



Les hypertextes : au delà des liens, la connaissance

Marc Nanard

► **To cite this version:**

Marc Nanard. Les hypertextes : au delà des liens, la connaissance. Sciences et Techniques Educatives, Hermes, 1995, 2 (1), pp.31-59. <edutice-00001453>

HAL Id: edutice-00001453

<https://edutice.archives-ouvertes.fr/edutice-00001453>

Submitted on 27 Oct 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les hypertextes : au delà des liens, la connaissance

Marc Nanard,
Professeur des Universités au CNAM,
LIRMM CNRS
161 rue Ada, 34392 Montpellier cedex
mnanard@lirmm.fr

Résumé

Pour un lecteur, un hypertexte rend possible et aisée une lecture non linéaire d'un ensemble de documents, en navigant le long de liens entre documents. La notion de *lien* apparaît alors comme une généralisation de celle de référence.

Nous proposons une analyse plus profonde des mécanismes mis en jeu dans les hypertextes. Un lien est fondamentalement une relation sémantique entre deux informations. Ceci conduit à distinguer les rôles de deux entités distinctes, l'une perçue de façon classique par le lecteur, que nous appelons *information*, et l'autre utilisée par la machine pour permettre la navigation. Cette dernière est appelée *connaissance* (au sens informatique de représentation de connaissance). Elle matérialise pour la machine des relations sémantiques entre informations et sert à cette dernière à effectuer les rapprochements des informations présentées au lecteur lors d'une navigation.

Cette interprétation de ce qu'est un hypertexte permet de mieux comprendre le rôle de l'auteur d'un hyperdocument. En tissant un réseau de liens sur un ensemble d'informations, il lui ajoute sa propre connaissance, il valorise ces informations. Plus encore, ce modèle traduit bien le fondement de systèmes multi-webs tels Intermedia qui est utilisé dans l'enseignement. Il montre de façon simple comment des hyperdocuments distincts peuvent être construits sur un même ensemble d'informations par des auteurs différents. Un hyperdocument est une structure d'observation plaquée sur un ensemble de documents.

Abstract

From the reader's side, an hypertext helps him escape the sequential reading of documents, taking advantage of links to browse from one information to another. Thus links are often considered simply as references between documents.

In this article, we propose a deeper analysis of the mechanisms which take place in hypertexts. A link is a semantic relationship between information chunks. As a consequence, two entities are to be considered. On the one hand, information is presented to the reader as node contents. On the other hand, knowledge (with the A.I. meaning) is used by the machine to make browsing possible. Semantic relationships between information chunks represent the knowledge, which is used by the computer to make information closer .

This view of hypertext helps understanding what an author does when he or she designs an hypertext. The web of links that an author creates over information adds value to it and reflects the author's own knowledge. This model also shows very simply the bases of systems such as Intermedia which clearly dissociate the document base from the

hypermedia network. This allows several people to express distinct knowledge about a given information base. An hypertext is an organized view of an information set.

Mots clés : Hypertexte, liens, connaissances.

Key Words: Hypertext, links, knowledge.

1. INTRODUCTION

Les sauts technologiques qui ont ponctué l'histoire de l'humanité ont été chaque fois la cause de profondes mutations culturelles et sociales qui ont eu des répercussions importantes sur la façon de transmettre le savoir. L'écriture, l'imprimerie puis la télévision en sont des exemples triviaux. L'avènement des technologies de la communication et des hypermédia constitue un de ces sauts technologiques majeurs.

L'évolution de la quasi totalité des innovations technologiques passe par quatre phases principales. Une première phase de gestation scientifique permet d'élaborer et de rassembler les fondements nécessaires à la réalisation. Ensuite vient une phase d'émergence, où un nombre limité de prototypes ou de produits préindustriels voient le jour sans connaître un essor économiquement viable et contribuent à la maturation de la technique. Enfin la phase d'expansion, où la technique étant maîtrisée, son essor au sein de la société devient significatif. Cette phase s'accompagne sur le plan culturel d'une maîtrise progressive de l'usage de la technique, donnant éventuellement naissance à des formes d'art nouvelles. Enfin, vient une phase de vieillissement dans laquelle d'autres technologies prennent progressivement le pas sur la précédente et contribuent à son extinction progressive. Même si ces phases se chevauchent largement, elles reflètent l'âge relatif d'une technique dans la société.

Comme l'avait rappelé Ian Ritchie dans son exposé d'ouverture de la conférence ECHT'92 [Lucarella 92], un ensemble de techniques apparentées, désignées sous les noms d'hypertexte, hyperdocuments, multimédia, hypermédia commence tout juste à entrer dans la phase d'expansion. Au cours de ces deux dernières années l'essor brutal des publications hypermédia disponibles, la disponibilité de matériels bon marché supportant le multimédia, l'essor tout aussi spectaculaire de l'accès direct à des informations réparties au moyen de World Wide Web, sont des événements majeurs pour l'évolution de la société, et tout particulièrement pour les techniques éducatives.

Cet article fait le point sur les techniques sous-jacentes, et montre ce que les résultats récents en recherche dans ce domaine peuvent apporter aux applications éducatives. Pour cela nous commençons par mettre en évidence ce qui unifie le domaine des hypertextes, et en dégageons un modèle fonctionnel. Nous étudions ensuite trois grandes classes d'applications hypertextes, qui correspondent à des attitudes différentes de l'utilisateur face à l'information, et les situons dans le contexte des activités éducatives. Nous montrons comment, indépendamment de leurs différences de finalité, les systèmes hypertextes correspondants répondent à ce modèle fonctionnel. Enfin, à titre d'illustration, deux systèmes sont présentés plus en détail.

2. QU'EST CE QU'UN HYPERTEXTE?

Cette première section a pour but de présenter le domaine des hypertextes sous son aspect fonctionnel. Usuellement, les hypertextes sont surtout connus grâce à

l'automatisation de l'exploitation de références entre documents qu'ils rendent possible. La notion de *lien* est souvent considérée comme une généralisation de celle de référence. Celle de *nœud* désigne les documents élémentaires. Un lien permet de passer directement d'un nœud à un autre. L'assimilation de la notion de lien à celle de référence est une simplification qui ne s'applique qu'aux hypertextes les plus simples.

Nous proposons une modélisation moins technique mais plus unifiée de la notion d'hypertexte. Elle repose sur la séparation entre le contenu informationnel brut porté par les documents présents dans l'hypertexte et les connaissances explicitement disponibles dans la machine pour permettre l'exploration de l'hypertexte. Ces dernières sont responsables des aspects dynamiques de l'hypertexte. Les liens sont des relations sémantiques entre informations. La navigation correspond au rapprochement d'informations reliées par une relation donnée. En séparant information et connaissances, l'activité de l'auteur d'un hypertexte est mieux cernée.

Mais tout d'abord quelques précisions sur le vocabulaire sont souhaitables.

2.1. Terminologie : Hypertexte, Multimédia, Hypermédia

Le domaine des hypertextes est actuellement typiquement dans la période où l'émergence scientifique et le début de l'expansion économique se chevauchent, amplifiant ainsi les écarts entre les possibilités des prototypes de recherche et celles des produits commerciaux disponibles. Bien que certains efforts d'unification du vocabulaire et de formalisation de sa signification ont été couronnés de succès [Halasz 94], il n'y a toutefois pas encore de définition universelle de ce qu'est un hypertexte.

Le terme *hypertexte* a été introduit par Ted Nelson [Nelson 65]. Nous ne nous attardons pas sur les aspects historiques des hypertextes, le lecteur intéressé peut se reporter par exemple à [Nielsen 90]. On trouve également dans [Bruillard 94] une présentation de l'intérêt des hypertextes dans l'éducation. La notion d'hypertexte évoque aujourd'hui principalement une organisation permettant la lecture non linéaire des documents à l'initiative du lecteur. Elle n'est pas limitée au seul aspect textuel. Elle concerne toutes les formes de documents, y compris bien entendu les formes multimédia [Garzotto 94]. C'est pour cela que le terme d'hypermédia le remplace progressivement. Fondamentalement les hypermédiés sont des applications qui gèrent et exploitent des *relations sémantiques* entre des documents multimédiés pour permettre à un lecteur d'atteindre très facilement une information cherchée.

La notion de *multimédia* concerne l'aptitude pour une machine (et une application) à gérer plusieurs canaux de communication sensorielle avec l'utilisateur, -usuellement le son et l'image animée-. La notion de multimédia n'implique nullement celle d'hypermédia [Nanard 94-c]. La présence d'une image animée et de son n'est pas le garant d'une véritable interactivité entre usager et machine. Remplacer une image par une animation vidéo et ajouter un fond sonore dans un document en fait un document multimédia mais pas un hypermédia.

Dans cet article, nous continuons à employer le terme hypertexte quel que soit le type de média utilisé dans des documents.

2.2. Un essai d'unification

Si l'on cherche à caractériser ce qu'ont en commun des systèmes hypertextes aussi différents que KMS, SEPIA, VIKI, Superbook, MacWeb, Guide, WWW, etc., on constate qu'il est quasi impossible de les comparer tant sur le plan fonctionnel, que sur

celui de leur architecture interne effective. Seule l'intentionnalité et l'esprit de l'approche de l'interaction homme-machine leur est commune. Hypertexte, étymologiquement évoque 'plus que' 'texte'. 'Plus' ne doit pas être interprété comme 'plusieurs' textes (interconnectés) mais bien comme texte plus un élément d'une autre nature.

Un hypertexte est fondamentalement une entité composite (au sens des matériaux composites de l'industrie) composé d'un ensemble de *documents* et d'une seconde entité, qui, en termes informatiques, est appelée *connaissance*, les deux étant fortement connectées par un mécanisme dit *d'ancrage*. La nature 'matériau composite' des hypertextes est la raison profonde de leur force et de leur souplesse. Elle permet de concilier les avantages et les contraintes propres aux deux partenaires de l'interaction.

Examinons les deux composants.

- Les documents sont des entités fondamentalement destinées à l'homme. Ils constituent l'unique surface d'échange entre l'homme et l'hypertexte. L'homme les voit tels quels et les interprète¹ directement. Les documents, par leur nature familière ont un pouvoir évocateur très fort pour le lecteur. Ils véhiculent ainsi une très forte quantité d'information, en particulier par le non-dit qui leur est associé. Le lecteur apprendra beaucoup plus sur l'atmosphère méridionale en suivant une interprétation d'une scène de Pagnol par Raimu sur une station multimédia qu'en lisant le texte -même commenté- de cette même pièce sur un terminal ASCII. Par contre, les documents en tant que tels ont très rarement un rôle pour la machine. Celle-ci se limite à les restituer. Toute interaction entre le lecteur et un hypertexte s'appuie sur les documents. L'utilisateur en consulte un puis un autre, il navigue. La navigation est certes à l'initiative de l'utilisateur mais elle ne peut se faire qu'en fonction des propositions de destinations que lui offre l'hypertexte.

- La connaissance² (au sens informatique) est un ensemble de relations spécifiquement destinées à la machine qui ont pour rôle de lui permettre de décider ce qui doit être proposé au lecteur comme suites possibles de sa lecture en fonction du contexte courant. La représentation la plus pauvre de cette connaissance est le lien "en dur"³ qui relie deux documents. La forme la plus élaborée peut être gérée par un système complexe s'appuyant sur la modélisation du domaine, celle du lecteur, tenant compte de l'historique des lectures, de ses goûts observés, etc.

¹ Souvent, le lecteur croit être le seul responsable objectif de l'interprétation des documents. L'un des problèmes éthiques que posent les hyperdocuments est la partielle transparence de l'auteur pour le lecteur. En pouvant accéder, dans les documents, à des faits bruts (ex: dans des documents historiques d'archives), et non à l'exposé d'une interprétation des faits par l'auteur, le lecteur croit disposer pleinement de sa liberté d'interprétation. En fait l'auteur de l'hypertexte en renforçant les relations entre certains documents influe sur l'ordre de lecture et donc d'interprétation des documents. En rapprochant des informations, il facilite leur lecture conjointe et donc oriente fortement l'interprétation des faits, la précedence temporelle en lecture étant implicitement perçue comme un relation de causalité.

²Le terme connaissance n'est pas utilisé ici avec son sens commun. Nous introduisons le terme (informatique) de connaissance pour insister sur l'aspect formel et calculable de la connaissance qui est mise en jeu du côté de la machine. Nous l'opposons à celui d'information (véhiculée dans les documents) qui est destinée à être interprétée par l'homme. Un extrait d'un discours est une information, le rapprochement pertinent entre cet extrait et un document relatant un autre fait est la matérialisation d'une connaissance.

³défini en extension -par exemple une référence explicite-, par opposition à un lien dynamique, calculé par la machine lors de sa traversée. Le lien dynamique est défini en intension -par exemple une propriété que doit vérifier le document cible-.

- L'ancrage est le liant entre ces deux composants. L'ancrage est une fonction très importante dans les hypertextes -et hélas souvent à la fois mal comprise et mal implémentée-. L'ancrage est le moyen d'associer librement la connaissance aux documents. Son sens usuel est l'attachement d'information à l'extrémité d'un lien -pour nous, la mise en relation de deux informations portées par les documents-.

La métaphore de *l'ancrage* est particulièrement bien choisie. Elle suggère l'indépendance et la flexibilité des relations dans un hypertexte. Le réseau de connaissances "flotte" au dessus des documents ; il n'en fait pas partie! L'auteur de l'hypertexte peut jeter une ancre là où il veut. Le mouillage d'une ancre est une opération naturelle pour un bateau : elle ne nécessite pas que l'on construise un port là où l'on veut mouiller (-en termes informatiques, il n'est pas nécessaire de modifier les documents pour réaliser l'ancrage, c'est le système hypertexte qui mémorise l'ancrage et non les documents).

Dans un système hypertexte (bien fait) le réseau de connaissance n'altère pas les documents, et est le plus souvent indépendant de ceux-ci. Ainsi dans Intermedia [Yankelovitch 88], un même ensemble de documents est partagé par plusieurs hypertextes indépendants, chacun matérialisant une connaissance différente des informations des documents. Chaque étudiant peut d'ailleurs continuer à construire d'autres hypertextes sur ces documents. Chaque relation exprimée entre documents est juste "ancrée".

Une bonne métaphore pour comprendre la nature réelle des hypertextes est d'imaginer un ensemble de documents accrochés à un mur et protégés par une feuille de plastique transparent. La connaissance ancree sur ces documents serait matérialisée par des annotations, des lignes colorées reliant des parties de documents tracées au marqueur effaçable sur le plastique. Il est alors possible de faire évoluer cette connaissance sans altérer les documents. Le matériau composite constitué des documents, du plastique qui assure l'ancrage (le positionnement) et des dessins faits sur le plastique a typiquement la structure caractéristique des hypertextes. Cette métaphore n'est pas gratuite : certains systèmes tels Dolphin [Haake 94] l'implémentent au moyen de tableaux actifs (Liveboards), où l'utilisateur voit les documents projetés sur le tableau et exprime les relations entre informations par ses gestes sur l'image.

Un autre métaphore est parfois employée à la place de l'ancrage : "harponner de l'information". La suggestion de blessure qui est attaché à ce terme traduit moins l'indépendance qui existe entre les documents et la connaissance que l'on plaque dessus.

2.3. Intérêt de cette unification

Nous préférons de très loin l'interprétation sémantique des mécanismes hypertextes basée sur la connaissance associée aux documents -beaucoup plus générale- à celle de la traditionnelle description basée sur la technique d'une organisation en nœuds et liens. Elle permet d'unifier des systèmes aussi différents que Superbook [Remde 87] -où les liens sont absents!- dans lequel la navigation est basée sur des mécanismes de recherche d'information, MacWeb⁴ [Nanard 91], où les relations entre documents peuvent faire appel à de l'inférence, et des systèmes plus simples tels WWW où les relations entre documents sont des références mono-directionnelles explicites. D'autres systèmes utilisent explicitement des mécanismes d'apprentissage [Boy 91], des systèmes experts

⁴Il s'agit du système MacWeb développé sous ce nom au LIRMM depuis 1988, et non du visualiseur WWW développé au MCC, et qui a repris ce même nom en 1994.

[Deschamps 94] ou des réseaux neuronaux [Lelu 92] pour gérer la navigation. Tous répondent à ce modèle. Même sous sa forme la plus fruste (documents et liens en dur), les liens matérialisent la connaissance que l'auteur a des documents. Le travail de l'auteur -ou de l'enseignant- lors de la construction d'un hypertexte est typiquement une explicitation de connaissances. Toutefois, contrairement à la rédaction d'un discours, qui serait orienté directement vers l'homme, cette explicitation de connaissances se fait sous une forme purement orientée vers la machine. Elle ne transparait pour le lecteur qu'au travers du guide ténu qui oriente discrètement (ou parfois contraint) l'enchaînement de ses lectures.

Cette unification permet aussi de distinguer ce qui est un hypertexte de ce qui n'en est pas un. Un hypertexte doit fondamentalement être composé de ces deux entités - connaissance et documents-. Si l'une d'elles manque, il ne s'agit pas d'un hypertexte. Un système candidat au test de Turing [Shieber 94], même s'il était capable de 'discuter intelligemment' de peinture avec un humain ne serait pas pour autant un hypertexte, si l'interaction se fait sans s'appuyer sur les tableaux sur lesquels porte la discussion. Même si l'aspect gestion de connaissance est excellent dans un tel système, il enfreint la règle du support de l'échange qui -dans un hypertexte- ne peut être que le document. Inversement un manuel de maintenance qui contient des animations vidéo à la place des images pour illustrer les gestes à effectuer n'est pas non plus un hyperdocument, car ces animations font partie intégrante du document lui-même. Il n'y a donc pas intervention de connaissances externes au document pour proposer des choix de lecture. Enfin, un système d'EIAO dans lequel un ensemble de connaissances sur le domaine étudié, sur l'apprenant, sur son niveau, sur les stratégies pédagogiques, ... etc. qui serait utilisé pour piloter la progression des documents proposés à l'apprenant serait bien un hypertexte, en raison de la dualité documents-connaissances sur les documents.

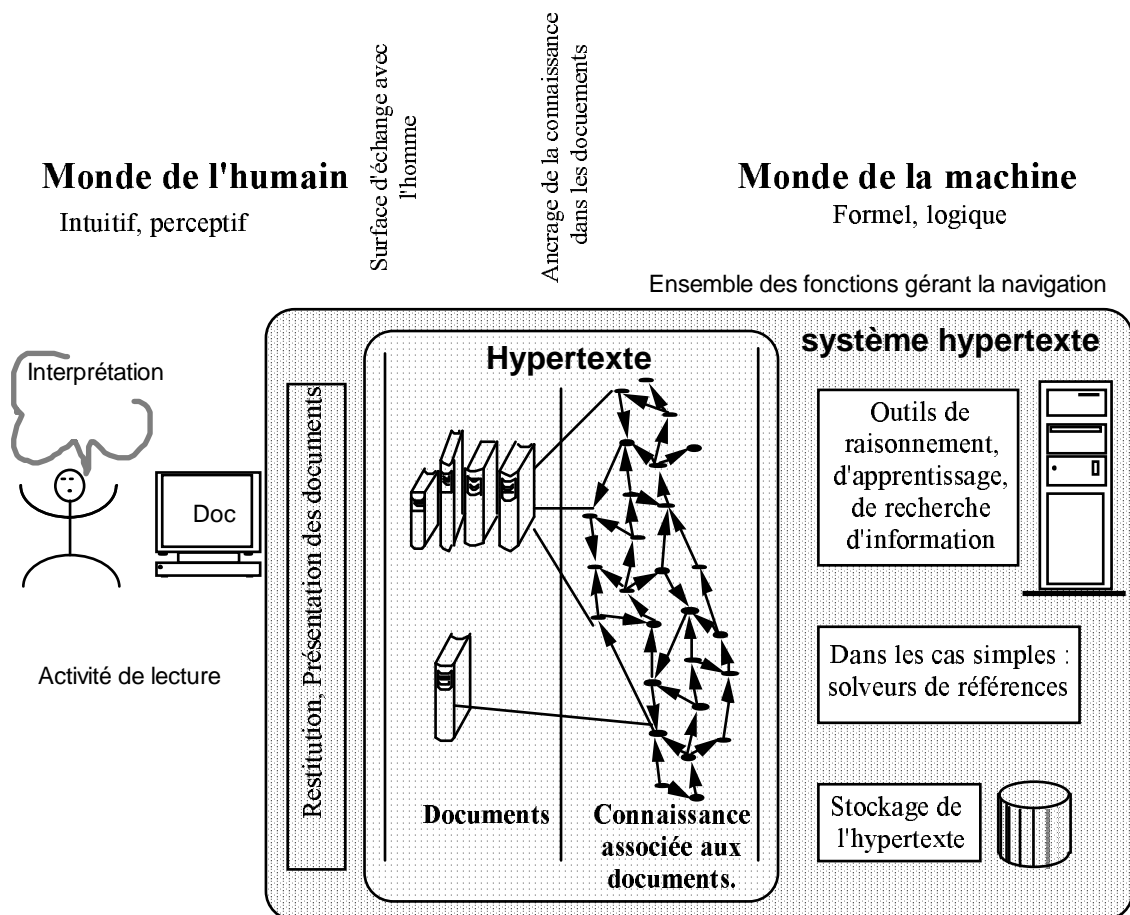


Fig.1 : *Modèle fonctionnel de ce qu'est un hypertexte.* La traditionnelle notion de lien est en fait l'établissement d'une relation entre deux informations au travers de la connaissance associée aux documents.

C'est essentiellement le couplage entre les deux composants documents-connaissances qui caractérise de façon profonde la notion d'hypertexte. L'hypertexte est typiquement une interface entre l'homme et la machine, une membrane au sens chimique avec ses deux faces spécifiquement adaptées aux milieux qu'elle sépare : d'un côté les documents pour le monde perceptif et émotionnel de l'homme -seul capable d'apprécier un poème, une peinture-, de l'autre la connaissance formelle et logique qui permet à la machine de guider efficacement le processus de lecture non linéaire, par exemple en déterminer en une fraction de seconde une similitude entre deux passages de textes parmi des milliers.

3. L'HYPERTEXTE FACE AUX SCIENCES ET TECHNIQUES EDUCATIVES

3.1. Documents et connaissances

La modélisation des hypertextes en tant que matériau composite -documents / connaissance- est particulièrement pertinente pour comprendre le rôle et l'intérêt des hypertextes dans les techniques éducatives, tant du côté de l'enseignant que de celui de l'enseigné.

- L'auteur, -l'enseignant-, en construisant un hypertexte, organise un réseau de relations entre des informations ; il les valorise en leur greffant des connaissances additionnelles. Il crée aussi des documents pour constituer le contenu de l'hypertexte. Ces documents ont

une valeur informationnelle intrinsèque indépendamment de leur appartenance à l'hypertexte. Par contre, la richesse de l'hypertexte, qui est responsable de l'aspect dynamique de l'exploration dépend essentiellement de la structure de connaissance qui est ajoutée à ces documents et permet de les rapprocher.

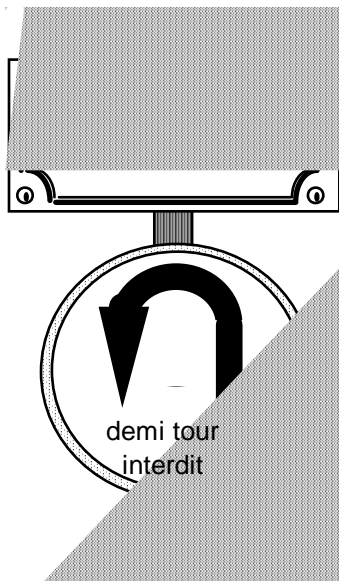
Remarquons que contrairement aux livres où la connaissance apportée par l'auteur est intimement mêlée à l'information -et de ce fait indissociable d'elle-, la connaissance peut rester disjointe de l'information dans un hypertexte. Ceci est mis à profit dans des outils comme Intermedia où enseignants et étudiants opèrent sur une même base d'information, chacun exprimant sa propre connaissance en rapprochant des informations, mais sans les altérer.

- Le lecteur, au cours de son activité de lecture de l'hypertexte, met en jeu implicitement deux mécanismes : la lecture des documents (le plus souvent immuables), et l'évolution dans un cadre de lecture souple -la structure hypertexte- qui engendre un mécanisme de montage⁵ des informations [Balpe 90]. Un même document observé au cours de deux cheminements distincts aura le plus souvent des interprétations différentes. Ceci est mis en pratique dans les hypertextes dits de "fiction" dans lesquels le lecteur choisit à sa guise l'évolution d'une histoire. Les mêmes passages peuvent être utilisés avec des sens différents dans des versions différentes car leur interprétation par l'homme est toujours contextuelle.

Le choc de la juxtaposition d'informations a souvent un impact plus important sur le lecteur que le contenu du document lui-même. Cette technique qui est bien maîtrisée sous sa forme séquentielle par les producteurs d'émission de télévision, s'avère beaucoup plus riche, plus complexe (et plus sournoise) dans le cadre non linéaire des hypertextes. Cette technique de juxtaposition a un intérêt pédagogique extrêmement important, par exemple pour la formation contextuelle en situation de travail. Ainsi, dans des postes de travail/formation mixtes, la détection par le système d'une difficulté de l'utilisateur déclenche l'affichage de documents de formation spécifiquement adaptés au contexte. La juxtaposition temporelle de l'occurrence de la difficulté rencontrée et de l'apparition du document pédagogique associé permet de mieux le mémoriser : l'information délivrée répond à un besoin. Cet exemple est typique d'une approche hypertexte riche puisque le système propose un document (ensemble d'information préexistante) en s'appuyant sur l'analyse d'une situation (connaissance gérée par le système).

Typiquement ces deux aspects -contenu et parcours- sont les reflets de la dualité de la structure des hypertextes.

⁵ au sens cinématographique. Chaque document est lu et interprété dans un contexte, c'est une scène qui ne prend sens qu'en fonction de ce qui a déjà été vu. L'ordre de lecture des documents engendre un montage dynamique [J Nanard 94], qui ordonne les documents comme les scènes d'un film. Ce procédé est d'ailleurs mis à profit dans les hypertextes de fiction, où différentes histoires s'engendrent par des parcours différents d'un même ensemble de documents. Bernstein fait remarquer que la présence de cycles n'est pas gênante, le même texte relu dans un contexte différent ayant souvent une interprétation très différente.



La dépendance induite entre ces deux panneaux du fait de leur montage sur le même support entraîne l'émergence d'un sens nouveau, suggère une histoire, qui n'est pourtant portée en tant que telle par aucune des structures informationnelles présentes. Ce mécanisme est du point de vue cognitif la base de l'approche hypertexte. Les relations entre informations explicitées dans l'hypertexte sont les germes d'informations nouvelles. Mais elles ne prennent naissance que lors de l'interprétation qu'en fait le lecteur en fonction du montage réalisé lors du parcours effectif de l'hypertexte. C'est ce qui fait la force mais aussi la difficulté de conception d'hypertextes pour son auteur.

Fig. 2 : La notion de montage.

- Le lecteur et l'auteur sont parfois la même personne. C'est en particulier le cas lorsqu'un apprenant enrichit un hypertexte de ses propres annotations, et surtout met en évidence des relations nouvelles entre des informations déjà présentes dans l'hypertexte. En ancrant de nouvelles relations entre des documents, il matérialise sa connaissance du domaine ainsi que la synthèse qu'il a réalisé des informations qui lui ont été présentées. Ainsi, un étudiant qui utilise Intermédia, matérialise sa compréhension d'un cours d'histoire de l'art en établissant des liens pertinents entre un ensemble d'œuvres. Cette connaissance ainsi exprimée dépasse le simple contenu informationnel des documents qui la supportent.

3.2. Les types d'usage des hypertextes

Bien que les hypertextes répondent tous au modèle fonctionnel que nous venons de présenter, à savoir la mise en relation d'informations par des connaissances- ils diffèrent largement en fonction de l'usage qui est fait de cette connaissance et de la finalité de leur mise en œuvre par le lecteur/auteur. Ces différences sont corrélées au type d'application qui est fait de l'hypertexte. D'ailleurs le style de l'interface de l'hypertexte est lui-même fortement dépendant de l'usage qui en est attendu.

Il y a en gros trois types d'attitudes face à l'information :

- L'extraction d'information brute dans des bases d'information, apparentée à la recherche documentaire.
- L'organisation d'information existante pour mieux la présenter, faciliter son accès, augmenter son impact chez le lecteur, en un mot, la valoriser.
- La production d'information ou des structures de connaissances nouvelles.

Ces attitudes face à l'information sont homologues à celles que l'homme a eu face à la matière (inerte ou vivante) : la mine, la transformation, l'agriculture (jardinage). Les termes anglo-saxons de *Information Mining*, *Information Manufacturing* et *Information Gardening* introduits par Mark Bernstein [Bernstein 93] dans une métaphore intéressante sont plus évocateurs. Ces attitudes conduisent à des types de systèmes hypertextes très différents dans leur forme et leur fonctionnalité dans la mesure où le besoin à satisfaire chez l'utilisateur est différent. Toutefois, lors de la conception de

document hypermédias complexes, on constate que l'auteur évolue constamment entre ces trois activités, recherchant de la matière première, l'organisant et la valorisant.

3.2.1. L'extraction d'information

L'hypertexte est un remarquable outil d'extraction d'information. Par son style d'interaction basé sur les documents, il est plus naturel pour le novice que les systèmes à base de requêtes. Notons d'ailleurs que le modèle proposé ne suppose rien sur l'architecture interne du système. Celui-ci peut très bien mettre en jeu un mécanisme interne à base de connaissance pour calculer des similitudes de documents et proposer un ensemble de documents voisins de celui ou de ceux sélectionnés. Plus encore, la présence de relations explicites entre informations permet d'obtenir des réponses plus pertinentes [Ben Mrad 94].

Toutefois, dans ce type d'approche, le lecteur est avant tout concerné par l'aspect recherche d'information. Il est plus intéressé par la rapidité d'accès à des documents pertinents et par l'abondance des informations obtenues que par les autres aspects, tels leur valorisation immédiate, la présentation, ou même la qualité des relations qu'un auteur aura ancré dans ces documents. L'utilisateur est alors d'abord un fouineur plus qu'un lecteur. L'information pour lui est un minerai qu'il exploitera lui-même. En conséquence dans ce type d'approche, la structure de connaissance est souvent gérée de façon automatique, ce qui est une quasi nécessité compte tenu des volumes importants d'information mis en jeu. De plus, le type d'organisation de cette connaissance doit se refléter dans la surface de l'hypertexte par la présence de nombreux documents de style table des matières facilitant une navigation en profondeur.

L'hypertexte destiné à la recherche d'information est un outil pédagogique extrêmement flexible. En raison même de l'interaction directe avec les documents, il peut être utilisé de deux façons :

- La première est l'exploration libre "vagabondage", dans laquelle l'utilisateur feuillette l'hypertexte comme on feuillette sans but initial une encyclopédie. Ceci est très caractéristique des premières prises de contact des utilisateurs avec ces types de systèmes, par exemple WWW.
- La seconde est la recherche intentionnelle d'information. C'est le type d'emploi finalisé des hypertextes de cette catégorie. Elle a un rôle formateur extrêmement important, ceci pour deux raisons. D'une part, l'homme aura de plus en plus à acquérir d'information par lui-même dans les décennies qui arrivent. Il est donc très important de le former à ce type d'activité. D'autre part, en l'obligeant à se concentrer sur un but précis et à résister aux tentations de digressions qui sont nombreuses lors d'une navigation, ce type d'exercice à un rôle formateur extrêmement important, même chez les plus jeunes.

3.2.2. La valorisation de l'information

L'information sera au siècle qui arrive ce qu'a été l'énergie à celui-ci. Depuis des siècles l'homme a amassé beaucoup d'information mais l'a peu valorisée. Par valorisation nous entendons une transformation destinée à la rendre mieux et plus rapidement assimilable par l'homme. En dissociant l'aspect structurel du contenu, en permettant grâce au multimédia la consultation de documents au pouvoir évocateur fort, en permettant grâce aux réseaux l'accès à de l'information répartie, les hypertextes sont le vecteur d'avenir de la diffusion d'information valorisée. Les grands éditeurs ont déjà perçu l'intérêt de ce marché, et les nombreux CD-ROMs disponibles n'en sont que le signe avant coureur.

Le rôle d'une équipe pédagogique qui développe un hypertexte afin de valoriser de l'information est typiquement d'extraire des éléments informationnels, de les interconnecter pour permettre au lecteur de satisfaire sa curiosité instantanée. Ainsi, un hypertexte sur "la flûte enchantée" permettra bien sûr d'écouter l'œuvre, de la voir sur l'écran, mais une connaissance supplémentaire constituera le propre de l'hypertexte. Ancrée (au sens défini précédemment) dans l'œuvre, elle ne la perturbe en rien, mais permet à tout instant au lecteur d'accéder au livret s'il ne comprend pas un passage, de se renseigner sur un personnage, sur le contexte, de voir la partition, d'observer le visage d'un acteur...etc. En établissant des relations entre des informations existantes, l'auteur valorise l'information. Les documents proposés au lecteur conservent le plus souvent le même contenu informationnel mais leurs interrelations matérialisée par les connaissances de l'hypertexte, facilitent leur interprétation par le lecteur.

Pour l'auteur, ce type d'activité, par son caractère créatif, dans laquelle le nouvel hypertexte peut être vu comme une œuvre à part entière, est fondamentalement différent de celui de la simple introduction de nouvelles informations dans un hypertexte destiné à la recherche documentaire. Comme dans les livres, la qualité esthétique est importante pour assurer une bonne transmission du message au lecteur. Toutefois, beaucoup d'outils privilégiant l'aspect recherche d'information la négligent encore ou la relèguent à un rôle très secondaire.

En tant qu'outil pédagogique, les hypertextes basés sur la valorisation de l'information sont surtout utilisés en phase de lecture, ce qui est leur finalité. Mais, il semble important d'insister sur le rôle formateur important que constitue pour un apprenant la *construction* d'hypertexte. Cette activité créatrice nécessite chez l'apprenti auteur d'abord de maîtriser la base informationnelle sur laquelle il construit l'hypertexte, mais surtout développe chez lui le sens de l'observation, de l'évaluation, et des facultés de jugement pour décider quels sont les points à valoriser dans son document. Enfin une qualité très importante est promue dans ce type d'activité : l'évaluation de la pertinence des structures et de la présentation des informations. Etre auteur, c'est d'abord être capable de se mettre à la place du futur lecteur.

Le domaine de la valorisation d'informations étant celui le mieux connu et le plus développé des hypertextes, nous ne le décrivons pas davantage.

3.2.3. Engendrer de l'information ou des structures de connaissances

Le troisième type d'activité de l'utilisateur se servant d'un hypertexte est de faire émerger de nouvelles informations, de nouvelles connaissances. Les systèmes hypertextes destinés à ce type d'activité sont souvent appelés *gestionnaires d'idées*. Mais encore une fois, ce terme est trop imprécis (et a été galvaudé par la commercialisation de produits médiocres) pour être repris ici. Nous emploierons donc encore celui d'hypertexte.

3.2.3.1. La métaphore du jardinage

La métaphore du jardinage introduite par Cathy Marshall, puis développée par Bernstein est très riche et subtile [Bernstein 93]. Elle est extrêmement importante pour saisir l'esprit de ce type d'activité. Détaillons la.

Le jardinier d'information fait pousser de nouvelles informations et de nouvelles structures à partir d'informations existantes. Il est le vrai producteur d'information.

L'observation puis l'explicitation de relations entre les informations conduit à l'émergence de structures qui s'ancrent sur les documents. Ces relations "poussent" en quelque sorte sur les documents et représentent la connaissance de l'hypertexte qui est en train d'être construit. Notons au passage la dualité de cette connaissance qui est en même temps un aspect du savoir personnel de l'auteur et la connaissance explicitée disponible dans l'hypertexte. Mais il faut pousser la métaphore plus loin. Telles les plantes, les relations entre informations portent aussi des fruits. En connectant des éléments informationnels, ces relations les rapprochent et donnent à leur couple un sens nouveau. Il est bien connu que la juxtaposition d'informations est susceptible de générer un sens qui n'est explicite dans aucune d'elles (voir fig 2). Cette démarche très classique de tout auteur qui se documente dans une bibliothèque, rapproche des informations, en déduit de nouvelles, etc., est désormais plus qu'une démarche intellectuelle. Elle est une suite d'actions, prises en charge par un système qui les aide, en garde trace, et surtout les conserve directement sous forme d'une structure hypertexte re-exploitable.

Le jardinage d'information vise à rendre explicites des connaissances (au sens usuel) implicitement portées mais non apparentes dans un ensemble de documents. Le temps (au sens durée) et l'évolution (sous son aspect instabilité) sont donc fondamentalement des caractéristiques majeures de cette activité qui vise la compréhension profonde d'informations et l'explicitation de connaissances. Contrairement aux approches relatives à la valorisation d'informations dans lequel l'objet principal est le produit fini, ici l'objet principal est le processus mental qui conduit le jardinier d'information à constater l'émergence d'une nouvelle idée. Le jardinier vit pleinement son travail, la floraison de nouvelles informations en sont la conséquence, alors que pour le producteur d'informations valorisées, tendu vers le résultat visé, le travail n'est que le moyen de l'atteindre. Par nature l'un peut savoir où il va, l'autre ne le peut pas. C'est en cela que le jardinage d'informations appartient aux activités créatrices.

Un second parallélisme est important à observer. De la même façon que les plantes ont besoin de soins et de temps avant d'arriver à maturité, une idée réellement neuve n'émerge jamais sous sa forme la plus élaborée. Pour permettre de prendre en compte le processus mental réel qui conduit de la première apparition d'une idée à sa mise en forme, il est fondamental que le système hypertexte qui est le support des activités de jardinage d'informations permette (et prenne en charge) l'évolution des connaissances et autorise le passage par d'inévitables états transitoires incohérents et instables qui font partie inhérente de toute démarche de pensée. Les hypertextes destinés au jardinage doivent donc disposer de fonctionnalité spécifique pour gérer et manipuler facilement les états transitoires qui sont le plus souvent ignorés ou même méprisés dans la plupart des approches. En ce sens, les outils de jardinage de l'information sont en prise directe avec les processus mentaux des créateurs. Pour être efficaces ils doivent faciliter les démarches intuitives et l'émergence progressive de structure.

3.2.3.2. Les outils pour le jardinage d'information

La disposition d'outils respectant vraiment les processus mentaux propres aux activités créatrices est extrêmement importante pour la formation. Ces outils doivent faciliter le processus créatif lui-même et non se limiter à permettre la description du résultat de ce processus. Toutes les personnes impliquées dans une activité créatrice, -même les chercheurs-, reconnaissent que leur démarche interne est une alternance de phases ascendantes et descendantes, même si l'énoncé final est plus facile sous une forme descendante. De nombreuses études expérimentales [Hayes 80??], en particulier sur le

processus d'écriture ont mis en évidence l'aspect opportuniste et l'intrication des types d'activités au sein d'un processus créatif. La qualité d'une œuvre (dans un contexte créatif) est fortement corrélée à la complexité et à l'instabilité de son processus d'élaboration.

Malheureusement, très peu d'outils prennent réellement en charge les aspects transitoires de la pensée. Beaucoup de systèmes ne s'intéressent qu'à l'état final, jetant un voile pudique sur ce qui a permis d'y arriver. Les quelques systèmes hypertextes typiquement orientés vers le jardinage d'information Aquanet, Viki, Sepia, Dolphin, Macweb⁶ sont des outils dont l'emploi facilite grandement toute l'activité de construction incrémentale d'hypertexte. Ces outils, en permettant une édition interactive naturelle de structures, en supportant l'évolution et certaines incohérences de structures permettent de coller au plus près d'un processus de création. Les aspects techniques de ces outils sont décrits dans [Nanard 94a] tandis que l'aide au processus créatif est présenté dans [Nanard 94b].

Naturellement un outil doit être spécifique à la tâche à laquelle il est destiné. Un système hypertexte n'aura pas les mêmes fonctionnalités selon qu'il est utilisé pour de la recherche d'informations, la valorisation de celles-ci, ou l'émergence de structure. Ainsi, lorsque l'on construit un hypertexte destiné à la valorisation d'informations, on peut se permettre de spécifier le produit final, d'en décrire la structure, le contenu, ... etc. sous une forme textuelle. Un compilateur de description d'hypertexte produit l'hypertexte final [Caloini 92]. Et on l'observe alors. Mais on ne peut le modifier directement. Il faut intervenir sur sa spécification écrite et recommencer tout le cycle. Le système hypertexte ne supporte⁷ pas l'effort de l'auteur. Celui-ci opère en retrait du système.

Au contraire dans un environnement destiné au jardinage, l'auteur agit directement sur l'hypertexte lui-même, il l'observe, l'évalue et incrémentalement le modifie. Il y a continuité dans son action, la rétroaction entre l'observation et l'édition peut être directe. Plus encore, il peut incrémentalement bénéficier des éléments structurels qu'il a élaborés au cours des étapes précédentes pour faciliter sa tâche. Cette liberté de travail n'est pas pour autant une cause de désordre et de structuration anarchique. Elle permet la revalorisation des démarches ascendantes qui avaient été dénigrées au profit des démarches purement descendantes.

3.2.3.3 . *Le jardinage d'information et la formation*

La démarche de jardinage d'information est pratiquée couramment, mais le plus souvent sans outils, par la plupart des enseignants qui préparent un cours. Au mieux, utilisent-ils des outils très primitifs qui tiennent plus du bloc note électronique que de l'hypertexte, où à l'inverse, des outils lourds trop proches d'environnements de programmation. Les systèmes hypertextes offrant un mode auteur réellement convivial sont rares mais apportent une aide réelle à l'organisation des informations, et permettent surtout de passer de façon continue de la phase d'émergence des idées au document pédagogique où les informations sont valorisées.

Mais, c'est surtout du côté de l'apprenant que l'apport des outils de jardinage d'informations semble le plus profitable. Un pas important vers le développement de

⁶le nom MacWeb est ici celui du prototype de recherche développé au LIRMM. Depuis le milieu 94 ce nom a été repris à des fins commerciales par un autre produit développé au MCC destiné à l'accès à WWW, qui n'a rien à voir avec le prototype de recherche cité ici.

⁷au sens sportif.

l'initiative et le vers le retour de la spontanéité consisterait à faire prendre conscience à un apprenti que l'élaboration d'une œuvre, -qu'elle soit littéraire, culturelle, scientifique ou autre- n'est ni un miracle ni l'application d'une méthode efficace et stérilisante, qu'elle a ses grandeurs et ses faiblesses qui en font légitimement partie. Lui offrir un outil qui lui permette une démarche d'essai erreur, et qui, plus encore, l'aide dans cette démarche, lui permette de s'appuyer réellement sur les résultats des étapes intermédiaires pour avancer plus vite dans sa création, est à mon sens, une activité extrêmement salutaire dans la formation. En faisant prendre conscience de la réalité du processus mental de sa démarche créative à l'apprenti, on libère sa spontanéité, sa créativité profonde, on lui apprend à penser, à créer librement, et non à savoir seulement mettre en œuvre des méthodes stéréotypées.

4. QUELQUES EXEMPLES D'APPLICATIONS HYPERTEXTES

Dans cette section, nous donnons quelques exemples d'applications hypertextes pour illustrer les trois facettes de l'usage des hypertextes : l'extraction, la valorisation et le jardinage. Les travaux sur l'emploi d'hypertextes dans l'enseignement, en particulier à l'IRIS⁸, sont suffisamment connus pour qu'il ne soit plus nécessaire de les décrire. C'est pour cela que nous présentons des applications hypertextes moins classiques.

4.1. World Wide Web

World Wide Web est un système hypertexte réparti, de couverture mondiale, accessible librement au travers de l'INTERNET⁹. Il a été créé par le CERN dans le but de mettre en commun les résultats scientifiques, les publications et la documentation dont avaient besoin les chercheurs. Basé sur une approche ouverte et collective, très rapidement ce système a séduit la communauté scientifique, ...et même au delà¹⁰. Actuellement WWW est de très loin, tant en volume et qu'en nombre d'interactions, le plus grand hypertexte au monde. Son taux de croissance en 94 a été spectaculaire, le trafic dans Internet concernant WWW doublant sensiblement tous les 60 jours.

Le principe de fonctionnement, conforme au modèle général des hypertextes, est très simple. Matériellement, WWW est organisé selon une architecture client-serveur. Chaque serveur de documents ne gère et ne connaît que l'ensemble des documents dont il est responsable. Il reçoit des requêtes de clients lui demandant tel document qu'il délivre alors selon un codage standard. WWW repose sur deux standards: HTTP Hypertext Transfert Protocol qui concerne l'aspect réseau et est utilisé pour le dialogue entre les clients et les serveurs et HTML Hypertext Markup Language qui est utilisé pour décrire les documents. Les documents sont principalement textuels, mais peuvent contenir des images, des animations vidéo, des sons dans la plupart des formats usuels.

Dans cette approche, la structure de connaissance est très primitive. Elle est le plus souvent matérialisée par des liens unidirectionnels statiques qui sont conservés avec le

⁸Institute for Research on Information and Scholarship, Brown University. IRIS est à l'origine de INTERMEDIA [Yankelovitch88] et de nombreuses expériences en vraie grandeur d'utilisation des hypertextes dans l'enseignement.

⁹Internet est le réseau fédérateur des divers réseaux principaux universitaires.

¹⁰On trouve même déjà des menus de restaurants ainsi que des catalogues de vente par correspondance sur WWW. Des serveurs spécialisés louent maintenant de l'espace d'information à des entreprises commerciales.

document source. Ils sont notés en HTML au sein de la description du document. Le client qui a reçu un document, a reçu aussi les liens qui partent de ce document. Un interpréteur HTML sur la station client assure la mise en page décrite en HTML et gère l'ancrage dans le texte et les images. Les logiciels clients les plus répandus et les plus conviviaux sont actuellement Mosaic, développé au NCSA¹¹ et maintenant Netscape. Chaque lien repère une ancre, un document destination et identifie le serveur de ce document sur le réseau sous forme d'une URL (Unified Ressource Locator). Le choix effectué par le lecteur en sélectionnant une ancre est matérialisé par l'envoi d'une requête au serveur qui gère le document cible.

Dans ce type d'architecture, l'ensemble de la connaissance mise en jeu lors d'une navigation est typiquement une connaissance collective répartie. Chaque auteur ne gère qu'un ensemble limité de documents dans lesquels il est le seul autorisé à définir des références vers d'autres documents. Par contre n'importe quel autre auteur peut, dans les documents dont il est responsable, référencer les documents de tout autre auteur sans que ce dernier ait besoin d'en être informé. La structure de liens est asymétrique. Elle permet d'interconnecter librement des cellules de connaissances. Ainsi chaque spécialiste matérialise sous forme d'un hypertexte la connaissance qu'il a de son domaine, ainsi que les relations entre ces connaissances et d'autres documents appartenant à d'autres spécialistes. Il devient alors possible de naviguer à partir d'un document, en profitant à tour de rôle de la connaissance des divers auteurs, jusqu'à atteindre la cible recherchée. Il me semble très important de se dégager de l'aspect technique (liens, requêtes...) pour observer macroscopiquement et sociologiquement ce phénomène et constater que ce qui fait réellement sa force vient de l'interconnexion de savoirs répartis.

Le fonctionnement social de WWW repose sur le partage non lucratif de la connaissance. Chacun des auteurs met bénévolement ses connaissances à la disposition de la collectivité sachant qu'en retour il accède gratuitement, en s'appuyant sur le savoir des uns et des autres à des informations qui n'étaient pas en sa possession, et peut les utiliser librement. Il peut indifféremment les référencer (mais dans ce cas il ne peut ni les modifier ni leur ajouter de liens), ou les copier dans son propre domaine où il peut les compléter librement, et en particulier ajouter d'autres références et les mettre ainsi à la disposition de la collectivité. Toutefois, toute copie détache l'objet de son auteur initial; par suite l'évolution de la connaissance reportée sur l'original -tant en contenu qu'en liens- ne sera pas répercutée.

Après un peu plus d'un an d'expansion, WWW est actuellement une mine d'information extrêmement diversifiée, très riche. La facilité d'emploi liée à l'approche hypertexte en particulier grâce à Mosaic en fait un outil de choix pour l'enseignant qui veut explorer un domaine, rassembler de l'information pour préparer un enseignement. Le travail d'acquisition d'information fait traditionnellement à partir de documents en papier, longs et chers à obtenir, se fait alors très rapidement en accédant en quelques secondes à des documents situés dans des universités distantes de plusieurs milliers de kilomètres. Plus encore, dans des matières où des archives historiques, des images de sites, d'objets, etc. ont une importance pédagogiques majeure, l'accès à l'information se trouve grandement facilité. On peut ainsi très facilement accéder à des enregistrements vidéo d'événements historiques, à des images de pièces de musée, etc. De nombreux cours sont déjà disponibles sous WWW. Les documents pédagogiques développés sous WWW offrent à

¹¹National Center for Supercomputing Applications, le centre informatique de l'université de l'Illinois à Urbana Champaign.

l'apprenant un espace de connaissances explorable de façon continue. La trame principale d'un nouveau document pédagogique, en référençant d'autres documents ouvre largement et librement l'espace d'exploration, contribuant à développer l'esprit d'initiative chez l'apprenant.

Naturellement, comme dans toute consultation de document hypertexte, les risques de digression ne sont pas négligeables. Toutefois, les outils de consultation offrent des fonctions de reprise directe du parcours en un point quelconque, qui à défaut d'éviter au lecteur de s'égarer, lui permettent de reprendre plus facilement le fil de son exploration.

World Wide Web, au delà d'une application hypertexte réussie est d'abord un phénomène de société. Il est appelé à avoir un impact considérable dans les années à venir, tant sur la façon d'accéder à l'information, que sur celle de l'organiser pour diffuser le savoir. WWW se comporte macroscopiquement comme un système collaboratif socialement organisé : des comportements émergent et sont acceptés de façon consensuelle sans qu'aucune autorité n'en ait la responsabilité. Chacun agit à sa guise, mais souvent en s'inspirant de ce qu'ont fait d'autres. De ce fait, lorsqu'un choix est macroscopiquement bon, il est progressivement adopté par la communauté sans qu'aucune règle explicite ne le prescrive. En ce sens, WWW est un remarquable terrain expérimental pour l'étude de l'émergence de comportement collectifs. Notons aussi que la structure répartie de WWW conduit à une dépersonnalisation¹² de l'information apparaît pour le lecteur. Le lecteur passe sans transition notable des informations fournies par un auteur, à celles fournies par un autre. Le lecteur ne pense plus à un auteur -comme dans le cas d'un livre-, mais à l'information obtenue de façon plus anonyme.

Dès que la baisse du coût des télécommunications et l'augmentation du débit des réseaux rendra cette technique réellement accessible à une part importante de la population, son impact sera comparable à celui de la télévision. La différence fondamentale étant qu'alors le lecteur aura l'initiative...

4.2. L'édition interactive d'hypertexte

A l'opposé de World Wide Web dans lequel l'aspect social prime sur l'aspect individuel, certains systèmes hypertextes ont pour vocation d'aider un individu -enseignant ou enseigné- à organiser personnellement ses connaissances. On trouvera dans [Marshall 91], [Marshall 94], [Streitz 89], [Streitz 92], [Haake 94] des documents très intéressants sur des hypertextes orientés vers l'édition dynamique de structures complexes. Nous présentons ici, de façon plus technique, comment cet aspect est géré dans le système MacWeb. L'objectif de cette partie de la présentation est d'illustrer concrètement quels sont les outils disponibles pour les activités de jardinage d'information, et par là même montrer que ce type d'approche, malgré la liberté qu'elle prône, s'appuie sur des bases rigoureuses. Cette partie de l'exposé résume les idées énoncées dans [Nanard 94]. On trouvera aussi dans [Nanard 94-2] une réflexion plus générale sur les méthodes de conception d'hypertextes et leurs implications sur les fonctionnalités souhaitables des outils.

4.2.1. Quelques rappels sur MacWeb

MacWeb est un prototype de recherche de système hypertexte en constante évolution depuis 88, développé au LIRMM ; il sert de banc d'essai à différents travaux sur les

¹²bien que la plus part de pages soient signées.

hypertextes. Il gère une structure appelée "web"¹³, matérialisée par un réseau de nœuds typés reliés par des liens typés. Conformément au modèle de Dexter [Halasz 94], le contenu informationnel des nœuds ("within component layer") est séparé du réseau lui-même ("storage layer"). Le mécanisme d'ancrage permet d'ancrer les liens sur des éléments du contenu des nœuds. L'interface auteur et lecteur est conçue pour une approche très interactive de la construction et de la mise à jour de l'hypertexte ("run time layer").

Contrairement à beaucoup d'autres systèmes, MacWeb attache une importance particulière au typage des informations et des relations qui les relient. Le typage, et en particulier le modèle à objet qui est utilisé dans MacWeb est un moyen pratique pour représenter les connaissances dans l'hypertexte. Les types de nœuds et de liens sont définis librement, mais un mécanisme réflexif permet de spécifier la sémantique du typage : les types sont eux-mêmes représentés par des nœuds reliés par des liens typés qui définissent les relations entre types. MacWeb permet de spécifier et d'éditer à tout moment interactivement les relations entre types et les relations entre instances à l'aide du même éditeur de graphe. Le réseau hypertexte peut être vu comme composé d'un "web principal" et d'un "web des types". Les nœuds et les liens du web principal sont considérés comme des instances des types définis dans le web des types. Ce dernier définit la structure entre les types comme un réseau de nœuds et de liens selon un modèle à objets avec héritage entre les types. Des nœuds de type "Script" attachés aux nœuds représentant des types spécifient les méthodes associées à ces types vus comme des classes.

4.2.2. Le choix de la structuration incrémentale dans MacWeb

Dans cette présentation, nous nous intéressons plus spécialement à l'aide, à l'émergence et à la réutilisation de structures régulières dans la construction incrémentale des hypertextes. Les mécanismes proposés présentent une compatibilité cognitive avec les processus mentaux du concepteur et avec la démarche opportuniste qui caractérise toute activité de conception. Ils s'appliquent essentiellement à la production *interactive* d'hypertextes. Il s'agit de fournir à l'auteur au cours du processus de création interactive des outils aptes à favoriser l'émergence et le réemploi de microstructures apparues librement.

L'approche retenue dans MacWeb s'appuie sur le prototypage¹⁴ et la fusion de structures pour ne pas entraver la créativité et épargner la définition a priori de schémas. En effet, les activités de jardinage d'information reposent sur des stratégies typiquement ascendantes, c'est à dire orientées de l'exemple vers l'abstraction. Il est donc fondamental, pour aider ces types d'activités, de fournir des mécanismes orientés vers la post-structuration d'information.

Quatre types de mécanismes sont présentés ici :

- la réutilisation d'éléments structurels par héritage grâce à une approche par objet
- la réutilisation de structures par clonage de prototypes
- la fusion de sous-structures produites isolement

¹³Web est le mot consacré depuis très longtemps dans la communauté scientifique étudiant les hypertextes. Il évoque la notion de toile d'araignée caractéristique des réseaux hypertextes. On le retrouve ainsi, par exemple, dans World Wide Web. Ce terme ayant été introduit pour désigner certaines structures de MacWeb, il est conservé ici, malgré sa consonnance anglo-saxonne.

¹⁴au sens Objet, des approches par prototypes.

- l'abstraction progressive et la désignation d'éléments structurels.

Ces choix fonctionnels visent à assurer une compatibilité cognitive avec les activités mentales du concepteur, en particulier :

- ne pas imposer de cadre structurel *a priori* pour la construction de la structure globale de l'hypertexte, mais aider à renforcer la structure qui émerge *a posteriori*. L'auteur -le créateur- ne doit pas être contraint sans raison. Par contre, il convient de l'aider à faire émerger une macrostructure cohérente par réutilisation d'éléments isomorphes à la microstructure qu'il a lui-même créé.
- permettre de réutiliser *facilement* des éléments structurels qu'il a lui-même composés, en utilisant comme prototype n'importe quel sous-ensemble du réseau hypertexte. La facilité d'emploi de ce mécanisme est garant de son emploi et par suite de l'homogénéité de la macrostructure construite.
- permettre de concevoir des sous-structures de façon indépendante, puis de les raccorder à des structures existantes ("Fusion") en propageant dans celles-ci les apports propres de la sous-structure.
- permettre à l'auteur une abstraction progressive de son travail de structuration, en dégageant progressivement et nommant les sous-structures utiles. Ceci revient à faire émerger explicitement des prototypes ou des modèles mieux identifiés et donc plus facilement réutilisables.

Il est très important de noter la grande liberté d'emploi de ces outils qui doit être laissée à l'auteur. Les outils doivent aider l'utilisateur et non l'entraver par le respect d'une structure qui en réalité ne prendra essence que lorsque la part créative du travail de l'auteur sera achevée

4.2.3. Les outils de la structuration incrémentale

4.2.3.1. La réutilisation par héritage

Le typage des nœuds et des liens est un des moyens efficaces pour renforcer la structuration dans les hypertextes. Le typage sert d'une part de façon interne dans la couche de représentation des connaissances de l'hypertexte, et, d'autre part de façon externe pour suggérer la structure sous-jacente au lecteur. On peut ainsi lui fournir, de façon anticipée, une information sur la sémantique de la traversée qu'il s'apprête à faire, l'aidant ainsi à se repérer au cours de sa navigation.

L'héritage dans MacWeb permet la réutilisation, au niveau des propriétés structurelles, des propriétés d'interaction, des propriétés concernant le contenu et la présentation des documents (textes, graphiques, son, vidéo). Un mécanisme plus général que celui disponible sous Hypercard permet de composer l'image d'un nœud par superposition d'images provenant, pour le premier plan, de son contenu propre et, pour l'arrière plan, de la superposition des différentes images spécifiées par le graphe d'héritage de son type. Cette image est active, elle possède des ancres permettant d'activer des liens définis pour ce type.

Ainsi tous les éléments de structure visuelle et d'interaction partagés par un ensemble de nœuds peuvent n'être décrits qu'une seule fois au niveau du nœud spécifiant leur type. L'héritage permet de hiérarchiser et de nuancer ces descriptions au gré des spécialisations ou des généralisations. Pareillement, les scripts accessibles depuis un nœud sont organisés selon le paradigme à objets. Chaque nœud dispose d'un ensemble de

souvent comme prototypes à partir de portions de web déjà construites avant même la création de modèles de structure.

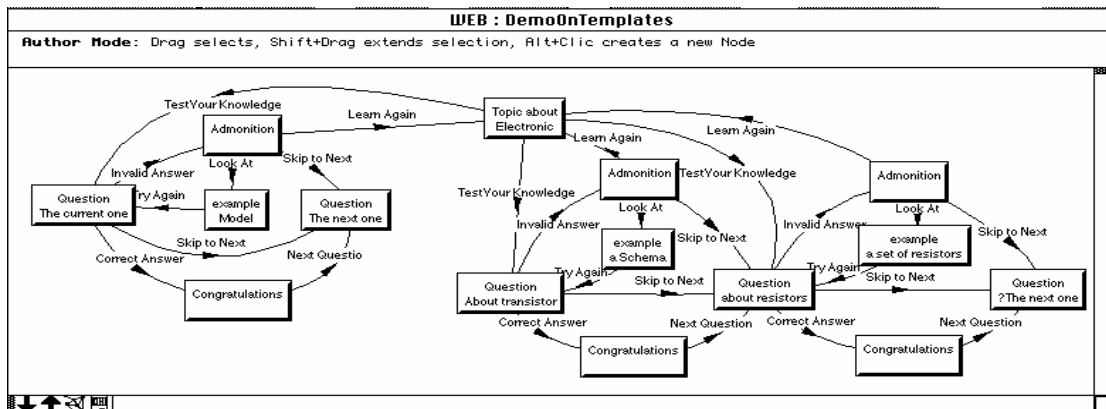


Fig. 4 : *Un exemple de structure obtenue par clonage.* La partie centrale, créée librement a servi de germe. Elle a été utilisée pour créer le motif de droite. Puis, l'intérêt se manifestant , l'auteur a isolé un prototype pur, à gauche, qu'il réutilisera pour engendrer la suite de la structure. Notons que toutes ces opérations se font par manipulation directe sur le graphe de la structure.

MacWeb permet le clonage à partir d'un prototype qui peut être n'importe quel sous graphe du Web. L'auteur qui a construit de façon libre une portion d'hypertexte, peut l'utiliser comme prototype pour engendrer d'autres portions qui auront la même organisation. Pour cloner un sous-web, il suffit de sélectionner les nœuds constituant ce sous-web et d'invoquer l'opération "clone" dans un menu. Le clonage réalise d'une part une duplication structurelle à partir du sous-web sélectionné, d'autre part reproduit aussi les contenus avec l'ancrage de leurs liens.

Ainsi, l'auteur peut bénéficier directement de la réplication de structures pré-remplies. Mais il a souvent intérêt à disposer plutôt d'un prototype "pur", ne contenant que les éléments qui seront utiles à tous ses descendants. Il construit ce prototype pur facilement par clonage d'une portion existante qui est ensuite expurgée des spécificités de l'instance initiale. MacWeb permet indifféremment de cloner un sous-web sélectionné ou un prototype explicitement nommé.

Contrairement à une approche qui serait guidée par un schéma préexistant, l'approche proposée permet une réutilisation opportuniste, palliant les difficultés de l'utilisateur à définir préalablement un modèle de structure. De plus, il est toujours possible de modifier interactivement la structure du descendant engendré. Le web résultant est un graphe possédant de nombreuses régularités mais dont la structure globale n'est pas nécessairement engendrée par la seule répétition d'un ou plusieurs motifs. Par exemple, il est possible de disposer d'une structure type utile pour décrire une ville, un individu, un événement... Il est, par contre, bien plus difficile de pré-définir dans quel contexte seront évoquées ces structures. Ceci évoque une analogie chimique : au sein de structures complexes il est usuel de reconnaître des groupements caractéristiques : tels amine, acide gras, cétone etc... Au sein des hypertextes, et malgré la liberté de création du réseau, il est souhaitable et utile de mettre en évidence et de favoriser la répétition de sous-structures.

L'approche proposée permet, en disposant de prototypes créés en situation réelle par l'auteur, de faciliter la structuration dans le cadre d'une démarche opportuniste.

4.2.3.3. Fusion de structures

Le phénomène de réutilisation de structure et les mécanismes de clonage ou d'instanciation qui permettent de le mettre en œuvre produisent des sous-webs souvent peu connectés au reste de l'hypertexte au moment de leur production. L'utilisation de ces sous-webs comme motifs intervenant dans la structure globale pose le problème de leur raccordement. En effet, seuls les homologues des liens existant entre le prototype et le reste du web assurent la connexion de la partie dupliquée. La mise à jour du réseau hypertexte suite à un clonage ne se réduit pas à la création de nouveaux liens mais nécessite aussi de pouvoir fusionner des nœuds.

En effet, dans les sous-webs clonés, il peut se faire que les nœuds engendrés doivent être identifiés à des nœuds existant déjà dans l'hypertexte. Par exemple, dans la structure décrivant un homme célèbre, apparaît un nœud relatif à la "ville" où il est né. Il se peut que cette ville soit déjà décrite dans l'hypertexte pour d'autres raisons.

Le mécanisme de fusion a un double but :

- 1) remplacer un ou plusieurs nœuds devant jouer des rôles identiques, par un nœud unique, en modifiant la structure en conséquence.
- 2) rassembler en un nœud unique l'information, l'ancrage et les liens qui appartenaient aux divers nœuds fusionnés.

En pratique, le contenu du nœud résultant de la fusion est formé de la concaténation des contenus. Les liens et les ancres suivent les contenus auxquels ils sont attachés.

Une étude expérimentale de la démarche des auteurs a mis en évidence l'intérêt de pouvoir invoquer la commande de fusion de deux façons :

- La *fusion explicite* dans laquelle l'auteur désigne explicitement dans le réseau les nœuds devant être fusionnés indépendamment de leur nom. Elle est souvent utilisée de façon naturelle immédiatement après un clonage de prototype, pour raccorder explicitement la nouvelle structure au reste du réseau. A l'issue du clonage, chaque nœud cloné se voit attribué un nom provisoire préfixé par le symbole "?" (fig. 2). Lorsqu'il utilise la fusion explicite, l'auteur n'a pas besoin d'affecter leur nom définitif aux nœuds qui peuvent temporairement garder leurs noms provisoires. Le nœud résultant de la fusion prend automatiquement le nom de celui qui n'a pas un nom provisoire, sinon celui du premier sélectionné. Ainsi la fusion explicite permet d'indiquer très facilement quelles parties de deux sous-structures sont communes : si l'une d'elles a déjà ses noms définitifs, ceux-ci se conservent.

- La *fusion implicite* concerne tous les nœuds ayant le même nom dans un même type. Elle est utile lorsque l'auteur utilise la sous-structure clonée comme une vue partielle qu'il met à jour sans se préoccuper du reste du réseau et souhaite son raccord automatique.

Le sous-web issu d'un clonage permet à l'auteur de s'intéresser localement aux seules informations relatives à son point de vue courant, tel que le décrit le prototype. Par exemple, l'auteur écrit dans le nouveau nœud "ville" les seules informations spécifiques au fait qu'un certain homme y a vécu, indépendamment d'autres informations déjà présentes dans un nœud "ville" de la partie principale de la structure. L'identité de nom exprimée par l'auteur traduit la désignation d'une même entité.

La fusion implicite a pour rôle de réunir les nœuds de même type qui ont un même nom. Elle évite à l'auteur, lorsqu'il a nommé le nouveau nœud, d'avoir à sélectionner explicitement ses homologues.

Typiquement, cette opération facilite une construction incrémentale de l'hypertexte où différentes facettes d'un même nœud peuvent être élaborées dans le cadre de sous-structures différentes représentant chacune un contexte particulier dans lequel ces caractéristiques prennent naissance. Ceci contribue à clarifier le processus de conception.

Ces deux mécanismes peuvent être utilisés pour "fusionner" partiellement un sous-web nouvellement créé à un web existant.

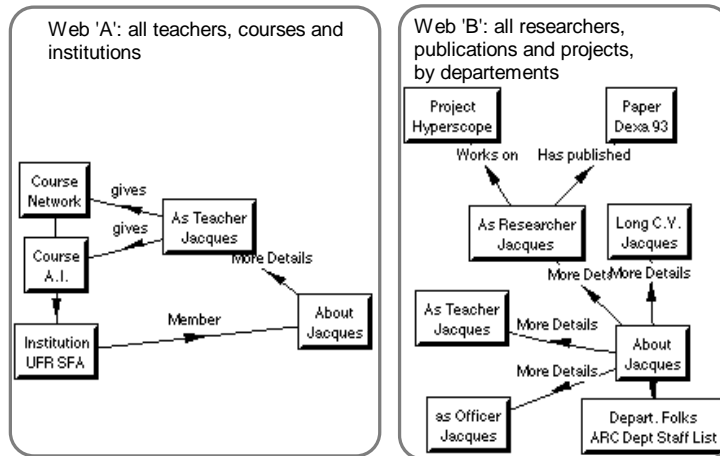


Fig. 5-1: La fusion de deux webs par fusion implicite. Les deux hypertextes 'A' et 'B' sont édités séparément, mais partagent des informations, ici le nœud About Jaques. Leur fusion produit le web ci dessous.

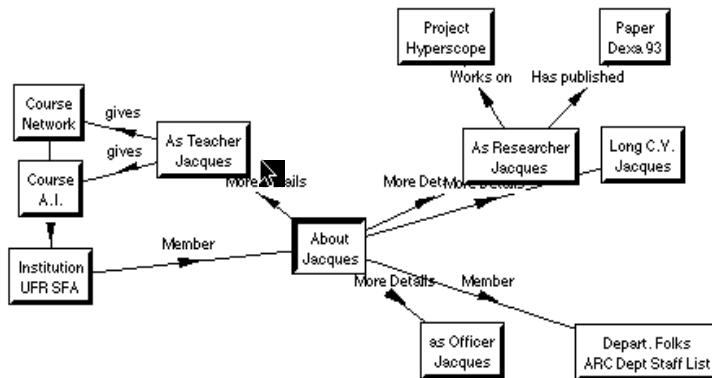


Fig. 5-2: Web résultant de la fusion de deux webs de la fig 3.

4.2.3.4. Abstraction progressive de la structure

Les mécanismes proposés permettent au processus de structuration de s'amorcer sans contrainte par clonage de parties du web. Dans ce processus, l'auteur fait progressivement émerger des microstructures modèles. Une fois reconnues, celles-ci acquièrent un statut de prototype. Cette montée progressive en abstraction implique qu'il devient souhaitable de pouvoir désigner globalement ces nouvelles structures. C'est le rôle des *groupes de nœuds*, qui permettent de manipuler tout sous-ensemble du web comme un nœud unique désignable par un nom.

Lorsque l'auteur a isolé un prototype, il lui suffit de demander la création d'un groupe pour désigner ce sous-web, et d'attribuer un nom à ce groupe. Il a ainsi abstrait le concept de cette microstructure. L'ensemble de cette démarche est typiquement un processus incrémental de structuration et donc d'acquisition et de représentation de connaissance. Il a pour résultat d'une part un web structuré et d'autre part les éléments générateurs de cette structure. La structure du web ne résulte pas d'un schéma préexistant; elle résulte *a posteriori* des choix de construction de l'auteur. La facilité

d'emploi des mécanismes proposés conduisent naturellement celui-ci à construire librement et incrémentalement des réseaux riches en régularités.

4.2.4. L'apport de la structuration incrémentale

MacWeb, en tant que banc d'essai de techniques hypertextes, est un outil extrêmement général dans lequel se côtoient des outils qui étaient initialement destinées à des types d'applications extrêmement différentes. MacWeb a été utilisé aussi bien pour des activités d'extraction d'information dans des corpus de textes techniques analysés de façon automatique, pour des activités de valorisation d'information, où l'aspect structurel et multimédia est prépondérant, que pour des tâches appartenant typiquement au jardinage d'information. Dans ces dernières c'est l'interactivité et l'incrémentalité de l'outil qui a constitué le point fort. Cette diversité d'expériences, et surtout le fait de pouvoir, pour chacune d'elles, disposer d'une large palette d'outils a permis d'expérimenter sans contraintes. Nous avons retenu de ces expériences que la liberté était une condition strictement nécessaire à la créativité, et que les outils devaient permettre et favoriser cette liberté. Mais nous avons constaté parallèlement que la liberté bien maîtrisée n'implique pas le désordre structurel. Les outils doivent donc simultanément ne pas contraindre l'auteur et l'entraîner vers plus de rigueur. Ceci est tout à fait réalisable lorsque les outils réduisent l'effort nécessaire pour accomplir une tâche de façon ordonnée, par rapport à sa réalisation désordonnée. Typiquement le clonage de prototype va dans ce sens : l'auteur n'est jamais précontraint; il peut innover structurellement chaque fois que cela lui semble souhaitable, par contre, partout où cela est possible, le clonage est plus économique pour lui et engendre une structure plus régulière.

5. CONCLUSION

Les hypertextes et les techniques éducatives sont deux domaines qui ont très tôt coopéré et qui ont bénéficié l'un de l'autre depuis de nombreuses années. Beaucoup a déjà été dit. C'est pour cela que j'ai évité d'aborder les classiques problèmes de navigation, d'aide, de désorientation,... etc. et ai plutôt cherché à offrir une vue synthétique du domaine, sans la couper des réalités.

Bien que les hypertextes répondent tous à un même modèle fonctionnel simple, qui justifie leur intérêt en tant qu'outil pour les techniques éducatives, il existe des différences importantes entre les systèmes hypertextes, en fonction des catégories d'application visées. Nous avons vu que l'attitude de l'utilisateur face à l'information induisait des grandes classes d'applications (extraction, valorisation, création) et montré en quoi leur spécificité a des répercussions importantes sur les systèmes hypertextes qui leur sont adaptés. En particulier nous avons insisté sur l'usage des hypertextes dans les activités créatrices, et vu comment des outils spécifiquement adaptés peuvent aider le processus créatif.

Un hypertexte est souvent considéré soit du côté de son auteur soit de celui du lecteur. Je plaide pour que l'on utilise aussi les hypertextes en même temps comme auteur et comme lecteur. Un hypertexte n'est pas seulement une encyclopédie sur CD-ROM, une mine à information sur l'Internet, c'est d'abord un instrument de travail, capable d'aider de façon continue l'enseignant depuis l'apparition de ses idées jusqu'à leur présentation. Enfin, l'hypertexte est un aussi merveilleux outil de travail pour l'enseigné, la construction d'hypertextes étant l'une des activités intellectuelles les plus formatrices et motivantes en raison de l'écart entre son apparente simplicité et sa complexité réelle.

6. BIBLIOGRAPHIE

- [Akscyn 88] R.M. Akscyn, D.L. McCracken, E. A. Yoder, "KMS: a distributed hypermedia system for managing knowledge in organizations", *CACM*, 31(7), 1988, pp. 820-835.
- [Balpe 90] J.P. Balpe, "Hyperdocuments, hypertextes, hypermédias", Eyrolles 1990.
- [Ben Mrad 94] M. Ben M'rad, " Agrégats et composition de requêtes dans les hypertextes virtuels ", Thèse, EPFL Lausanne, 1994.
- [Bernstein 92] M. Bernstein, M. Joyce, D. Levine, "Contours of constructive hypertexts", in [Lucarella 92].
- [Bernstein 93] M. Bernstein, "Enactement in information farming", *Proc. Hypertext'93*, ACM Press, Seattle, Nov. 1993.
- [Boy 91] G. Boy, " Indexing Hypertext Documents in Context", *Proc. Hypertext'91*, San Antonio, ACM Press, Dec. 1991.
- [Bruillard 94] E. Bruillard, B. de La Passardière, "Hypermédia et éducation", *Sciences et Techniques Educatives*, Afcet, Vol. 1, n°1, Hermès 1994.
- [Caloini 92] Caloini A., "Matching hypertext models to hypertext systems, a compilative approach", in [Lucarella 92], pp. 91-101.
- [Deschamps 94] R. Deschamps, "Bases de connaissances généralisées : une approche fondées sur un modèle hypertexte expert ", Thèse de Doctorat, Univ. de Pau, 1994.
- [Garzotto 91] F. Garzotto, P. Paolini, D. Schwabe, "HDM - a model for the design of hypertext applications", *Proc. Hypertext'91*, ACM Press, San Antonio (Tx), pp. 313-328, Dec. 1991.
- [Garzotto 94] F. Garzotto, L. Mainetti, P. Paolini, "Adding Multimedia collections to the Dexter Model". *Proc. ECHT94*, ACM Press, Edinburgh, Sept. 94.
- [Haake 94] J. Haake, *et al.*, "Coexistence and transformation of informal and formal structures, Requirements for more flexible hypertext systems", *Proc. ECHT94*, ACM Press, Edinburgh, Sept. 94.
- [Halasz 90] F. Halasz, M. Schwartz, "The Dexter Hypertext Reference Model", *Proc. Workshop on Hypertext Standardization*, pp. 95-133, 1990.
ré-édité dans :
- [Halasz 94] F.Halasz, M. Schwartz, "The DEXTER Hypertext Reference Model", *CACM*, 37(2), pp. 30-39.
- [Hayes 80] J.H. HAYES, L.S. FLOWER, "Identifying the Organization of Writing Processes," in *Cognitive processes in writing*, L.W. GREGG, E. R. STEINBERG, eds, Lawrence Erlbaum Associates, Publishers, 1980.
- [Lelu 92] Lelu A. , François C., "Hypertext paradim in the field of information retrieval : a neural approach", in [Lucarella 92], pp. 112-122.
- [Lucarella 92] *Proceedings ECHT'92, 4th ACM Conf. on Hypertext*, Milano, D. Lucarella, J. Nanard, M. Nanard, P. Paolini, eds, ACM Press, Dec.1992.
- [Marshall 92] C. Marshall et al., "Two Years before the Mist: Experiences with Aquanet", in [Lucarella 92].
- [Marshall 93] C.C. Marshall, F.M. Shipman III, "Searching for the missing link: Discovering Implicit structure in spatial hypertext", *Proc. ACM Conf. Hypertext'93*, ACM Press, Seattle, Nov. 1993.
- [Marshall 94] C.C. Marshall, F.M. Shipman III, J.H. Coombs, "VIKI: Spatial hypertext supporting emergent structure", *Proc. ECHT94*, ACM Press, Edinburgh, Sept. 1994.

- [Nanard 91] J. Nanard, M. Nanard, "Using Types to Incorporate Knowledge in Hypertext", *Proc. Conf. Hypertext'91*, ACM Press, San Antonio (Tx), pp. 329-344, Dec. 1991.
- [Nanard 93-a] M. Nanard et al., "La métaphore du généraliste : Acquisition et utilisation de la connaissance macroscopique sur une base de documents techniques", Actes des Journées Acquisition de Connaissances, Saint Raphaël, Mars 1993.
- [Nanard 93-b] J. Nanard, M. Nanard, "Should anchors be typed too? An experiment with MacWeb", *Proc. HTX93, 5th ACM Conf. on Hypertext*, ACM Press, Seattle, Nov.1993.
- [[Nanard 94-a] M. Nanard, J. Nanard, "Helping Emergence of Structures in Incremental Design of Hypertext", *Proc. Interface to Real and Virtual Worlds*, Montpellier, ISBN 2-2-906899-83-6, Mar. 1994.
- [Nanard 94-b] M. Nanard, J. Nanard, (1994). *Some Thoughts and Experiences on Approaches and Tools for Designing Hypertext Structure*, LIRMM RR 94028.
- [Nanard 94-c] J. Nanard, "Multimedia and Hypermedia facing the concept of active document", *Proc. BIWIT'94*, Basque international Workshop on Information Technology, Anglet 6-8 Mars 1994.
- [Nelson65] T.H. Nelson, "A File Structure for the Complex, the Changing and the Indeterminate", *Proceedings of the National Conference*, 20th. New York : ACM, 1965. pp. 84-100 (Document ancien, une description plus récente de ses travaux se trouve dans: [Nelson80]).
- [Nelson80] T.H. Nelson, "Replacing the printed word: a complete literary system", In : *Proceedings IFIP Congress*, 1980, North-Holland, 1980, pp. 1013-1023.
- [Nelson 81] T.H. Nelson, "Literary Machines", Swarthmore, Pensylvanie, 1981, ré-édité : Mindfull Press, Ca, 1993.
- [Remde 87] J.R. Remde *et al.*, "Superbook; an automatic tool for information exploration? , in *Proc Hypertext'87*, pp. 175-188.
- [Shieber 94] S.M. Shieber, "Lessons from a restricted Turing test", in *CACM* Vol. 37, n°6 pp.70-78, 1994.
- [Streitz 89] N.A. Streitz, J. Hanneman, M. Thuring, "From ideas and arguments to hyperdocuments: travelling through activity spaces", *Hypertext'89 Proceedings*, ACM Press, Baltimore, 1989, pp. 343-364.
- [Streitz 92] N. Streitz *et al.*, "SEPIA:a cooperative hypermedia authoring environment", in [Lucarella 92].
- [Yankelovitch88] N. Yankelovitch, B.J. Haan, N. Meyrowitz, S.M. Drucher, "Intermedia: the concept and the construction of a seamless information environment", *IEEE Computer*, 21 (1), 1988, pp. 81-96.