

CHAPITRE 10

LES MÉTHODES ON-LINE 1 : ANALYSE DES PARCOURS

FRANCK AMADIEU, CLAUDE BASTIEN et ANDRÉ TRICOT

1. INTRODUCTION

Comprendre l'activité de recherche d'information dans les documents et les systèmes documentaires est un objectif partagé par de nombreux chercheurs. Case (2002) estimait il y a 5 ans que 10 000 publications traitaient de cette question. Dans ce chapitre, nous abordons plus spécifiquement l'analyse des parcours des utilisateurs de documents électroniques. L'enjeu de ces travaux est triple :

- Peut-on interpréter le comportement des utilisateurs de documents pour le qualifier ou le quantifier en termes de performance ?
- Peut-on interpréter le comportement d'un utilisateur pour inférer la nature des difficultés que celui-ci rencontre dans son activité d'utilisation ?
- Dans quelle mesure les différents indicateurs des comportements peuvent nous renseigner sur les processus cognitifs sous-jacents ?

Différentes méthodes d'analyse des parcours sont utilisées depuis une quinzaine d'années, dont la plupart sont encore discutées, sinon en cours d'élaboration. Ces méthodes sont présentées dans la première partie du chapitre.

La seconde partie présente les méthodes consacrées à la description de la suite des actions de l'utilisateur : contrairement à la première catégorie de méthodes, celle-ci rend compte de l'ordre des actions. On est donc généralement conduit à élaborer une grammaire, un système de codage des actions. Elaborer de telles grammaires ne présente pas de difficulté particulière. Toutefois une grammaire n'a aucune pertinence en soi. Sa pertinence est relative à son pouvoir de description au regard des variables étudiées.

Ainsi, à l'image de ce que nous venons d'évoquer à propos des grammaires de parcours, chaque partie donne lieu à une discussion de la pertinence des mesures présentées. En effet, à travers l'analyse des différentes mesures (variables dépendantes) utilisées dans la littérature, nous défendrons le point de vue selon lequel l'intérêt d'une mesure réside dans le fait qu'elle permet de contribuer à l'élaboration d'un modèle psychologique (ou de comprendre un processus psychologique) plutôt que de quantifier la performance *per se*. L'accumulation de résultats empiriques n'a pas de signification psychologique en soi. Autrement dit, il nous semble important de pouvoir mettre en relation l'analyse du parcours et l'analyse de la performance à la tâche principale en l'inscrivant dans le cadre fourni par un modèle psychologique.

Les mesures que nous recensons sont utilisées dans un domaine de recherche en émergence, ce qui explique leur hétérogénéité et donc la difficulté à accumuler de façon raisonnée des résultats empiriques.

2. LES MESURES D'EFFICACITÉ DU PARCOURS

Dans les travaux que présentons ici, le parcours est considéré comme un ensemble et non pas comme une suite : l'ordre n'est pas pris en compte. Il s'agit de mesures de performances relativement sommaires. Cette simplicité explique sans doute le succès de ces variables dépendantes.

2.1. LES TAUX DE RAPPEL ET DE PRÉCISION

Une première catégorie de méthodes consiste à mesurer l'efficacité de l'ensemble des actions (des sélections de documents ou de sous-parties du document) réalisées par l'utilisateur lors de la réalisation de la tâche d'utilisation du document. Les mesures d'efficacité du parcours reprennent les mesures traditionnelles issues des sciences du document, utilisées au départ pour mesurer les performances de techniques documentaires (par exemple, pour comparer l'efficacité de deux techniques d'indexage) : les taux de rappel et de précision (Buckland & Gey, 1994). Les indices utilisés pour établir ces mesures sont :

T, le nombre total de nœuds ouverts ; « Nœud » est utilisé comme terme générique pour désigner une unité documentaire. Un nœud peut-être un article, une page, une image, *etc.* C'est ce à quoi on arrive quand on clique sur un lien hypertextuel. Un nœud ouvert est donc une unité documentaire consultée par l'utilisateur.

F, le nombre de nœuds ouverts plus d'une fois ;

D, le nombre de nœuds différents ouverts ;

R, le nombre de nœuds pertinents ou « cibles » (une cible est un nœud qui correspond au besoin de l'utilisateur, ou, dans le cadre d'une expérimentation, au but de la tâche, le « besoin de l'utilisateur » étant alors défini par l'expérimentateur) ; ou R', le nombre de nœuds du parcours optimal (c'est-à-dire le nombre de nœuds nécessaires et suffisants à la réalisation de la tâche) ;

O, le nombre de cibles effectivement ouvertes par l'utilisateur ; O' nombre de nœuds du parcours optimal effectivement ouverts ; O'' nombre de nœuds du parcours optimal effectivement ouverts au moins une fois.

U, le nombre de nœuds non pertinents ouverts par l'utilisateur. Nous présentons une analyse détaillée du tableau 10.1.

<i>nom de la mesure</i>	<i>formule</i>	<i>exemple de référence</i>
Taux de rappel	O/R	Buckland & Gey, 1994
	O/R'	Stanton, Correia, & Dias, 2000
	O'/R'	Dias & Sousa, 1997
	O (quand R est fixe)	Downing, Moore, & Brown, 2004
Taux de précision	O/T	Buckland & Gey, 1994
	O'/T	Dias & Sousa, 1997
	O' / (O' + U)	Dias & Sousa, 1997
	T	McDonald & Stevenson, 1998
	F	Ahuja & Webster, 2001
	O / (R + U)	Rouet, 2003

Tableau 10.1. – *Principales mesures d'efficacité du parcours de l'utilisateur d'un document.*

Ces mesures de performances correspondent donc au fait que l'utilisateur sélectionne le plus possible de documents pertinents (taux de rappel) et le moins possible de documents non pertinents, voire de

documents en général (taux de précision). La validité de ces mesures est donc relative au concept même de pertinence : s'il n'y a pas d'ambiguïté quand au statut de celle-ci, si un document défini comme pertinent par l'expérimentateur l'est réellement, alors ces mesures sont valides. Elles ne sont rien d'autre qu'une mesure du taux de « bonnes réponses » et de précision de ces réponses.

Comme le font remarquer de nombreux auteurs (par exemple, Salampasis, Tait, & Bloor, 1998), ces mesures ne sont valides que si le but de la tâche est de rechercher une information, définie de façon suffisamment précise. Ces mesures ne sont pas valides quand le but est parcourir de nombreuses informations, quand le but est flou, ou quand on ne sait pas quel document est pertinent.

L'utilisation de l'indice R' à la place de R ou de O' (ou même de O'') à la place de O semble pertinente dans les hypertextes où la cible n'est pas atteignable directement : il faut passer par un certain nombre de nœuds intermédiaires. Quand l'expérimentateur n'a pas de tâches à comparer entre elles, l'utilisation de T ou de F pour mesurer la précision devient possible. Enfin, l'indice de précision utilisé par Rouet (2003) présente l'avantage d'utiliser la même fraction de base que l'indice de rappel (O/R), en ajoutant U au dénominateur.

2.2. LES AUTRES MESURES D'EFFICACITÉ : MESURES DU TEMPS, DE LA REDONDANCE ET DE LA DÉSORIENTATION

Il existe d'autres mesures de l'efficacité des comportements de navigation. Parmi les plus utilisées on trouve bien entendu les temps de réalisation de la tâche. Cette mesure devient souvent l'indicateur de performance dans les tâches de recherche d'information, supplantant les performances relatives au nombre d'informations cibles atteintes par les participants (par exemple Lin, 2003 ; McDonald & Stevenson, 1998). A cette mesure est souvent associée une mesure du nombre de nœuds consultés pour atteindre le but de la tâche (ex. Nilsson & Mayer, 2002). Mais pour atteindre une meilleure compréhension des comportements de navigation, le nombre de nœuds consultés peut être pondéré par une dimension de redondance des parcours. Nous présentons ici deux indices de redondance (*cf.* tableau 10.2).

Economie	D / T	Tricot et al., 1999
Taux d'accès	$R' / (R' + U)$	
Indice de temps	Temps O' / Temps T	Dias & Sousa, 1997
		Lin, 2003
*Redondance	Nb nœuds activés plusieurs fois $(T-D) / T$	McDonald & Stevenson, 1998 Padovani & Lansdale, 2003
*Désorientation (L)	$(D/T - 1)^2 + (R/D - 1)^2$	Smith, 1996

Tableau 10.2. – *Autres mesures d'efficacité ou d'inefficacité (*)*.

Tricot, Puigserver, Berdugo et Diallo (1999) ont par exemple proposé un indice d'économie proche de l'indice de précision. L'indice d'économie renvoie au « nombre de nœuds ouverts différents par l'utilisateur divisé par le nombre total de nœuds ouverts par les participants », soit D / T . Ainsi, plus un utilisateur répète l'activation d'un même nœud, moins sa navigation est économique. Cet indice peut être considéré comme un indice de redondance, il ajoute une information supplémentaire à l'indice de précision, que les auteurs utilisent par ailleurs. La nécessité de multiplier les indices du comportement, peut être illustrée par l'étude de Lin (2003). Dans cette étude, l'auteur a évalué les comportements d'exploration libre d'un site, en relevant le nombre de nœuds ouverts (T) mais surtout le nombre de nœuds ouverts plus d'une fois (F) qui ici est considéré comme un indicateur de désorientation¹. L'utilisation des deux indices a permis de mettre en évidence pour un site structuré hiérarchiquement (c'est-à-dire nœuds organisés hiérarchiquement avec une liberté de navigation contrainte au respect de la hiérarchie), une exploration plus importante (nombre d'ouvertures) mais surtout avec moins de redondance (nombre de nœuds rouverts) comparativement à un site structuré en réseau (c'est-à-dire nœuds reliés entre eux par des liens référentiels et donc offrant une navigation totalement libre). Dans un protocole similaire, McDonald et Stevenson (1998) ont mis en évidence que des experts du domaine ouvraient plus de nœuds et faisaient moins de

¹ La désorientation d'un utilisateur dans un document électronique peut être définie comme un état cognitif traduisant des difficultés de repérage de l'organisation et des bornes du document et par conséquent de son propre parcours et position dans l'espace documentaire. L'utilisateur possède une représentation parcellaire ou erronée de l'espace dans lequel il évolue.

réouvertures que des novices, indiquant ainsi une exploration plus exhaustive et moins redondante et donc une faible désorientation des experts.

L'évaluation de la redondance des parcours permet donc d'évaluer l'étendue de l'exploration réalisée par les participants lorsqu'elle est associée à l'indice T. Padovani et Lansdale (2003) propose un autre indice de redondance des parcours dans une tâche de recherche d'information : $(T-D) / T$. Ceci permet d'estimer l'étendu du parcours réalisé par les utilisateurs dans un hypertexte, et donc traduit un comportement confiné à un ensemble de nœuds ou au contraire un comportement beaucoup plus exploratoire. Plus le nombre de nœuds différents ouverts est important, plus l'indice de redondance est faible (si $D=T$ alors la valeur de l'indice est de 0).

Les indices de redondance peuvent être également utiles dans les tâches de lecture - compréhension. Dee Lucas et Larkin (1995) ont par exemple utilisé un indice des stratégies de relecture des nœuds d'un hypertexte pour mettre en évidence qu'une structuration en liste d'un hypertexte induisait une stratégie délibérée de revue et de vérification, alors qu'une structure hiérarchique produisait davantage un parcours (ou feuilletage) répété du contenu du texte.

La désorientation dans les hypertextes est très souvent abordée via des mesures subjectives (par exemple Brinkerhoff, Klein, & Koroghlanian, 2001 ; Müller-Kalthoff & Möller, 2003). Or, comme nous l'avons vu précédemment, un indice de précision ou un indice de redondance peuvent être employés comme mesures de la désorientation de l'utilisateur. Smith (1996) propose une mesure d'efficacité des parcours qui offre une mesure quantitative plus fine basée sur la dégradation des performances. Celle-ci consiste à mesurer les déviations par rapport à un parcours optimal dans un hypertexte. Ce coefficient de désorientation L (cf. tableau 10.2) repose sur un indice de redondance $(D/T - 1)^2$ et un indice de précision $(R/D - 1)^2$. L signifie une très forte désorientation lorsqu'il tend vers 1 et inversement un parcours de recherche parfait lorsqu'il tend vers 0.

La plupart des mesures présentées ici portent sur le nombre d'ouvertures des nœuds composant le système, mais le temps passé sur chaque nœud ou sur un type de nœud peut fournir de nouvelles informations sur la nature des parcours. Par exemple, Dias et Sousa (1997) ont mesuré un indice de temps correspondant au temps passé dans les nœuds du parcours optimal sur le temps total passé sur l'ensemble des nœuds. Cet indice est en fait une transposition de

l'indice de précision (O/T) à des valeurs temporelles. Ainsi, l'ensemble des indices présentés précédemment pourrait être également appliqué à des données temporelles. Il est possible d'imaginer que la précision d'un parcours peut être modulée par les données temporelles. Par exemple, un utilisateur qui ouvre beaucoup de nœuds non pertinents aura plutôt un indice de précision faible. Or si le temps passé sur ces nœuds non pertinents est très faible (relativement au temps passé sur les nœuds pertinents), cela pourrait traduire une distinction réussie de la pertinence des nœuds par le sujet, à l'inverse, un temps de consultation important des nœuds non pertinents peut signifier que le sujet n'a pas repéré la pertinence des nœuds. Ainsi, le temps peut être un indice supplémentaire pour la compréhension des comportements. L'utilisation d'un indice seul n'est pas suffisante pour atteindre une compréhension suffisante des comportements. Il est nécessaire de multiplier les mesures afin de définir un pattern de performance de l'utilisateur dans une tâche. De plus, un comportement observé doit être interprété sur la base d'un modèle de l'utilisateur, car un même comportement peut avoir différentes significations. Par exemple, les retours en arrière dans un parcours peut être un indicateur d'une activité d'établissement d'une cohérence avec la page en cours de consultation, ou bien la nécessité de revenir sur la page précédente car elle sert de point de repère pour la navigation (rationalisation).

2.3. LES MESURES D'EFFICACITÉ PONDÉRÉE OU EFFICIENCE

Les indices d'efficacité de l'activité de navigation peuvent être pondérés ou normalisés afin de permettre des comparaisons inter-sujets, inter-groupes ou inter-documents. Des mesures d'efficacité sont alors utilisées. De façon générale, l'efficacité renvoie au rapport entre les performances (scores d'apprentissage, atteinte de l'objectif, précision de la réalisation de la tâche) et les ressources qui ont été mobilisées pour atteindre ces performances, d'où la nécessité de considérer des indices d'efficacité. Nous présentons ici l'exemple de la mesure d'un taux de précision pondéré par les scores à une tâche de recherche d'information (Padovani & Lansdale, 2003) et la mesure d'Otter et Johnson (2000) qui proposent de pondérer l'indice de désorientation de Smith (1996) par le poids sur la désorientation des

liens utilisés (par exemple un lien associatif aura un point plus important qu'un lien organisationnel).

Dans le champ d'étude des apprentissages, les mesures de performances sont souvent pondérées par d'autres mesures afin d'établir par exemple les réels bénéfices d'un dispositif d'apprentissage. Par exemple, certains travaux issus de la théorie de la charge cognitive pondèrent généralement les performances des apprenants par le coût mental qu'a représenté la tâche pour les sujets au travers d'un calcul d'efficacité mentale (Pass & van Merriënboer, 1993). Dans le domaine des hypermédias, Lee et Tedder (2003) ont par exemple obtenu des résultats indiquant qu'un hypertexte structuré en réseau tendait à favoriser les performances de rappel des participants comparativement à un hypertexte structuré hiérarchiquement. Or lorsque les temps de consultation ont été pris en compte comme covariable dans les analyses statistiques, la différence de performance entre les deux conditions disparaît. Ce type de constat permet de soutenir l'idée que les différences de performance observées peuvent être induites par des différences de temps d'apprentissage générées par les dispositifs étudiés.

Padovani et Lansdale (2003) ont mesuré l'efficacité de navigation dans une tâche de récupération de nœuds précédemment consultés. L'indice correspond à $(O/R) * (R') / T$. Ainsi l'efficacité est traduite ici par la relation entre le score relatif au nombre de cibles trouvées (O/R) et un taux de précision (R'/T). Ceci suggère que l'utilisation d'un taux de précision seul n'a de sens que lorsque les sujets ont trouvé toutes les informations cibles. L'indice d'efficacité de Padovani et Lansdale (2003) correspond à une normalisation du taux de précision faite sur la base des scores de récupération des cibles. Ceci permet de comparer les performances des participants quel que soient leurs scores de récupération.

Concernant les indices de désorientation, Otter et Johnson (2000) ont proposé une nouvelle mesure quantitative consistant en une pondération du coefficient de Smith (1996) par la nature des hyperliens composant le système. Les auteurs distinguent les liens associatifs, taxonomiques, séquentiels et annotationnels. Ces types de liens se distinguent sur la base de leur degré de prédictibilité de la localisation d'une information cible. Par exemple, les liens associatifs (liens arbitraires en les nœuds) participeraient beaucoup à la désorientation car ils sont porteurs d'un faible degré de prédictibilité. Ainsi, les auteurs ont révisé la mesure de Smith (L), précédemment

présentée, en la pondérant avec les poids des liens : $LWLM = L/(LW/4)$, LW renvoyant au poids des liens (*link weighting*). Une valeur faible de LW signifie que le lien ne favorise pas la prédictibilité et donc augmente la désorientation. Néanmoins, une étude expérimentale menée par les auteurs (Otter & Johnson, 2000) n'a pas véritablement confirmé une plus-value de leur coefficient face à celui de Smith (1996) en termes de corrélations avec des mesures subjectives de désorientation (échelles en 10 points évaluant par exemple le sentiment d'être désorienté).

Efficienne des parcours	$(O/R)*(R') / T$	Padovani & Lansdale, 2003
*désorientation (LWLM)	$L/(LW/4)$	Otter & Johnson, 2000

Tableau 10.3. – *Mesures pondérées d'efficacité ou d'inefficacité* (*).

2.4. PERTINENCE PSYCHOLOGIQUE DE CES VARIABLES DÉPENDANTES

De nombreux auteurs que nous venons de citer combinent, de façon fort pertinente, les mesures que nous venons de présenter avec des descriptions plus qualitatives. Certains décrivent l'appropriation des fonctionnalités du document par l'utilisateur (Bera & Liu, 2004 ; Padovani & Lansdale, 2003). Danielson (2002) par exemple utilise une palette très complète de mesures de l'utilisation des fonctionnalités comme moyen d'analyse de l'activité de navigation : il compte le nombre de clics sur des fonctions comme « retour en arrière », « page suivante », ainsi que les caractéristiques des déplacements entre nœuds : saut vers un nœud sur-ordonné, sous-ordonné, latéral, *etc.*

De nombreuses discussions ont porté sur la pertinence psychologique de ces mesures, parfois de façon virulente, et ceci depuis longtemps (Bernstein, Joyce & Levine, 1992). Selon le modèle sous-jacent, différentes interprétations peuvent être attribuées à la même valeur d'une variable. Il n'est pas le lieu de présenter ici ces controverses et ses contradictions. Nous proposons de considérer :

- que les taux de rappel et de précision peuvent être utilisés pour mesurer la performance des participants si et seulement si (a) la tâche consiste à rechercher une ou plusieurs informations précisément définies et (b) la recherche doit être réalisée de manière efficace ;
- les autres mesures d'efficacité, y compris le temps, d'inefficacité ou de redondance ne doivent pas être utilisées comme seules variables dépendantes, elles doivent être combinées avec d'autres variables, mesurant la performance principale des utilisateurs (ex. la compréhension, l'apprentissage) pour être interprétées, et ne sont pertinentes que pour les tâches qui doivent être réalisées de manière efficace ;
- les mesures d'efficience sont utilisables avec les mêmes contraintes ; elles sont utilisées quand des comparaisons doivent être réalisées.

3. LES DESCRIPTIONS DE PARCOURS COMME SUITE D' ACTIONS

La plupart des mesures présentées précédemment peuvent être qualifiées de « statiques » (nombre de nœuds ouverts, longueur de parcours, temps de consultation) car elles ne prennent pas en compte les caractéristiques d'ordre des parcours de navigation.

Or, il paraît évident que les parcours reposent également sur les dimensions d'ordre et de temps : la lecture, la consultation d'un document est un processus purement séquentiel. C'est pourquoi nous présentons dans ce chapitre différents formalismes employés pour la mesure de la navigation en considérant les différentes dimensions qui la caractérisent. A la frontière entre ces deux parties, nous pouvons citer Smith (1996), qui distingue de façon intéressante les mesures selon les différentes étapes de réalisation de la tâche principale (planification, recherche d'information ou lecture, retour aux questions, vérification ou révision, *etc.*).

Les descriptions de parcours dans le domaine des documents correspondent au domaine des grammaires d'actions en IHM (Kieras & Polson, 1985 ; Payne & Green, 1986).

3.1. MESURE DE LA DISTANCE ENTRE UN PARCOURS MODÈLE ET UN PARCOURS OBSERVÉ

Pour définir un modèle de navigation de l'utilisateur il convient dans un premier temps de constituer un modèle du dispositif et du contexte. Le modèle de navigation ne prendra sens que confronté aux modèles du dispositif dans lequel l'utilisateur se déplace ou bien à d'autres modèles de navigation. Par exemple, Lawless et Brown (1997) distinguent trois profils de navigation : (1) les chercheurs de connaissances (2) les explorateurs de caractéristiques et (3) les utilisateurs apathiques. Ainsi sur la base de modèles de navigation existants, l'étude de patterns de navigation des participants dans un hypertexte passe par la modélisation des comportements suivie par la confrontation de ces modélisations aux différents modèles.

Pour la modélisation des hypertextes et des comportements de navigation plusieurs types de mesures existent. La plupart des méthodes de mesure, comme celle de Salampasis et al. (1998), sont issues des recherches sur les hypermédias adaptatifs ou à partir de la théorie des graphes. Deux types de graphes sont distingués : d'une part, le graphe du site – modélisation de la structure des nœuds et des liens – et d'autre part, le graphe de l'utilisateur – modélisation des nœuds consultés et des liens activés par l'utilisateur (Herder, 2002). Mais ce type de modélisation n'intègre pas une donnée importante, celle de la temporalité des comportements. C'est pourquoi, d'autres attributs peuvent être implémentés dans la modélisation de ces graphes. Les mesures présentées dans les paragraphes précédents étaient unidimensionnelles, or l'analyse des structures de graphes permet d'appréhender plusieurs dimensions des hypertextes et des comportements de navigation comme : la taille d'un hypermédia (nœuds et liens), la complexité (degré de liberté de navigation), la linéarité (ordre linéaire implicite dans la structure d'un réseau), la distance (longueur d'un parcours entre 2 nœuds) ou encore l'organisation en classes (positions stratégiques de nœuds dans un bloc thématique).

Parmi les méthodes de modélisation on citera par exemple l'approche d'analyse et de visualisation des données de navigation de Berendt et Brenstein (2001) : STRATDYN. La méthode de visualisation des comportements propose trois manières indépendantes de groupements perceptuels : ordre des visites des nœuds (souligne la

découverte de l'information), les mouvements dans l'hypertextes (souligne la nature de la navigation) et la tâche expérimentale (représentation des nœuds et liens pour les différentes tâches expérimentales). Ces auteurs abordent alors les représentations obtenues des comportements de navigation en les comparant à des modèles de patterns comme les « patterns de recherche en largeur en premier » (l'utilisateur explore un a un tous les nœuds accessibles à partir d'un nœud A - A, B1, A, B2, A, B3, A, B4), les « patterns de recherche en longueur en premier » (la consultation des noeuds se fait en une séquence ininterrompue - A,B,C,E...) ou encore les « patterns de points de repère » (un nœud est un point de repère lorsque il a plus de 5 liens entrant). STRATDYN peut aussi traiter le contenu sémantique des liens entre nœuds et ainsi, rendre compte d'une certaine sémantique du parcours (Berendt, 2002). Il semble en effet très pertinent de considérer que le passage d'un nœud à un autre, qui correspond à l'activation d'un lien, a un contenu sémantique, une signification. Par exemple, aller du nœud A au nœud B, si B « est un exemple de » A, n'a sans doute pas la même signification que si B « est un concept analogue à » A.

Boechler et Dawson (2002) utilisent également une mesure multidimensionnelle. Cette méthode statistique permet de découvrir les régularités structurales qui sont ensuite représentées sous forme graphique. Cette méthode utilise des mesures de proximité pour représenter les configurations de données et ensuite représente les patterns de comportements de navigation par des configurations de points dans l'espace où les coordonnées représentent les propriétés sous-jacentes. Les points sont configurés selon la proximité des objets étudiés (ex. nœuds). Par exemple, dans l'étude réalisée par les auteurs les objets étudiés étaient les pages-transition, soit le nombre de fois qu'un utilisateur se déplace d'un nœud particulier vers un autre. Les fréquences des transitions étaient utilisées comme mesure de proximité. Cela a permis par exemple d'étudier comment les différents outils de navigation (structures des hypertextes) sont utilisés par les individus.

Enfin, la méthode proposée par McEneaney (2001) calcule la centralité d'un nœud (nœuds-départ, nœuds-destination), la compacité (c'est-à-dire le degré de connectivité du réseau) et les strates (indiquent le degré de linéarité du réseau). Ces indices permettent de caractériser non seulement le réseau mais aussi les parcours des utilisateurs. McEneaney propose une représentation et une analyse de

la navigation dans les hypertextes au travers de deux formes d'analyse des parcours de navigation : la 1ère forme repose sur des métriques structurales de chemin spécifique, la 2nd sur une méthode graphique qui permet d'illustrer les parcours sous forme de patterns navigationnels visuellement distincts. McEneaney utilise le modèle traditionnel « nœuds et liens » des hypertextes qui conduit à 2 formalismes : les « matrices d'adjacence » (tables de 1 et de 0 indiquant l'existence (1) ou non (0) d'un lien entre 2 nœuds) et les « graphiques étiquetés » (représentation du réseau et de ses caractéristiques structurales). Ces deux formalismes peuvent alors servir de base pour le calcul de matrice de proximité, ou de mesures de connectivité et de linéarité.

Les mesures proposées par l'auteur reposent principalement sur la notion de centralité d'un nœud qui traduit la prééminence d'un nœud dans un réseau en tant que « destination » (de nombreux liens conduisent à ce nœud) ou « départ » (ce nœud présente de nombreux liens vers d'autres nœuds). Cette notion de centralité comprend également les aspects connectivité et directionnalité. Les deux principales mesures pour la modélisation ou la comparaison de réseaux sont la compacité et la strate. La compacité (C_p) représente le degré de connexion des nœuds d'un réseau entre eux, une compacité de 1 traduisant un réseau totalement connecté (chaque nœud est connecté à tous les autres nœuds) et une compacité de 0 un réseau déconnecté (il n'y a aucun lien). La strate (St) quant à elle renvoie à l'ordre linéaire implicite dans la structure d'un réseau. Une valeur de 1 indique une séquence linéaire stricte des nœuds et une valeur de 0 indique un réseau totalement connecté.

Ces métriques de compacité et de strate permettent la caractérisation des parcours des utilisateurs. Par exemple à partir de la succession de nœuds consultés par l'utilisateur un graphique de parcours peut par exemple être représenté. Une matrice de parcours peut également être construite. Elle indique le nombre de transitions de chaque nœud vers tous les autres nœuds dans un parcours. Notons qu'afin de comparer des parcours de différents utilisateurs (compacité et strate), il est nécessaire de normaliser les mesures. Ainsi selon les objectifs de l'expérimentateur, la normalisation se fera plutôt sur la base de la structure de l'hypertexte si l'objectif est d'étudier l'impact d'une structure d'un hypermédia sur les parcours des utilisateurs, par contre elle se fera sur la base du parcours relevé (nombre de nœuds

représentés dans la matrice de parcours) si l'objectif est de comparer des parcours différents dans un même hypermédia.

3.2. LES GRAMMAIRES DE PARCOURS

Dans la littérature abordée ci-dessous on trouve des études qui considèrent le document comme un ensemble indifférencié de nœuds, et ne prennent pas en compte les types de nœuds existants. D'autres distinguent par exemple les fonctionnalités des nœuds de contenu.

Puisque la mesure de l'efficacité ou de l'efficience des parcours implique une connaissance précise de la signification de l'efficacité des parcours (connaissance que nous n'avons pas), on peut vouloir simplement décrire les parcours en fonction de questions que l'on se pose. Tricot et Coste (1995) se sont demandés quelles relations les utilisateurs établissent entre les contenus et les fonctionnalités d'accès à ces contenus (les menus). Ils se sont aussi demandés comment les sujets passent d'un domaine de contenu à l'autre, la tâche prescrite impliquant ces passages. Ils ont donc distingué les nœuds de contenu et les nœuds de menu. Les nœuds de menu donnent des informations sur les relations entre les nœuds de contenus, à un niveau global (organisation générale du système) ou local (relation entre deux nœuds). Les nœuds de contenu peuvent faire partie d'un même thème ou non (il n'y a aucune restriction concernant l'ajout de catégories de nœuds différentes). Ils notent :

- α_i un nœud de menu
- a_i, a_j deux nœuds de contenu d'un même thème
- a_i, b_i deux nœuds de contenu de thèmes différents

Dans leur notation, il y a donc quatre types de déplacements possibles entre deux nœuds. Chacun des quatre déplacements possibles s'écrit dans les deux sens :

- $(a \ b) = - (b \ a)$:
- L (largeur) = $a_i \ b_i$
- O (orientation) = $\alpha_i \ \alpha_j$
- S (surface) = $\alpha_i \ a_i$
- T (thème) = $a_i \ a_j$

Nous définissons une séquence $[x \ y(n)]$ comme une suite de déplacements $(x \ y)$ comprenant n éléments. Par exemple, $[\alpha_i \ a_i \ \alpha_j \ a_j]$ s'écrit $[\alpha \ a \ (4)]$ soit $[S4 ; \text{un parcours de type S et de longueur 4}]$

Règle : si deux écritures d'une même séquence sont possibles, alors choisir celle où (n) est le plus grand.

Ce formalisme a par exemple permis de montrer que lors de la phase de découverte d'un document complexe, les utilisateurs ont tendance à faire des parcours de type S (une alternance de consultation de nœuds de menu et de nœuds de contenu), et non pas, comme on aurait pu s'y attendre, d'abord des parcours O (l'utilisateur apprend à se servir de l'outils) puis des parcours L ou T (l'utilisateur se sert de l'outil : il traite des contenus).

3.3. PERTINENCE PSYCHOLOGIQUE DE CES ANALYSES

Si les différents formalismes et mesures présentés dans cette seconde partie semblent intéressants, on remarque qu'ils ne procèdent pas à partir d'un véritable modèle de parcours prédit par un modèle du fonctionnement psychologique. L'interprétation psychologique de telles mesures est donc délicate, le plus souvent *ad hoc*. On pourrait envisager par exemple de fonctionner selon une approche rationnelle comme cela se fait souvent en psychologie, notamment dans le domaine de la résolution de problème, et mesurer la distance entre un parcours observé et un parcours optimal. Il n'a pas échappé au lecteur en effet, que les mesures utilisées dans le domaine de l'analyse des parcours documentaires ressemblent fort aux mesures utilisées dans l'analyse de la résolution de problème. Les mesures définies dans le tableau 10.1 ne sont rien d'autre que les mesures issues de la résolution de problème considérées comme un parcours dans un espace de recherche constitué d'états et de transformation d'états, c'est-à-dire un automate d'états (Newell & Simon, 1972) et que les états sont appelés ici des noeuds et qu'un noeud ouvert est un état visité (par exemple Zanga, Richard & Tijus, 2004). Ainsi, l'analyse de l'activité avec un document électronique, dont on peut décrire la structure comme un ensemble de nœuds reliés par des liens, et celle de résolution de problème (comme celui de la Tour de Hanoï) peuvent être réalisées au sein du même cadre. Cette analogie est peut-être plus profonde encore, comme le suggèrent Anderson et Milson (1989), qui soulignent que les mesures utilisées en documentation peuvent aussi être utilisées dans le domaine de l'étude du rappel volontaire en mémoire. Il serait d'ailleurs intéressant de retracer l'histoire parallèle de ces mesures dans les deux disciplines, la psychologie cognitive et

les sciences du document. L'analogie que nous venons de souligner existait déjà dans les années 1960. Il est aussi important de noter que les grammaires d'action en IHM, qui ne sont rien d'autre que des méthodes d'analyse de la résolution de problèmes transférées au domaine des IHM, ont intégré une dimension sémantique (Poitrenaud, 1995) il y a plus de dix ans ; cette dimension sémantique a pu être reprise dans l'analyse de la résolution de problèmes (Richard, Clément & Tijus, 2002) ainsi que, comme nous l'avons évoqué plus, dans l'analyse de l'utilisation de documents électroniques (Berendt, 2002).

4. DISCUSSION

L'analyse des parcours permet de mieux comprendre la performance mesurée *a posteriori* (apprentissage - compréhension, recherche d'information, résolution de problème, ou autre). Le tableau 10.4 récapitule les principales mesures utilisées dans le domaine.

Taux de rappel	mesure le fait que l'utilisateur sélectionne le plus possible de documents pertinents
Taux de précision	mesure le fait que l'utilisateur sélectionne le moins possible de documents non pertinents, voire de documents en général
Economie	mesure le fait que les documents sélectionnés par l'utilisateur sont différents (<i>i.e.</i> qu'il n'ouvre pas plusieurs fois les mêmes)
Redondance*	mesure le fait que l'utilisateur sélectionne plusieurs fois certains documents
Temps	mesure le temps passé à sélectionner des documents pertinents, rapporté au temps total passé à réaliser l'activité ; (parfois cette seconde mesure est utilisée seule)*
Désorientation*	mesure la distance entre le parcours réalisé et le parcours optimal ; cette mesure peut être pondérée par le poids des liens activés
Efficienc	mesure la réussite de la tâche principale (par exemple, apprentissage) rapportée au coût de l'activité de navigation

Tableau 10.4. – *Récapitulation des principales mesures utilisées dans l'analyse des parcours. Les * indiquent que la mesure est inverse : plus la valeur est grande, moins la performance est bonne.*

Les mesures on-line prennent sens à partir du moment où elles sont abordées au travers d'un modèle psychologique. Mais une confrontation des données on-line aux données en sortie permet également d'obtenir des réponses. Par exemple Potelle et Rouet (2003) montrent qu'il existe une relation forte pour des novices du domaine, entre l'ordre de rappel des idées principales d'un document hypertexte dans un résumé et l'ordre de consultation des nœuds lors de l'activité de lecture. Salmerón, Cañas, Kintsch et Fajardo (2005) montrent que le niveau de cohérence (coréférentielle) des parcours entre les nœuds d'un hypertexte prédisait des performances de compréhension différentes selon que le lecteur possédait ou non des connaissances antérieures dans le domaine. C'est dans ce sens qu'il faut aller, nous semble-t-il.

Il nous semble pour finir que l'approche rationnelle, telle que définie par Anderson et al. (2004) et employée dans le domaine de la recherche d'information par Pirolli et Card (1999) ou par Fu et Gray (2006), fournit un cadre suffisamment précis pour prédire des parcours optimaux plausibles et s'en servir de base de comparaison avec des parcours observés.

5. BIBLIOGRAPHIE

- Ahuja, J.S., & Webster, J. (2001). Perceived disorientation : an examination of a new measure to assess web design effectiveness. *Interacting with Computers*, 14, 15-19.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y . (2004). An integrated theory of the mind. *Psychological Review*, 111, 1036-1060.
- Anderson, J.R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, 96, 703-719.
- Bera, S., & Liu, M. (2004). Cognitive tools, individual differences, and group processing as mediating factors in a hypermedia environment. *Computers in Human Behavior*, 22, 295-319.

- Berendt, B. & Brenstein, E. (2001). Visualizing Individual Differences in Web Navigation: STRATDYN, a Tool for Analyzing Navigation Patterns. *Behavior Research Methods, Instruments, & Computers*, 33, 243-257.
- Berendt, B. (2002). Using Site Semantics to Analyze, Visualize, and Support Navigation. *Data Mining and Knowledge Discovery*, 6, 37-59
- Bernstein, M. Joyce, M., & Levine, D., (1992). Contours of constructive hypertexts. *ECHT Proceedings* (pp. 161-170). New York : ACM Press.
- Boechler, P.M., & Dawson, M.R.W. (2002). Effects of navigation tool information on hypertext navigation behavior: A configurational analysis of page-transition data. *Journal of Educational Multimedia and Hypermedia*, 11, 95-115.
- Brinkerhoff, J.D., Klein, J.D., & Koroghlanian, C.M. (2001). Effects of overviews and computer experience on learning from hypertext. *Journal of Educational Computing Research*, 25, 427-440.
- Buckland, M., & Gey, F. (1994). The relationship between recall and precision. *Journal of the American Society for Information Science*, 45,12-19.
- Case, D. (2002). *Looking for information : A survey of research on information seeking, needs, and behavior*. San Diego : Academic Press.
- Danielson, D.R. (2002). Web navigation and the behavioral effects of constantly visible site maps. *Interacting with Computers*, 14, 601-618
- Dee-Lucas, D., & Larkin, J.H. (1995). Learning from electronic texts: Effects of interactive overviews for information access. *Cognition & Instruction*, 13, 431-468.
- Dias, P., & Sousa, A.P. (1997). Understanding navigation and disorientation in hypermedia learning environments. *Journal of Educational Multimedia and Hypermedia*, 6, 173-185.
- Downing, R.E., Moore, J.L., & Brown, S.W. (2005). The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in Human Behavior*, 21, 195-209.
- Fu, W.-T., & Gray, W.D. (2006). Suboptimal tradeoffs in information seeking. *Cognitive Psychology*, 52, 195-242.
- Herder, E. (2002). Metrics for the Adaptation of Site Structure. Paper presented at the *German Workshop on Adaptivity and User Modeling in Interactive Systems ABIS02*, Hannover, Germany, 9-11 October.
- Kieras, D., & Polson, P.G. (1985). An approach to the formal analysis of user complexity. *International Journal of Man Machine Studies*, 22, 365-394
- Lawless, K.A., & Brown, S.W. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional Science*, 25, 117-131.
- Lee, M.J., & Tedder, M.C. (2003). The effects of three different computer texts on readers' recall: based on working memory capacity. *Computers in Human Behavior*, 19, 767-783.

- Lin, D.-Y.M. (2003). Hypertext for the aged: Effects of text topologies. *Computers in Human Behavior, 19*, 201-209.
- McDonald, S., & Stevenson, R.J. (1998). Effects of Text Structure and Prior Knowledge of the Learner on Navigation in Hypertext. *Human Factors, 40*, 18-27.
- McEneaney, J.E. (2001). Graphic and numerical methods to assess navigation in hypertext. *International Journal of Human-Computer Studies, 56*, 761-786.
- Müller-Kalthoff, T., & Möller, J. (2003). The effects of graphical overviews, prior knowledge, and self-concept on hypertext disorientation and learning achievement. *Journal of Educational Multimedia and Hypermedia, 12*, 117-134.
- Newell, A., & Simon, H. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nilsson, R., & Mayer, R.E. (2002). The effects of graphic organizers giving cues to the structure of a hypertext document on users' navigation strategies and performance. *International Journal of Human-Computer Studies, 57*, 1-26.
- Otter, M., & Johnson, H. (2000). Lost in hyperspace: metrics and mental models. *Interacting with Computers, 13*, 1-40.
- Paas, F., & van Merriënboer, J. (1993). The efficiency of instructional conditions: an approach to combine mental effort and performance measures. *Human Factors, 35*, 737-743.
- Padovani, S., & Lansdale, M. (2003). Balancing search and retrieval in hypertext: context-specific trade-offs in navigational tool use. *International Journal of Human-Computer Studies, 58*, 125-149.
- Payne, S.J., & Green, T.R.G., (1986). Task-Action Grammars: a model of the mental representation of task languages. *Human-Computer Interaction, 2*, 93-133.
- Pirolli, P. & Card, S. (1999). Information foraging. *Psychological Review, 106*, 643-675.
- Potelle, H., & Rouet, J.-F. (2003). Effects of content representation and readers' prior knowledge on the comprehension of hypertext. *International Journal of Human Computer-Studies, 58*, 327-345.
- Poitrenaud, S. (1995). The PROCOPE semantic network: An alternative to action grammars. *International Journal of Human-Computer Studies, 42*, 31-69
- Richard J.-F., Clément, E., & Tijus C.A. (2002). Les différences de difficulté dans la résolution des problèmes isomorphes comme révélatrices des composantes sémantiques dans la construction de la représentation du problème. In J. Pitrat (Ed.), *Homage à H. Simon. Revue d'Intelligence Artificielle, 16*, 191-219
- Rouet, J.-F. (2003). What was I looking for? The influence of task specificity and prior knowledge on students' search strategies in hypertext. *Interacting with Computers, 15*, 409-428.

- Salampasis, M., Tait, J., & Bloor, C. (1998). Evaluation of information-seeking performance in hypermedia digital libraries. *Interacting with Computers*, 10, 269-284.
- Salmerón, L., Cañas, J.J., Kintsch, W., & Fajardo, I. (2005). Reading strategies and hypertext comprehension. *Discourse Processes*, 40(3), 171-191.
- Smith, P. A. (1996). Towards a practical measure of hypertext usability. *Interacting with Computers*, 8, 365-381.
- Stanton, N., Correia, A.P., & Dias P. (2000). Efficacy of a map on search, orientation and access behaviour in a hypermedia system. *Computers & Education*, 35, 263-279.
- Tricot, A., & Coste, J.-P. (1995). Evaluating complex learner-computer interaction : What criteria for what task ? *EARLI'95 Conference*, Nijmegen, August 26-31.
- Tricot, A., Puigserver, E., Berdugo, D., & Diallo, M. (1999). The validity of rational criteria for the interpretation of user-hypertext interaction. *Interacting with Computers*, 12, 23-36.
- Zanga, A., Richard, J.-F., & Tijus, C.A. (2004). Implicit learning in rule induction and problem solving. *Thinking & Reasoning*, 10, 55-83