

---

# Une approche fondée sur le raisonnement à partir de cas pour l'aide à la navigation sur le Web

Brigitte Trousse — Michel Jaczynski<sup>1</sup> — Rushed Kanawati

INRIA Sophia Antipolis, Action AID  
2004, route des Lucioles BP938  
06902 Sophia Antipolis  
trousse@sophia.inria.fr

---

*RÉSUMÉ.* Après une introduction à la recherche d'informations sur le Web et une classification des assistants d'aide à la navigation sur le Web, cet article décrit une nouvelle approche pour l'aide à la navigation sur le Web centrée fouille de données nommé Broadway. L'approche Broadway repose sur l'utilisation des techniques de raisonnement à partir de cas (RàPC) afin de recommander à un utilisateur des pages qui ont satisfait des utilisateurs ayant navigué d'une manière similaire. Un exemple de l'implantation de l'approche Broadway est décrit. Enfin, avant de conclure, une comparaison avec d'autres approches est donnée.

*ABSTRACT.* After an introduction on web information retrieval approaches and a classification of browsing assistants on the Web, this article describes a new data mining oriented approach, called Broadway, for supporting hypermedia browsing. The Broadway approach relies on using case-based reasoning techniques (CBR) in order to recommend to a user pages that have satisfied users having a similar behaviour. An example of how implementing the Broadway approach is given. Then, before concluding, a comparison with others approaches is given.

*MOTS-CLÉS :* Recherche d'information, Web, Aide à la navigation, Comportement utilisateur, Raisonnement à partir de cas.

*KEYWORDS :* Information retrieval, Web, Browsing support, User behaviour, Case-based reasoning,

---

## 1. Introduction : Recherche d'information sur le Web

Le World Wide Web (ou le Web) est un gigantesque hypermédia sans cesse croissant. La fraîcheur d'informations disponibles sur le Web et la disponibilité d'outils conviviaux d'accès à ces informations (i.e. NetScape, Internet Explorer) font

---

<sup>1</sup> M. Jaczynski est actuellement consultant chez Ilog SA, Gentilly.

que de plus en plus d'utilisateurs considèrent le Web comme la première source d'information. Une variété d'approches de recherche et de localisation d'information sur le Web sont proposées dans la littérature. Nous proposons les quatre critères suivants pour classer les systèmes de recherche d'information (RI) sur le Web :

— *La nature de la tâche prise en compte.* Selon le degré de la précision du but poursuivi et des connaissances de l'utilisateur sur les informations disponibles dans l'espace qui répondent à ce but, on peut identifier les classes de tâches suivantes : la localisation, la recherche, l'exploration, et la veille [TRI 97].

— *La métaphore de recherche.* Ceci réfère au type des formats d'entrées qu'un système de RI accepte de l'utilisateur dans le but de trouver l'information recherchée. Quatre métaphores de recherche sont identifiées : la navigation, la consultation d'un répertoire thématique (ex Yahoo<sup>2</sup>), la formulation de requêtes et la recherche grâce à l'adhésion à une communauté d'intérêts.

— *Le mécanisme de recherche.* Ceci réfère à la procédure sous-jacente utilisée par le système de RI correspondant à une métaphore de recherche adoptée par le système. Quatre principaux mécanismes de recherche d'information sont identifiés :

- *L'interrogation* : il s'agit de chercher les documents dans une base documentaire (i.e. suivant un index),

- *La navigation* : il s'agit d'envoyer des robots sur Internet pour retrouver des documents qui correspondent à la requête,

- *La recherche sociale* : il s'agit de localiser des personnes qui savent où trouver l'information sur les documents pertinents à la requête,

- *L'approche hybride* : il s'agit de déléguer la tâche de RI à un ou plusieurs systèmes de recherche qui peuvent utiliser des mécanismes différents. Les résultats sont agrégés avant de les retourner à l'utilisateur.

— Un quatrième critère de classification des systèmes RI est l'approche utilisée pour améliorer *l'efficacité de la recherche* (« *effectiveness* » en anglais) employée par le système. Il ne s'agit pas ici d'efficacité au sens rapidité de la recherche. Différentes mesures de l'efficacité sont proposées, dont les plus utilisées sont principalement les critères de *précision* et de *rappel (recall)* [KOR 97]

Dans cet article, nous nous intéressons aux systèmes fondés sur le mécanisme de navigation. Le mécanisme de recherche d'information par navigation consiste à se déplacer d'une page Web à une autre, par décision à chaque étape du lien à suivre parmi un ensemble de liens possibles. Deux grandes approches complémentaires sont proposées dans la littérature pour construire des systèmes d'aide au choix de liens et plus généralement pour le calcul de recommandation [RES 97] : 1) l'approche fondée sur l'utilisation de *profils utilisateurs* et 2) l'approche fondée sur des techniques de *fouille de données*. Le profil utilisateur est une structure de données qui décrit les centres d'intérêts de l'utilisateur. Une fois une telle structure construite, on peut l'utiliser soit pour filtrer les liens disponibles (on parle alors de filtrage basé sur le contenu) [SCE 97 ; LIE 95], soit pour recommander à l'utilisateur ce qui a satisfait d'autres utilisateurs ayant un profil similaire (on parle alors de filtrage

---

<sup>2</sup> <http://www.yahoo.com>

collaboratif). Les approches centrées fouille de données reposent sur l'analyse des activités de navigation sur le Web d'un utilisateur ou d'un groupe d'utilisateurs afin d'extraire de *recommandations* pour l'aide à la navigation sur le Web [LIE 95 ; Yan 96 ; WEX 97]. Une approche hybride peut reposer sur l'utilisation des techniques de fouille de données pour la construction d'un profil utilisateur [MLA 96].

Dans cet article, nous présentons une nouvelle approche pour l'aide à la navigation sur le Web centrée fouille de données nommée Broadway. L'approche Broadway repose sur l'utilisation de techniques de raisonnement à partir de cas (RàPC) afin de recommander à un utilisateur des pages qui ont satisfait des utilisateurs ayant navigué d'une manière similaire. A la différence de la plupart des autres approches centrées fouille de données, l'approche Broadway prend en compte *l'ordre des actions* des utilisateurs comme un élément central dans la calcul de recommandations.

Le reste de l'article est organisé comme suit. Dans la section 2, nous proposons une classification fonctionnelle des différentes approches d'aide à la navigation proposées dans la littérature. Ensuite, dans la section 3, nous présentons l'approche Broadway. Un exemple d'implantation de l'approche Broadway dans un système destiné à aider un groupe d'utilisateurs à partager leurs expériences de navigation sur le Web [JAC 98a] est présenté dans la section 4. Une étude comparative de notre approche avec les approches existantes est faite dans la section 5.

## 2. Les assistants à la navigation sur le Web

Un nombre croissant de systèmes d'aide à la navigation dans des hypermédia (ex. le Web) sont proposés dans la littérature. Nous citons pour exemples : Alexa (<http://www.alex.com>), BroadwayV1 [JAC 98a], FootPrints [WEX 97], Hospitext [ELK 97], Hypercase [MIC 96], Letizia [LIE 95], Personal WebWatcher (PWW) [MLA 96], WBI [BAR 97], WebWatcher [ARM 97], et le système de Yan et. Al [YAN 96]. Ces assistants offrent différents types d'aide à la navigation et utilisent différents mécanismes de calcul de recommandations. Pour faciliter la comparaison entre les différents assistants à la navigation nous proposons dans la suite une classification fonctionnelle qui s'appuie sur les critères suivants :

— *La nature de l'assistance fournie.* Différents types d'aide peuvent être proposés par les assistants à la navigation [BRU 97; PER 97] notamment les types suivants : 1) cacher des liens existants, 2) trier les liens en fonction de leur pertinence au contexte de navigation courant, 3) proposer des raccourcis et 4) annoter les liens.

— *La personnalisation.* Il s'agit de la capacité du système à personnaliser l'aide à fournir en fonction des besoins individuels de chacun de ces utilisateurs.

— *L'adaptabilité.* Il s'agit de la capacité du système à apprendre à partir des interactions de ses utilisateurs.

Le tableau 1 montre les caractéristiques de quelques assistants à la navigation. Dans la colonne décrivant la nature de l'aide apportée, nous utilisons les abréviations suivantes : C pour cacher des liens, T pour trier des liens, R pour offrir des raccourcis et A pour utiliser des annotations.

Assistants de navigation	AIDE (C,T,R,A)	Personnalisation	Adaptabilité
Alexa	(-, -, R, -)	-	✓
BroadwayV1	(-, -, R, -)	✓	✓
Footprints	(C, -, R, A)	-	✓
Hospitext	(C, -, -, -)	✓	-
Hypercase	(C, -, -, -)	-	✓
Letizia	(-, -, R, -)	✓	✓
PWW	(-, -, R, -)	✓	✓
Walden Path	(C, -, -, -)	-	-
WBI	(-, -, R, A)	✓	✓
WebWatcher	(-, T, R, -)	✓	✓
Yan et al	(-, -, R, -)	✓	✓

**Tableau 1.** Caractérisation des principaux assistants à la navigation

### 3. L'approche Broadway

Cette section vise à décrire tout d'abord notre approche de calcul de recommandations basé sur les techniques de raisonnement à partir de puis à illustrer l'application de notre approche pour l'aide à la navigation sur le Web via le système Broadway V1 [JAC 98a].

#### 3.1. Description générale

L'approche Broadway de calcul de recommandations s'appuie sur un filtrage basé sur la réutilisation de comportements de l'utilisateur et/ou d'un groupe d'utilisateurs. Ce filtrage peut être classé comme un type particulier de « filtrage collaboratif » [RES 97]. En effet les assistants basés sur l'approche Broadway permettent une recherche d'information où l'expérience d'un groupe de personnes est mise à profit. La prise en compte du *comportement des utilisateurs*, constitué notamment de la *séquence des pages visitées* et de leur contenu, doit également permettre de mieux cerner l'intérêt de l'utilisateur pour affiner les recommandations.

##### 3.1.1. Raisonnement à partir de cas

De manière générale, la réutilisation de comportements utilisateur s'appuie sur le raisonnement à partir de cas qui est une approche de résolution de problèmes basée sur la réutilisation par analogie d'expériences passées appelées *cas* au cours d'un cycle de raisonnement [KOL 96]. Un cas est composé de deux parties décrivant un *problème* et la *solution* qui lui a été appliquée. Un cas est une expérience passée permettant au système de résoudre des problèmes plus efficacement ou d'éviter les échecs passés : un cas doit donc être *utile* pour le raisonnement. Dans ce but, un ensemble de caractéristiques est extrait ou sélectionné dans la représentation de chaque cas. Ces caractéristiques, appelées *indices*, déterminent dans quelle *situation* un cas est applicable et utile. Les indices doivent tout d'abord être prédictifs, c'est-à-dire qu'ils doivent être responsables de l'élaboration de la solution mise en œuvre dans le cas et permettent ainsi de déterminer dans quelle mesure une solution est applicable pour un autre problème. Les indices doivent également être suffisamment

abstraites pour permettre la réutilisation des enseignements du cas pour différents problèmes, tout en restant suffisamment concrets pour être facilement identifiables dans un nouveau problème.

Lorsqu'un *nouveau problème* est à résoudre, un *cas cible* est construit avec la partie solution inconnue, celle-ci devant être apportée par le raisonnement. Les *cas sources* représentent des expériences passées qui sont stockées dans une *mémoire*. Les cas sources et le cas cible partagent le même formalisme de représentation. L'objectif du raisonnement est de transférer des enseignements pertinents des cas sources pour élaborer la solution du cas cible. Ce raisonnement se décompose habituellement en quatre phases principales [AAM 94] qui peuvent être totalement ou partiellement automatisées (cf. Figure 1).

### 3.1.2. Utilisation du RàPC pour le calcul de recommandations

L'approche Broadway repose principalement sur la réalisation d'un serveur de recommandations qui réutilise des cas principalement composés 1) de la description du comportement d'un utilisateur à un instant de référence dans sa navigation et 2) d'une liste de pages recommandées pour ce comportement. Alors que les sessions de navigation représentent les données brutes d'observation, les cas permettent de manipuler des associations pertinentes issues des navigations entre un comportement à un instant de référence et des pages évaluées :

*Cas = (comportement utilisateur issu d'une navigation à un instant de référence, liste de pages évaluées et recommandées)*

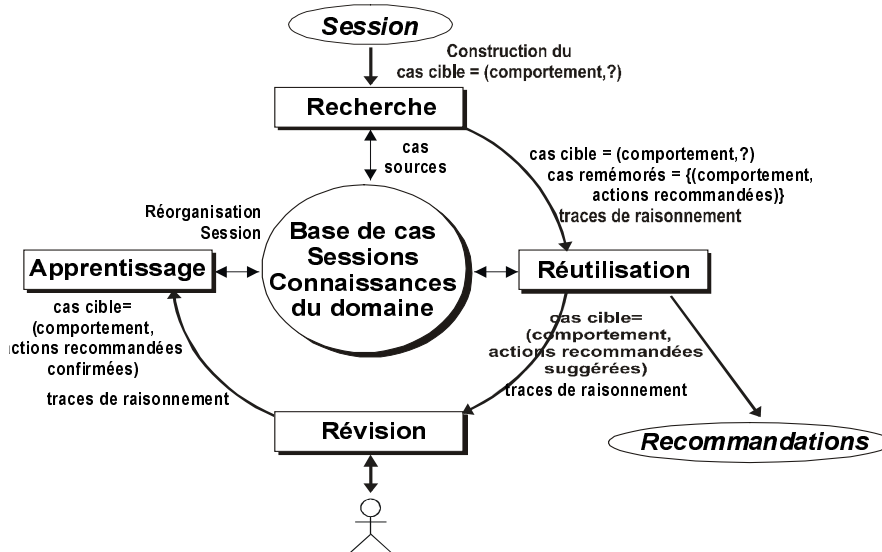


Figure 1. Cycle du raisonnement

Le calcul des recommandations est alors effectué durant un raisonnement en quatre phases, étant donné la navigation courante d'un utilisateur :

— *Recherche*. La base de cas est parcourue pour identifier les cas dont la partie comportement peut être mise en correspondance avec la navigation courante. Les meilleurs cas sont alors retournés.

— *Réutilisation*. Chaque cas retrouvé propose une liste de pages recommandées, et la phase de réutilisation opère la fusion de ces pages puis ordonnance à l'aide de divers critères les pages recommandées pour constituer la liste des *pages recommandées suggérée* à présenter à l'utilisateur courant.

— *Révision*. La phase de révision est effectuée par l'utilisateur lorsqu'il continue sa navigation et qu'il évalue les pages visitées s'il le désire. La révision se poursuit jusqu'à la fermeture de la navigation courante et met en attente tous les raisonnements effectués pour une navigation (pages recommandées confirmées).

— *Apprentissage et maintenance de la mémoire*. L'apprentissage prend en compte la navigation courante et l'ensemble des raisonnements associés pour en faire leur synthèse. La mémoire est alors mise à jour par la création de nouveaux cas, par la mise à jour des cas et par l'ajout de la navigation courante. A côté de ce processus de raisonnement, les cas et les navigations devenus obsolètes sont régulièrement effacés de la mémoire.

### 3.2. Mise en œuvre de l'approche Broadway

La mise en œuvre de l'approche Broadway nécessite la manipulation des cas dont les indices sont constitués de *séquences d'événements*. Pour cela, nous nous appuyons pour l'implantation de cette approche sur le « modèle d'indexation par situations comportementales » [JAC 98a] qui est conçue spécialement pour permettre la manipulation de ce type de cas. Le modèle est implanté dans une plateforme objet nommé CBR\*Tools [JAC 99] (conçue avec l'atelier Rose de Rational et implantée en Java) afin de faciliter le développement des systèmes implantant l'approche Broadway. Pour l'utilisation de ce modèle, nous avons identifié quatre étapes importantes :

1. Identification des variables d'observation.
2. Détermination de la sémantique d'un enregistrement (navigation) et de son contexte.
3. Définition de la représentation des cas et des situations comportementales (éléments issus d'une navigation pour caractériser un cas).
4. Conception des phases de raisonnement.

Le modèle d'indexation proposé repose sur l'utilisation de deux types de cas : *concrets* et *potentiels*. Un cas concret référence un instant précis dans une navigation et se compose principalement d'un couple (*problème, solution*). Le problème donne notamment la situation comportementale et la solution représente une liste de pages évaluées relativement à cette situation. De plus, des informations de gestion sont également ajoutées (date de création, nom de l'utilisateur). Les cas concrets sont issus de l'instanciation d'un patron de cas potentiels. Ce patron repose sur des règles de construction d'une situation comportementale et d'une liste de page évaluées pour un instant donné à l'intérieur d'une navigation. : selon ces règles, la situation comportementale est alors composée d'une partie instantanée contenant le contexte de la navigation (par exemple, noms des sites et mots clefs) et d'une partie comportementale référençant un instant précis dans la navigation. Cette dernière permet de sélectionner les indices comportementaux d'un cas par rapport à l'instant de référence et peut comprendre :

— les  $p$  dernières pages sur les variables de l'adresse et du contenu pour identifier le contexte immédiat de réutilisation d'un cas dans une navigation (notion de restriction),

— un ensemble de pages passées qui sont sélectionnées pour leur pertinence: soit pour une évaluation élément d'une liste  $e_p$ , soit pour un ratio d'affichage dépassant le seuil  $a_p$ ,

— et un ensemble de contraintes temporelles permettant d'ordonner les pages sélectionnées.

Ainsi, les conditions de réutilisation d'un cas sont manipulées plus facilement à travers la restriction (structure figée) et peuvent être précisées avec la liste des pages passées sélectionnées et ordonnées. La partie solution du cas qui donne une liste de pages évaluées est également construite en sélectionnant les pages pertinentes futures par rapport à l'instant de référence de manière similaire à la sélection des pages passées (évaluations éléments de  $e_p$  et seuil  $a_p$ ).

#### **4. Exemple d'utilisation**

Deux principaux assistants d'aide à la navigation basés sur l'approche Broadway ont été réalisés : Broadway V1 pour l'aide à la navigation sur le Web et Broadway V2 dans le cadre d'un contrat CTI-CNET pour l'aide à la navigation dans un site Web donnée. Nous illustrerons ici l'approche Broadway par le système Broadway V1 et qui a donné lieu à une validation expérimentale menée par des étudiants en Psychologie & Ergonomie Cognitive [HEB 98]. Nous invitons ceux qui souhaiteraient avoir plus de détails sur l'application de notre approche pour le système Broadway V1 à lire l'article [JAC 98b] et le chapitre 5 du document de thèse [JAC 98a]. L'approche proposée s'appuie sur l'application de notre modèle d'indexation par situation comportementale qui s'appuie sur les quatre étapes suivantes:

##### **4.1. Identification des variables d'observation.**

L'observation du comportement des utilisateurs est fondée sur l'utilisation de quatre variables représentant l'adresse, le contenu, l'évaluation et le ratio d'affichage des pages Web visitées. Ces variables qui sont notées respectivement Adresse, Contenu, Evaluation et Affichage, prennent des valeurs successives au cours du processus de recherche d'information pour chaque utilisateur.

##### **4.2. Navigation et contexte.**

La figure 2 donne l'exemple d'une navigation enregistrée par BroadwayV1 dans laquelle un utilisateur désire retrouver des informations concernant Broadway lui-même. Il commence alors sa navigation sur la page d'accueil de l'INRIA, passe par les pages de l'équipe de recherche AID et finalement trouve la page relatant les développements de logiciels et notamment ceux de Broadway.

**Navigation #16 from mjaczyn**

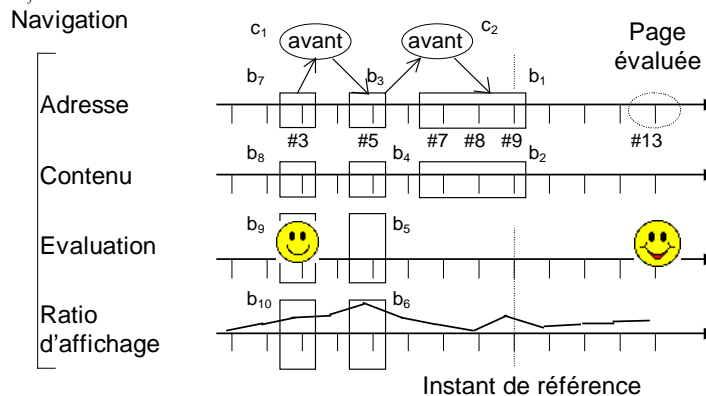
start date: Wed Nov 26 19:28:20 GMT+03:30 1997  
end date: Wed Nov 26 19:30:51 GMT+03:30 1997  
keywords (stems): PEOPLE, REPORT, INRIA, PUBLIC, CONFER, ACTIV, STAGE, MICHEL, THEME, SITE, SOFTWAR, DEADLIN, JACZYNSKI, ACTION, RESEARCH, PREFER  
hosts: www.inria.fr

#	Page address	Content (stems)	Display	Evaluation
0	<a href="http://www.inria.fr/Recherche/activites-eng.html">http://www.inria.fr/Recherche/activites-eng.html</a>	INRIA, ACTIV, RESEARCH	1.0	0
1	<a href="http://www.inria.fr/Themes/Theme3-eng.html">http://www.inria.fr/Themes/Theme3-eng.html</a>	INRIA, THEME	1.0	0
2	<a href="http://www.inria.fr/Equipes/AID-eng.html">http://www.inria.fr/Equipes/AID-eng.html</a>	ACTION	3.0	0
3	<a href="http://www.inria.fr/aid/aid-eng.html">http://www.inria.fr/aid/aid-eng.html</a>	ACTION	6.0	😊
4	<a href="http://www.inria.fr/aid/people.html">http://www.inria.fr/aid/people.html</a>	PEOPLE, ACTION	8.0	0
5	<a href="http://www.inria.fr/aid/personnel/Michel.Jaczynski/michel-eng.html">http://www.inria.fr/aid/personnel/Michel.Jaczynski/michel-eng.html</a>	MICHEL, SITE, JACZYNSKI, PREFER	38.0	0
6	<a href="http://www.inria.fr/aid/personnel/Michel.Jaczynski/pub-eng.html">http://www.inria.fr/aid/personnel/Michel.Jaczynski/pub-eng.html</a>	REPORT, PUBLIC, MICHEL, JACZYNSKI	10.0	0
7	<a href="http://www.inria.fr/aid/personnel/Michel.Jaczynski/michel-eng.html">http://www.inria.fr/aid/personnel/Michel.Jaczynski/michel-eng.html</a>	MICHEL, SITE, JACZYNSKI, PREFER	1.0	0
8	<a href="http://www.inria.fr/aid/people.html">http://www.inria.fr/aid/people.html</a>	PEOPLE, ACTION	4.0	0
9	<a href="http://www.inria.fr/aid/aid-eng.html">http://www.inria.fr/aid/aid-eng.html</a>	ACTION	8.0	0
10	<a href="http://www.inria.fr/aid/conf.html">http://www.inria.fr/aid/conf.html</a>	CONFER, DEADLIN, ACTION	0.0	0
11	<a href="http://www.inria.fr/aid/aid-eng.html">http://www.inria.fr/aid/aid-eng.html</a>	ACTION	2.0	0
12	<a href="http://www.inria.fr/aid/training.html">http://www.inria.fr/aid/training.html</a>	STAGE, ACTION	1.0	0
13	<a href="http://www.inria.fr/aid/software.html">http://www.inria.fr/aid/software.html</a>	SOFTWAR, ACTION	2.0	😊

**Figure 2.** Extrait d'une navigation

#### 4.3. Représentation des cas.

La figure 3 donne l'exemple d'un cas concret créé à partir de l'instanciation d'un patron de cas potentiel dans la navigation de la figure précédente (cf. Figure 1) pour un instant correspondant à la page #9. La restriction (i.e. l'ensemble d'indices utilise dans un premier filtrage des navigations) permet de sélectionner les trois dernières pages ( $p=3$ ) ainsi que la page #3 pour son évaluation ( $e_p=\{2\}$ ) et la page #5 pour le ratio d'affichage ( $a_p=20$ ). A cette situation comportementale, la page #13 est associée ( $e_f=\{2\}$ ) et représente la seule recommandation de ce cas.



**Figure 3.** Exemple de cas concret

#### 4.4. Conception des phases de raisonnement

Il s'agit ici de la spécification des quatre phases du RàPC. Nous invitons le lecteur intéressé à lire les deux références suivantes [JAC 98a ; JAC98b].



## 5. Comparaison avec d'autres approches d'aide à la navigation

Nous comparons, dans cette section, l'approche Broadway avec les autres approches d'aide à la navigation fondées sur le fouille de données. Pour cela, nous nous comparons aux principaux systèmes existants décrits complètement dans la littérature et structurons notre étude comparative selon les axes suivants :

### 5.1. Le type de la situation de recommandation.

Nous appelons *situation de recommandation* l'ensemble d'informations relatif à la navigation d'un agent qui sont exploitées par l'assistant afin de fournir des recommandations. Trois types de situation de recommandation construite sont identifiés :

— *Situation locale* : où seuls des informations relatives à la page courante de l'agent sont prises en compte. C'est le cas par exemple de FootPrints qui affiche l'ensemble des pages les plus visitées à partir d'une page donnée [WEX 97].

— *Situation résumée* : où un résumé de la navigation de l'utilisateur est calculé et utilisé par le système pour calculer ses recommandations. C'est le cas par exemple de Letizia [LIE 95] où le système identifie les centres d'intérêts de l'utilisateur à partir des pages qui sont déjà visitées dans la navigation courante et envoie des agents pour filtrer les pages reliées.

— *Situation comportementale* : l'ordre des pages visitées est pris en compte. C'est le cas de Hypercase mais aussi de Broadway V1 [JAC 98a] où plus généralement l'ordre des valeurs de toutes les variables observées/calculées de la composante comportementale est pris en compte. Notons que la spécification du système Radix [COR 98], propose une mesure de similarité de type événementielle entre sessions de recherche d'information.

### 5.2. La technique de fouille de données.

Différents techniques sont utilisées pour l'extraction des connaissances nécessaires au calcul de recommandation. La plupart des approches sont fondées sur des techniques de statistique (FootPrints, Alexa, Letizia). L'approche de Yan et. al. [YAN 97] consiste à appliquer une technique de catégorisation (« *clustering* ») de navigations passées ; une catégorie est assimilée à une navigation type. Une fonction d'appariement est appliquée à une navigation courante afin de déterminer la catégorie la plus proche et le système recommande les pages citées dans la catégorie identifiée et qui ne sont pas encore visitées dans la navigation courante. Une dernière technique est le raisonnement à partir de cas comme c'est le cas des systèmes Broadway V1, mais aussi de Hypercase [MIC 96].

### 5.3. La stratégie d'apprentissage.

Trois types d'apprentissage existent : le cas limite où il n'y a *pas d'apprentissage*, l'apprentissage centré sur un *unique utilisateur* et l'apprentissage à partir d'un *groupe* d'utilisateurs. Hypercase est un exemple d'assistant du premier type. Il ne peut pas apprendre ni s'enrichir à partir des expériences des utilisateurs, puisqu'il utilise uniquement un ensemble de cas prédéfinis construits par des experts. Letizia, et WBI sont des exemples d'assistant mono-utilisateur tandis que

WebWatcher, Radix, Footprints et Broadway V1 sont des assistants qui apprennent d'un groupe d'utilisateurs.

#### 5.4 Autres critères et synthèse

Nous ajoutons aux trois critères cités ci-dessus les deux critères de l'adaptabilité et de la personnalisation de l'aide fournie par l'approche comme nous les avons définis dans la section 2. Le tableau suivant résume les caractéristiques des principales approches d'aide à la navigation sur le Web.

Assistant de navigation	Situation de recommandation	Technique de fouille de données	Personnalisation	Adaptabilité	Apprentissage Origine des navigations
Alexa	Locale	Statistique	-	✓	Groupe
Broadway V1	Comportementale	RàPC	✓	✓	Groupe
FootPrints	Locale	Statistique	-	✓	Groupe
Hypercase	Comportementale	RàPC	✓	-	-
Hospitext	Locale	RàPC	-	-	-
Letizia	Résumée	Statistique	✓	✓	Mono-utilisateur
WBI	Locale	Statistique	✓	✓	Mono-utilisateur
WebWeatcher	Locale	Statistique	✓	✓	Groupe
Yan et. al.	Résumée	Catégorisation	✓	✓	Groupe

**Tableau 2.** *Quelques approches d'assistance à la navigation*

#### 5. Conclusion et travaux futurs

Nous avons présenté dans cet article l'approche Broadway pour l'aide à la navigation dans un hypermédia. Cette approche est basée sur la réutilisation de comportements observés *au niveau variable* s'appuyant sur des techniques de "raisonnement à partir de cas" et où l'ordre des actions utilisateurs est pris en compte. Les travaux en cours et futurs portent d'une part sur des validations expérimentales relatives à nos deux assistants à la navigation basés sur l'approche Broadway et d'autre part sur de nouvelles propositions au niveau des techniques de fouille de données à partir des logs et de la mémoire de sessions utilisateurs dans le but d'améliorer l'indexation et la maintenance de la mémoire (navigations, cas). Notons d'autre part que l'approche Broadway [TRO 99] est plus générale que l'aide à la navigation dans un hypermedia décrite dans cet article ; en effet cette approche adresse tout système de recommandation réutilisant les comportements utilisateurs *observés au niveau variable*, que ce soit pour l'aide à la navigation mais aussi par exemple pour l'aide à la recherche d'informations sur le Web [KAN 99].

#### 7. Références

- [AAM 94] A. Aamodt et E. Plaza. Case-Based Reasoning: Foundational Issues, Methodological Variations and Systems. AI Communications, vol. 7, n°1, p 36-59, 1994.
- [ARM 97] R. Armstrong, T. Joachims et T. Mitchell. WebWatcher : A Tour Guide for the World Wide Web. Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'97), pages 770-775, Morgan Kaufmann, 1997

- [BAR 97] R. Barrett, P. P. Maglio et D. C. Kellem. How to Personalize the Web. Proceedings of the Conference on Human-computer Interaction (CHI'97), pages 75–82, Addison-Wesley, 1997.
- [BRU 97] P. Brusilovsky, Efficient techniques for adaptive Hypermedia, In Intelligent Hypertext, advanced techniques for the world wide Web, C. Nicholas, J. Mayfield (Eds), Springer-Verlag LNCS No. 1326, pages 12-31, 1997
- [CAT 95] L.D. Catledge et J.E. Piko. Characterizing browsing strategies in the World Wide Web. Proc. of the 3rd International Conference on the World Wide Web, Darmstadt, Germany, 1995.
- [CHE 88] M-S. Chen et J. S. Park. Efficient Data Mining for Path Traversal Patterns. IEEE Transactions on Knowledge and Data Engineering, vol. 10, n°2, pages 209–221, 1988.
- [CHE 98] L. Chen et K. Sycara. WebMate : A Personal Agent for Browsing and Searching In Proc. of the 2nd Int Conference on Autonomous Agents, pages 132–139, ACM press, 1998.
- [CHO 97] J. Cho, H. Garcia-Molina and L. Page. Efficient Crawling Through URL Ordering 7th International Conference on the World Wide Web (WWW'7), Brisbane, Australia, 1997.
- [COR 98] F. Corvaisier, A. Mille et J.M. Pinon. Recherche Assistée de documents Indexés sur l'eXpérience (RADIX) : mesures de similarités des épisodes de recherche sur le Web. In Actes d'Ingénierie des Connaissances (IC'98), pages 153–163, INRIA, 1998.
- [ELK 97] S. Elkassar et J. Charlet. Représentation de connaissances et aide à la navigation hypertextuelle à partir de cas : application au dossier médical. In Actes de la journée Ingénierie des Connaissances et Apprentissage Automatique (JICAA'97), pages 387–401, INRIA, 1997.
- [HEB 98] C. Hébraud, M. Rech et A. Thébault. Test de l'utilité de l'outil d'aide à la navigation Broadway. Rapport interne, Université de Provence, UFR de psychologie, Aix en Provence, et INRIA Sophia Antipolis, Action AID, 19 pages, 1998.
- [JAC 98a] M. Jaczynski. Modèle et plate-forme à objets pour l'indexation des cas par situations comportementales : application à l'assistance à la navigation sur le Web. Thèse de Doctorat, Université de Nice, 286 pages, 1998. URL <http://aid.inria.fr/papers/thesemj/>
- [JAC 98b] M. Jaczynski et B. Trousse. WWW Assisted Browsing by Reusing Past Navigations from a Group of Users In Proceedings of 4th European Workshop on Case-Based Reasoning (EWCBR'98), volume 1488 de Lecture Notes in Artificial Intelligence, pages 160–171, Springer, 1998. <http://aid.inria.fr/papers/>
- [JAC 99] M. Jaczynski et B. Trousse. Patrons de conception dans la modélisation d'une plate-forme à objets pour le raisonnement à partir de cas. L'objet, numéro spécial « Les patrons orientés objet », vol 5, n 2, 27 pages, Hermès, 1999.

- [KAN 99] R. Kanawati, M. Jaczynski, B. Trousse et J.M. Andreoli, Applying the Broadway Recommendation Computation Approach for implementing a query refinement service in the context of the CBKB meta search engine . Plate-forme AFIA'99, Conférence de raisonnement à partir de cas (RàPC'99) - Palaiseau, Trousse & Mille (editors), pages 17-26 juin 1999.
- [KOL 96] J. L. Kolodner et D. B. Leake. A Tutorial Introduction to Case-Based Reasoning.: Experiences, Lessons & Future Directions , AAAI press, MIT press, pages 31–65, 1996.
- [KOR 97] R. R. Korfhage. Information Storage and Retrieval .. Wiley & Sons, 349 pages, 1997.
- [LIE 95] H. Lieberman. Letizia : An Agent that Assists Web Browsing . In Proceedings of International Joint Conference on Artificial Intelligence (IJCAI'95), pages 924–929, Morgan Kaufmann, 1995.
- [MIC 96] A. Micarelli et F. Sciarrone. A Case-Based System for Adaptive Hypermedia Navigation . In Advances in Case-Based Reasoning, Proc. of the 3rd European Workshop on Case-Based Reasoning (EWCBR'96), vol 1168 deLNAI, pages 266–279, Springer, 1996.
- [MLA 96] D. Mladenic. Personal WebWatcher : Design and Implementation . Rapport technique IIS-DP-7472, School of Computer Science, Carnegie Mellon University, 18 pages, 1996.
- [RES 97] P. Resnick et H. R. Varian. Recommender systems . Communications of the ACM, vol. 40, n°3, pages 56-58, 1997.
- [TRI 97] A. Tricot et J.F. Rouet Chercher des informations dans un hypertexte : vers un modèle des processus cognitifs. Hypertextes et Hypermedia, 1(3), 1997
- [TRO 99] B. Trousse, M. Jaczynski, R. Kanawati, Using User behavior Similarity for Recommendation Computation : the Broadway Approach, IN Proceedings of the 8<sup>th</sup> International Conference on human computer interaction (HCI'99), Munich, Août, 1999.
- [WEX 97] A. Wexelblat et P. Maes. Footprints : Visualizing Histories for Web Browsing . In Actes de la 5<sup>ème</sup> conférence sur la Recherche d'Information Assistée par Ordinateur sur Internet (RIAO'97), Centre des hautes études internationales d'Informatique, Montréal, pages 75–84, 1997.
- [YAN 96] T.W. Yan, M. Jacobsen, H.Garcia-Molina et U. Dayal. From User Access Patterns to Dynamic Hypertext Linking . In Proceedings of the 5th International World Wide Web Conference, volume 28 de Computer Network and ISDN Systems, pages 1007–1014, Elsevier, 1996.