recherche par structure.

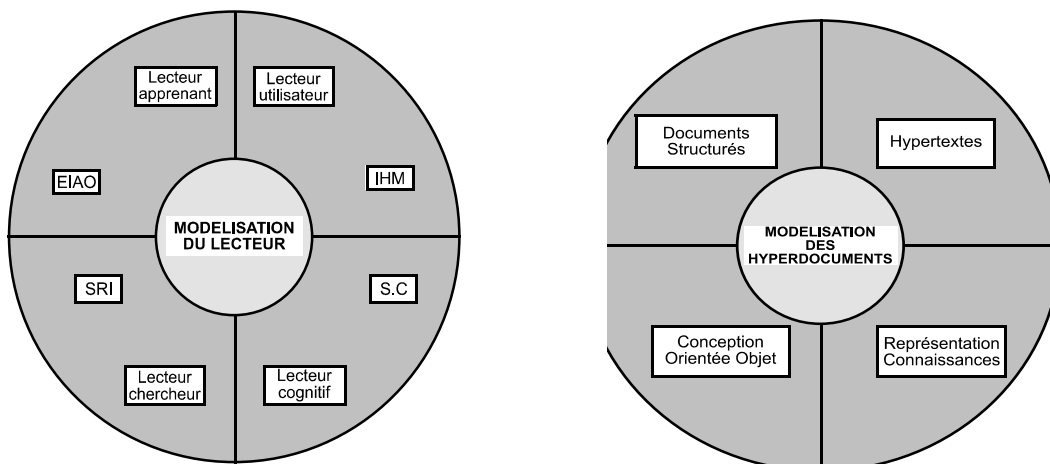


Figure 1: Modélisation du lecteur

Figure 2: Modélisation des hyperdocuments

3.2. Modélisation de l'hyperdocument

La modélisation des hyperdocuments (figure 2) peut être traitée par la combinaison des quatre approches : des documents structurés, des hypertextes, de la représentation des connaissances et de la conception orientée objets.

Dans l'approche «documents structurés» [16], la séparation de la représentation interne du document de la façon dont il sera présenté augmente la flexibilité et les possibilités de réutilisation du document. L'identification et la localisation précise de chaque information au sein du document permet d'automatiser l'exploitation de l'information contenue dans ce dernier, telle l'alimentation des applications de recherche documentaire et de consultation interactive des documents.

Par ailleurs, l'approche Hypertexte est particulièrement intéressante pour répondre à des exigences de l'Utilisation Assisté par Ordinateur de documents (UAOD) [17]. L'association des approches «documents structurés» et «hypertextes» conduit au concept «hyperdocument» [18].

Certains travaux comme [19,20,21] ont permis aux hypertextes de s'étendre à des techniques issues des bases de connaissances. Parmi les fruits de cette extension, nous citons la définition d'un modèle conceptuel d'expertexte [10], à partir d'une composante hypertexte et d'une composante experte qui communiquent par un dialogue événementiel.

Enfin, l'approche orientée objets offre une puissance de modélisation et un environnement de gestion particulièrement adaptés aux «nouvelles applications» nécessitant la prise en compte d'objets volumineux à structure complexe tels que les documents [22].

4. Lecture sur mesure : proposition d'un modèle général

Nous commençons d'abord par définir les modèles d'hyperdocument et de lecteur, considérés comme des modèles de bases, dans notre contexte de modélisation. Ensuite nous posons le problème de lecture sur mesure et nous lui associons le modèle général de lecture sur mesure, objet de cet article.

4.1. Le modèle d'hyperdocument

Le modèle d'hyperdocument (HD) est composé d'un espace de document (D) et d'un espace de connaissances du domaine (K). Le premier représente l'information proprement dite alors que le second représente les concepts du domaine sous jacents.

$$\mathbf{HD = (D,K)}$$

Les deux espaces de documents (D) et de connaissances de domaine (K) sont composés des éléments de structuration suivants :

- (C) désigne le contenu. Il est constitué des portions de contenu de base tels que le texte, le son, l'image fixe ou animée. Chacun de ces objets est codé selon une technique qui lui est particulière;
- (SP) désigne la structure physique. Elle est relative à la décomposition en partie plus petites (page, cadre, bloc) qui trouvent leur place dans une feuille (ou page) sur écran;
- (SL) désigne la structure logique. Elle est relative à la décomposition en parties toujours plus petites, compte tenu de la signification du contenu;

- (SR) désigne la structure du réseau. Elle désigne les liens hypertextes dans le cas de l'espace document et le réseau sémantique dans le cas de l'espace de connaissances du domaine.

4.2. Le modèle de lecteur

Le modèle de lecteur (L) est composé de cinq parties, le profil du lecteur (P), les connaissances (K), ses objectifs (O), ses préférences (F) et l'historique de ses lectures (H)

$$L = (P, K, O, F, H)$$

Nous expliquons dans ce qui suit chaque élément de ce modèle :

- (P) le profil du lecteur contient l'identité du lecteur, les stéréotypes (modèles de groupe) auxquels il appartient, son niveau scientifique, ses spécialités, ses langues, ses centres d'intérêts, etc. Au fait il s'agit de toute information sur le lecteur, qui peut être utile et qui est indépendante du domaine;
- (K) les connaissances du lecteur sur le domaine. C'est le sous ensemble des concepts du domaine qui est connu par le lecteur, on parle de modèle de couverture [23];
- (O) les objectifs du lecteur qui concernent le domaine de l'hyperdocument et l'activité du lecteur;
- (F) les préférences du lecteur qui traduisent ses choix en matière de lecture. Il s'agit notamment de choisir les portions de contenu à lire, les liens à activer, le niveau de détail souhaité, ... etc;
- (H) l'historique de lecture qui constitue une mémoire sur les lectures précédentes, accompagnées, éventuellement d'annotations du lecteur. Ces dernières lui permettent «d'écrire sa lecture».

Les cinq éléments du modèle de lecteur précités peuvent être définis indépendamment d'une session particulière. On parle alors de modèle de lecteur général. Dans le cas contraire on parle de modèle de lecteur contextuel. Les modèles contextuels du lecteur constituent une mémoire à court terme, qui alimente le modèle général du lecteur. Ce dernier constitue la mémoire à long terme.

4.3 Le modèle général de lecture sur mesure

Poser un problème de lecture sur mesure d'un hyperdocument (figure 3) consiste à définir un modèle d'hyperdocument (HD) et un modèle de lecteur (L) en vue d'adapter le premier au second.

Etant donné un problème de lecture sur mesure d'un hyperdocument [HD, L], on appelle modèle général de lecture sur mesure (figure 3) associé à ce problème le triplet :

$$[L, HD, A]$$

Où,

- (L) désigne le modèle de lecteur défini précédemment;
- (HD) désigne le modèle de l'hyperdocument défini précédemment;
- (A) désigne le processus d'adaptation qui transforme le modèle d'hyperdocument (HD) en fonction du modèle de lecteur (L). Le processus d'adaptation de la lecture décrit [23];
- Les techniques d'adaptation utilisées, les explications supplémentaires, les explications pré requises et les explications comparatives, les variantes d'explication, le tri des fragments d'information au sujet d'un concept donné, le guidage direct, le tri adaptatif, la dissimulation des liens, l'annotation adaptative ou l'adaptation des cartes;
- Les critères d'adaptation : il s'agit des éléments du modèle de lecteur nécessaires aux techniques

d'adaptation utilisées;

- Les éléments d'adaptation : ce sont les éléments du modèle d'hyperdocument, les portions de contenu (présentation adaptative) et les liens (support de navigation adaptative), qui sont retenus par les techniques d'adaptation utilisées.

En toute généralité, traiter un problème de lecture sur mesure revient à construire le triplet précédent.

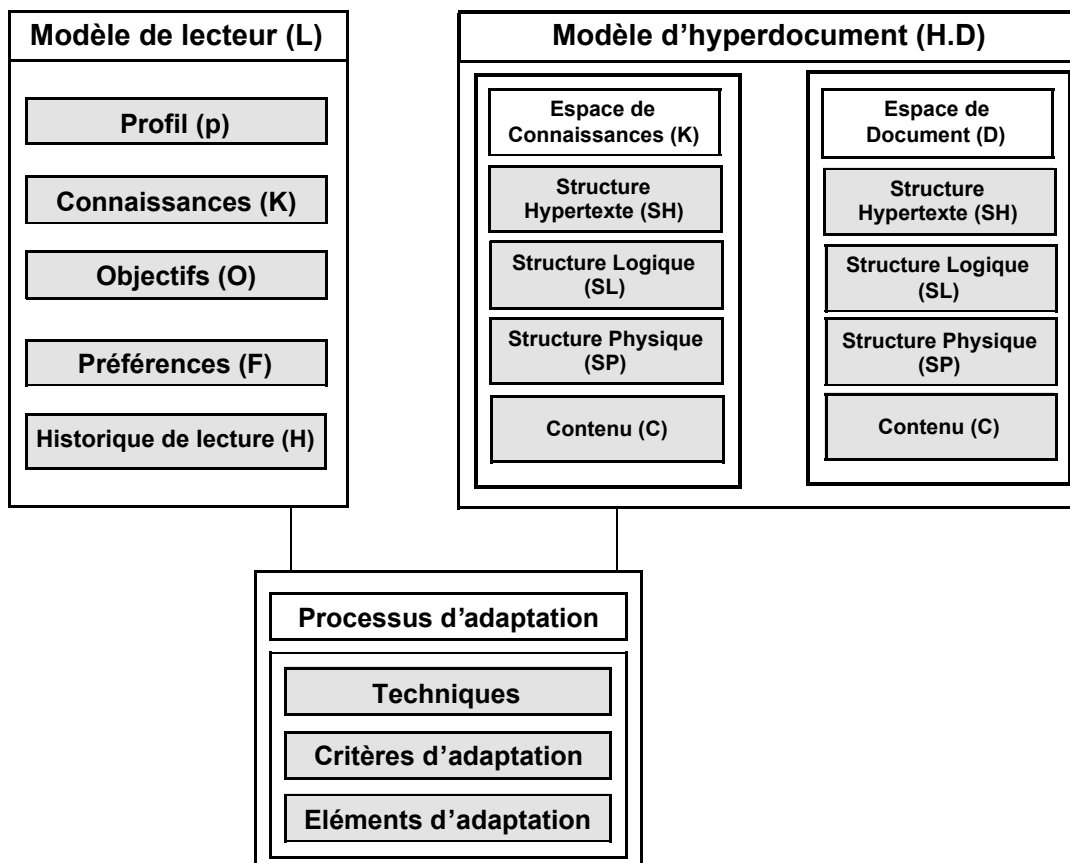


Figure 3. Modèle général de lecture sur mesure

5. Conclusion

Le but de cet article est de bien définir et modéliser le problème de lecture sur mesure.

Cette modélisation passe par la conception de trois modèles : le modèle de l'hyperdocument, le modèle de lecteur et le modèle général de lecture sur mesure. Le modèle de l'hyperdocument est caractérisé par la distinction de deux espaces de document et de connaissances de domaine. Le modèle de lecteur permet de prendre en charge des lecteurs typés (stéréotypés) ou non. Il est composé de cinq parties, le profil du lecteur, ses connaissances, ses objectifs, ses préférences et l'historique de ses lectures. Le modèle général sur mesure, composé des deux modèles précédents et d'un processus d'adaptation, est le cadre formel qui permet une analyse logique et transparente du processus de lecture sur mesure. Enfin nous restons conscients qu'une utilisation rationnelle de notre modèle nécessite une représentation, par exemple symbolique ou numérique, plus formelle des données modélisées.

Références bibliographiques :

- 1) FOLTZ P.W. «Comprehension, Coherence and Strategies in Hypertext and Linear Text» 1996, <http://www-psych.nmsu.edu/~pfoltz/reprints/HT-Cognition.html> fnB0
- 2) GOLDMAN S.R. & SAUL E.U., «Flexibility in text processing : A strategy competition model. Learning and Individual Differences», 2(2), 181-219p.
- 3) MYKA A. and GUNTZER U. and SARRE F. «Monitoring user actions in the hypertext system (hyperman)». In Going Online – Conference Proceedings of the SIGDOC 92 (Oct. 13-16, 1992, Ottawa, Canada), 1515 Broadway, New York 10036, 1992. The Association for Computing Machinery. 103-114p.
- 4) MYKA A. and SARRE F. and GUNTZER U. «Rule-Based machine learning of hypertext links». Upravlyaemye Sistemy Machiney, 1992, (7/8): 75-82p.
- 5) AGOSTI M., MELUCCI M., and CRESTANI F. « Automatic authoring and construction of hypermedia for information retrieval. Multimedia systems»1995,(3):15-24p.
- 6) COLLIER G.H.« THOTH – II : Hypertext with explicit semantics», In proceedings of the hypertext '87 , chapel Hill , November, ACM 1987, 269 –289p.
- 7) MAYFIELD J. and SNITCH C. N : « augmenting hypertext documents with a semantic net » , 1993.
- 8) RADA R. « Hypertext writing and document reuse : the role of a semantic net » Electronic publishing – Origination , Dissemination and Design, 1990 , 3 (3) : 125- 140 p.
- 9) WIESENER S.,KOWARSCHIK W., VOGEL P ., and BAYER R. « semantic hypermedia retrieval in digital libraries .», In To appear in : Advances in Digital Libraries, lecture Notes in Computer Science . Springer-Verlag , 1995.
- 10) GOUARDERES E» Spécification et formalisation d'un modèle de document réactif pour expertexte « Thèse de doctorat en informatique de l'université paul Sabatier de Toulouse du 4 juillet 1995
- 11) TERRAY P . « la navigation hypertextuelle et capitalisation des connaissances de spécialité : utilisation de techniques annotatrices , terminologiques et de sémantique hypertextuelle pour parcourir et élaborer un document Hypertexte». 1997 <http://www.hds.utc.fr/pmorizet/projets.recherche/projets-rech.html#pt>
- 12) BOY G.A. « Documents Actifs de conception « IHM 97 9ème journées sur l'ingénierie de l'interaction Homme – Machine du 10 au 12 septembre, 1997, 113 – 118p.
- 13) CHAUSS K. M. Rensselaer polytechnic Institute « Reader as User : Applying Interface design , Techniques to the Web», <http://english.ttu.edu/kairos/1.2/features/chauss> 1997
- 14) BARON M., GRS R ., NICAUD j.f. « environnements interactifs d'apprentissage avec ordinateur « Tome 1 , Editions Eyrolles , Paris , 19993.
- 15) SAVOY . J , DESBOIS D ., « hypertexte et réseaux d'inférences bayesiens », ICO 91, Montreal, 29 Avril –2 Mai 1991 , 140 – 148p.

- 16) COMPAROT J., POUSSIER C., « hyperbase : Formalisation et Architecture » thèse de doctorat en informatique de l'université Paul Sabatier de Toulouse du 31 Janvier 1994
- 17) CONCKLIN J., «hypertext: An introduction and Survey» IEEE Computer, 18 (9) , 1987, 17- 41 p.
- 18) STOTT P.D , FURUTA R. « Hypertexte 2000: Databases or Documents ?.» Electronic publishing , vol .4(2) , juin 1991, 119-121p.
- 19) DIAPER D., RADA R., «Expertext : Hyperizing expert systems and expertising hypertext « , in Hypermedia/Hypertext and Object – Oriented Databases , Ed Chapman &Hall , 1991
- 20) NANARD J ., NANARD M .,» Using Structured Types to incorporate knowledge in Hypertext «,Proceedings of the Third ACM Conference on Hypertext , San Antonio , Texas , 15- 18 Décembre1991, 329-343p.
- 21) NANARD J.,NANARD M ., MASSOTE A.M.,DJEMAA A, JOUBERT A.,BETAILLE H., CHAUCHE J.,» Integrating Knowledge – based Hypertext and Database for Task – oriented Access to Documents « ,DEXA ‘ 93, Actes, 1993, 721- 732p.
- 22) MEGHINI C.,RABTTI F.,THANOS C.»Conceptual Modeling of Mulimedia Documents « Computer , Octobre 1991, 23-30p.
- 23) BRUSILOVSKYP. « Methods & techniques of Hypermedia Systems» School of computer Science . Carnegie Mellon University. Pittsburg. 1996.