



RELIEFS : un système d'inspiration cognitive pour le filtrage adaptatif de documents textuels

Christophe Brouard

► **To cite this version:**

Christophe Brouard. RELIEFS : un système d'inspiration cognitive pour le filtrage adaptatif de documents textuels. *Revue des Sciences et Technologies de l'Information, Hermes*, 2002, 7 (1/2), pp.157–182. hal-00953817

HAL Id: hal-00953817

<https://hal.inria.fr/hal-00953817>

Submitted on 28 Feb 2014

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.

RELIEFS : Un système d'inspiration cognitive pour le filtrage adaptatif de documents textuels

Christophe Brouard

*MRIM Laboratoire CLIPS-IMAG
B.P. 53
38041 Grenoble Cedex 9 France
Christophe.Brouard@imag.fr*

RÉSUMÉ. L'objet de cet article est la présentation d'un nouveau système nommé RELIEFS (pour RElevance Information Extraction Fuzzy System) pour le filtrage adaptatif de documents textuels. Les grands principes de fonctionnement de ce système s'inspirent de mécanismes cognitifs intervenant dans les processus de sélection de l'information. Plus précisément, notre recherche part de l'analyse de modèles de la mémoire sémantique (accès et organisation des connaissances en mémoire) et de modèles qui rendent compte de phénomènes attentionnels (sélection des informations provenant de l'environnement). Des liens forts sont tissés entre ces modèles et des modèles traditionnellement utilisés en RI. Une nouvelle interprétation de la notion de pertinence est proposée. L'analyse nous conduit à extraire un ensemble de mécanismes de base renvoyant aux notions d'activation et de propagation d'activation pour la sélection d'information " pertinentes ". Ces mécanismes sont implémentés et testés avec succès dans la tâche de filtrage adaptatif de TREC9.

ABSTRACT. This paper deals with the description of a new adaptive filtering system called RELIEFS (for RElevance Information Fuzzy System). The main principles of this system draw their inspiration from cognitive models of information selection. More precisely, our research is based on the analysis of semantic memory models (knowledge access and knowledge organisation) and attentional models (information selection from the environment). We also propose to bridge a gap between these cognitive models and traditional information retrieval models. A new interpretation of relevance is given. The analysis leads us to extract simple mechanisms referring to the concepts of activation and spreading activation for a system which aims at selecting "relevant" information. These mechanisms have been implemented and successfully tested in the TREC9 adaptive filtering task.

MOTS-CLÉS : filtrage adaptatif, pertinence, cognition.

KEY WORDS : adaptive filtering, relevance, cognition.

1. Introduction

Comment ne pas noter comme le souligne Barry [BARR94], le décalage important qui existe entre la pertinence d'un document calculée par un système traditionnel de recherche d'information et son mode d'évaluation basé sur un jugement humain. Cette remarque rejoint le débat récurrent en Recherche d'Information (RI) qui oppose la pertinence dite "système" renvoyant à une mécanique formelle d'évaluation et la pertinence dite "utilisateur" qui demeure mal identifiée et que certains jugeront trop ambiguë ou multiforme pour être formalisée. Selon Saracevic [SARA75], il est difficile de trancher en faveur de l'un ou l'autre de ces points de vue qui se complètent. Green [GREE95a] pour sa part considère que l'utilisateur est le véritable juge de ce qui est pertinent mais que d'un autre côté, il n'est peut-être pas le meilleur puisqu'il ne dispose pas nécessairement des connaissances qui lui permettraient d'évaluer la pertinence du document.

Quoiqu'il en soit, on peut penser que si le but ultime du système est de sélectionner l'information pertinente pour un utilisateur, il est indispensable d'intégrer plus amplement l'utilisateur dans la conception des systèmes de RI. Cette voie de recherche est d'ailleurs de plus en plus privilégiée comme en témoigne le numéro spécial de "Journal of the American Society for Information Science" consacré à la pertinence en RI [FROE94]. Ainsi, certaines recherches tentent de faire l'inventaire des critères qui peuvent jouer un rôle dans l'évaluation de la pertinence par l'utilisateur [BARR94], [HOWA94]. D'autres se proposent d'étudier l'interaction de l'utilisateur avec le système [BRUC94] ou recherche des modes d'interaction efficaces [DENO97].

Une autre alternative permettant de prendre en compte l'utilisateur consiste à intégrer son propre fonctionnement interne dans le système. Il s'agit de s'intéresser aux modèles du système cognitif humain portant sur des fonctions liées au processus de sélection d'information et d'essayer de les reproduire dans les systèmes de RI. Ainsi, les modèles de la mémoire sémantique qui portent sur l'organisation et l'accès des connaissances en mémoire sont directement liés au problème de la sélection d'information de même que les modèles qui sont en mesure d'expliquer pourquoi l'attention se focalise sur un objet de l'environnement plutôt qu'un autre.

L'objectif de la recherche que nous présentons ici est de proposer une nouvelle approche du filtrage d'information (FI) et plus généralement de la formalisation de la notion de pertinence en nous appuyant sur des modèles cognitifs de la sélection d'information. La première partie de cet article présente cette approche. Elle part de l'analyse de modèles de la mémoire sémantique et de l'attention (section 2) pour proposer après avoir tissé des liens avec les modèles traditionnellement utilisés en RI (section 3), un ensemble de grands principes cognitifs pour la RI. Une nouvelle formalisation de la pertinence s'exprimant comme une résonance dans un réseau de concepts associés est ainsi proposée. La seconde partie de l'article concerne la mise en œuvre de l'approche dans la tâche de filtrage adaptatif de documents textuels. Nous décrivons le système RELIEFS (section 4) ainsi que les expériences réalisées

avec ce système dans le cadre de la conférence TREC9. Enfin, nous concluons par une discussion générale en résumant notre approche et en dégageant les nouveaux éléments qu'elle apporte par rapport au problème de la formalisation de la notion de pertinence.

2. Modèles cognitifs pour la sélection d'information

L'organisation et l'accès des connaissances en mémoire conditionnent notre façon de sélectionner l'information. Quels mécanismes mnésiques interviennent dans une situation donnée pour sélectionner en mémoire une connaissance particulière et la prendre en compte dans un raisonnement? De même, quels mécanismes sont responsables de la focalisation de notre attention sur certains objets de l'environnement? Ces questions ont été abordées en Psychologie Cognitive. Des données comportementales ont été recueillies, des modèles ont été proposés dont certains ont été testés avec succès. Dans cette section, nous proposons de considérer certains de ces modèles comme des sources d'inspirations pour un système de recherche d'information.

2.1. Modèles de la mémoire sémantique

2.1.1. La mémoire comme un réseau de concepts

La mémoire est un vaste champ de recherche en Psychologie Cognitive. On distingue aujourd'hui différents types de mémoires, par exemple, la mémoire à court terme et la mémoire à long terme sur la base de la durée du stockage, la mémoire procédurale et la mémoire déclarative sur la base de la nature de l'information stockée. Notre objectif est simplement ici de pointer sur des travaux liés à l'organisation des connaissances, et leur mode d'accès pour lesquels on parle de mémoire sémantique.

Le premier modèle de la mémoire sémantique est dû à Ross Quillian [QUIL78] qui propose de représenter l'organisation des connaissances en mémoire par un réseau de concepts et qui définit une typologie des différents liens entre concepts. Dans ce réseau, la définition de la notion de concept est très générale, elle peut s'appliquer par exemple à des objets, des actions, des propriétés.

2.1.2. Des mécanismes de propagation de l'activation

Pour rendre compte de la façon dont cette mémoire est exploitée, Quillian définit aussi un ensemble de mécanismes de propagation de l'activation d'un nœud du réseau vers les autres nœuds du réseau [QUIL78], [COLL75].

Cette théorie rend compte de nombreux résultats expérimentaux concernant l'accès des connaissances en mémoire. Elle rend compte par exemple, des résultats

des expériences d'amorçage (priming en anglais) [NEEL91] qui étudient l'effet de la présentation d'un objet (l'amorce) sur le traitement d'un objet associé (la cible). Ainsi, le traitement du mot "chien" (dans une tâche de décision lexicale qui consiste à dire si le mot présenté existe dans la langue française) est facilité (on observe des temps de réponse plus courts) s'il est précédé du mot "chat". Dans ce cas, la facilitation peut s'expliquer par une pré-activation de la cible par l'amorce. Cette théorie est aussi en accord avec des observations portant sur les temps de recherche en mémoire [COLL89] qui montre que le temps de réponse à une question du type "A est-il un B ?" dépend de la longueur du chemin à parcourir dans la hiérarchie pour relier le concept A au concept B, comme on peut le prévoir si on considère que la réponse repose sur une propagation de l'activation entre les nœuds du réseau représentant respectivement A et B. Indépendamment de la plausibilité psychologique de cette théorie, il est intéressant de noter que ce mode de fonctionnement permet une mise à disposition automatique de la connaissance activée liée par exemple aux objets perçus dans le contexte courant, ou bien à l'objectif, donc à la connaissance susceptible d'être pertinente. C'est pour cette raison que Quillian considère que le mécanisme de propagation d'activation est bien adapté à la sélection d'information par des systèmes informatiques. Il évite ou réduit le processus coûteux de la recherche de connaissance particulièrement problématique lorsque les bases de données sont importantes [ANDE83b].

2.1.3. La nature associative de la mémoire

Anderson reprend dans son modèle ACT [ANDE83a] cette même idée de propagation d'activation dans un réseau de concepts pour rendre compte du cheminement d'un raisonnement (règle d'inférence à choisir à chaque étape) et considère des liens entre concepts de nature associative. Ainsi, il considère que la présence simultanée répétée de deux concepts (dans le cadre d'un raisonnement par exemple) conduit à une association entre ces concepts.

En fait, la nature associative de la mémoire est une hypothèse largement répandue et ancienne puisque déjà Aristote remarquait qu'une condition suffisante pour que deux événements soient liés en mémoire soit qu'ils se déroulent simultanément. De ce point de vue, les travaux sur l'apprentissage associatif [PAVL77] constituent une source intéressante, de même que ceux de Hebb [HEBB49] qui a montré la réalité de ce principe au niveau neurophysiologique. Il montre que "Quand un axone de la cellule A est assez proche pour exciter une cellule B et quand, de façon répétée et persistante, il participe à son activation, un certain processus de croissance ou un changement métabolique s'installe dans une cellule ou dans les deux tel que l'efficacité de A, en tant qu'elle est une des cellules qui active B, est augmentée"

2.2. Modèles de l'attention

2.2.1. Généralités

L'attention [CAMU96] peut être étudiée de deux façons : soit en tant que processus réglant (quels changements interviennent dans le traitement d'une information suivant le degré d'attention que je lui accorde) ou processus réglé (qu'est ce qui explique qu'elle puisse subitement changer de cible pour porter sur un nouvel objet).

Dans le premier cas (processus réglant), de nombreuses expériences convergent sur la facilitation du traitement d'information attendue et l'altération du traitement de l'information inattendue. Toujours dans cette perspective de processus réglant, l'attention fait aussi l'objet d'études en neurophysiologie. Certaines expériences [LABE 95] ont par exemple montré une amplification du niveau d'activité nerveuse manifestée dans certains sites cérébraux participant à l'élaboration de la représentation des stimuli engagés dans un processus d'attention (par rapport à une situation où ces stimuli ne sont pas engagés c'est-à-dire par rapport à une situation où ces stimuli ne sont pas pertinents pour le système cognitif).

Dans le second cas (processus réglé), on peut s'intéresser aux conditions dans lesquelles certains objets sautent aux yeux sans qu'il faille les chercher de façon délibérée. C'est ce qu'on appelle le phénomène de "pop-out" qui a été étudié par Treisman [TREI93]. Ces recherches ont par exemple permis de constater que les objets se différenciant des autres objets sur une propriété simple étaient plus saillants perceptivement (un rond rouge parmi des ronds bleus par exemple) que lorsqu'il fallait considérer une conjonction (un rond rouge parmi des ronds bleus, des carrés rouges et des carrés bleus).

Ce double aspect (processus réglant/processus réglé), met en avant un principe intéressant dans le mode de fonctionnement de notre système attentionnel. Il est en mesure de sélectionner des informations en fonction des objectifs (donc de jouer le rôle de filtre) et il permet aussi de réagir à d'éventuels événements inattendus. Cela renvoie à l'idée d'une interaction entre un processus descendant c'est-à-dire guidé par des connaissances et un processus ascendant partant des stimuli.

2.2.2. Le modèle ART (Adaptive Resonance Theory)

Le modèle ART de Grossberg [GROSS96] est cohérent avec l'idée d'amplification d'activation et reprend celle de double interaction entre processus ascendant et descendant. Cette théorie part de l'analyse d'un système en interaction avec son environnement. Elle adopte une approche située [CLAN 97] dans laquelle on s'efforce de comprendre le système cognitif en considérant le milieu dans lequel il évolue. Le système (S) prend en compte les informations du monde externe (E) tout en appliquant ses connaissances pour la sélection de ces informations (figure 1).

La sélection d'une information par le système fait donc intervenir deux processus de nature différente, un processus ascendant ($E \rightarrow S$) et un processus descendant ($S \rightarrow E$).

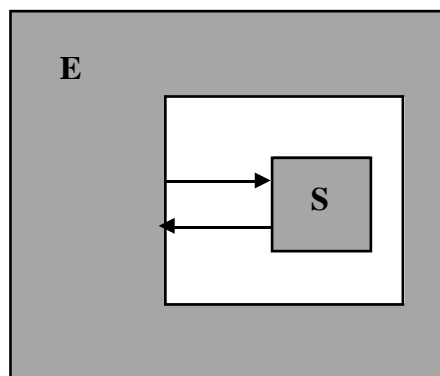


Figure 1: Interaction d'un système (S) avec son environnement (E). Extrait de [CLAN 97].

ART modélise cette interaction à partir d'un réseau de neurones à deux couches. La couche de neurone d'entrée (F1) permet de représenter le stimulus provenant de l'environnement. L'autre couche (F2) correspond à une représentation interne du système. Par exemple, un nœud de F2 peut correspondre à une catégorie dans laquelle les stimuli peuvent être rangés. Les deux couches sont liées par des connexions de nature associatives. Les connexions sont orientées et calculées selon une règle d'association similaire (dans le principe) à la règle de Hebb dans une phase d'apprentissage. Quand un nœud de F1 et un nœud de F2 sont simultanément activés leur connexion est renforcée.

Cependant toute l'information présentée au système n'est pas considérée pour l'apprentissage. D'une façon globale, la sélection d'une information dépend de son degré de résonance avec une représentation interne du système. Plus précisément, dans le réseau, un stimulus va être représenté par un pattern d'activation dans F1. Cette activation de F1 est propagée vers F2 et rétro-propagée (après compétition interne) vers F1 (figure 2). Si l'activation rétro-propagée à F1 s'apparie bien avec la distribution d'activation codant le stimulus dans F1, un état de résonance est atteint. En effet, l'activation étant retournée à la source, on peut parler de résonance dans le sens où cet appariement conduit à une amplification de l'activation des neurones codant pour le stimulus; amplification qui elle-même conduira à une amplification de l'activation des neurones codant pour la représentation interne etc...

D'un point de vue général, on peut s'interroger sur l'intérêt d'un tel mécanisme de sélection. Grossberg répond à cette interrogation en présentant la sélection par résonance comme une réponse à la question posée à tout système en interaction avec

un environnement : Comment préserver une connaissance précédemment acquise tout en continuant d'apprendre (le dilemme entre stabilité et plasticité [GROS 99]). Les connexions orientées de F1 vers F2 permettent de prendre en compte l'environnement (si elles n'existaient pas le système serait coupé du monde extérieur) et les connexions orientées de F2 vers F1 permettent de prendre en compte ses connaissances (s'il elles n'existaient pas, le système ne pourrait tenir compte des connaissances emmagasinées au cours de ses précédentes expériences). La sélection par résonance permet la prise en compte simultanée de ces deux types de connexions et garantit donc au système une plasticité tout en lui conférant une certaine stabilité.

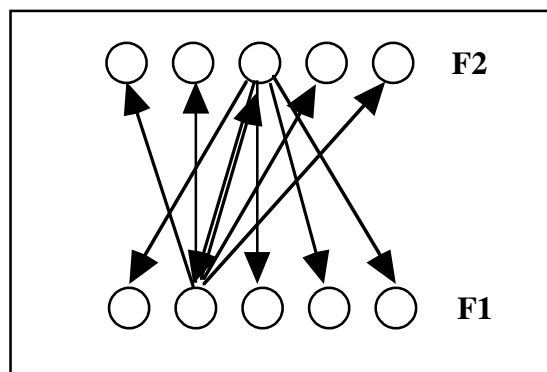


Figure 2. Propagation et rétro-propagation entre les couches F1 et F2.

3. Quels liens avec la Recherche d'Information ?

3.1. Modèles de la mémoire sémantique et RI

3.1.1. Utilisation et construction automatique de thésaurus

L'utilisation de thésaurus est maintenant très répandue en RI. Un thésaurus est une structure dans laquelle on représente des relations entre des termes. Il peut s'agir par exemple de relation de synonymie ou d'hyponymie. Un thésaurus permet d'étendre les requêtes en ne considérant pas seulement les termes saisis par l'utilisateur mais aussi les termes synonymes (par exemple) permettant ainsi d'augmenter les performances de rappel du système. Ils peuvent être assimilés avec des réseaux sémantiques. Certains thesauri comme Wordnet [FELL98], ont été construits manuellement, d'autres sont construits automatiquement. Concernant la construction automatique, le plus fréquemment, l'établissement d'un lien entre deux termes se fait sur la base du calcul d'une cooccurrence entre ces termes. Or, le

renforcement d'un lien entre deux termes présents simultanément (par exemple dans le même document) conduit à un calcul de cooccurrence. Les notions de réseau de concepts et celle de règle d'association présentes dans les modèles de la mémoire sémantiques ne sont donc pas étrangères à la RI.

3.1.2. Les méthodes de propagation d'activation

Le lien entre la RI et les méthodes de propagation d'activation proposées en psychologie cognitive ont déjà été exploités et ont donné dans certains cas des résultats intéressants. Par exemple, Cohen [COHE87], avec le système GRANT, obtient par propagation d'activation de meilleurs résultats que les systèmes classiques de recherche par mots clés. Ce système est chargé de trouver un organisme de financement pour un sujet de recherche donné. La méthode suivie consiste dans un premier temps à représenter dans un réseau sémantique les connaissances portant sur les sujets de recherche et les agences de financements, puis dans un second temps, à activer les connaissances relatives au sujet de recherche pour lequel doit être trouvé un financement et à laisser cette activation se propager dans le réseau vers les nœuds représentant les différentes agences de financement. Dans sa synthèse sur les méthodes de propagation pour la Recherche d'Information, Crestani [CREST97] distingue les méthodes de propagation «pures» et des méthodes de propagations plus contrôlées. Les méthodes de propagation «pures» consistant à laisser l'activation se propager sans introduire de contrainte ont un inconvénient, elles conduisent à la propagation de l'activation à l'ensemble du réseau ce qui ne leur permet donc plus de jouer leur rôle de sélection. Parmi, les contraintes communément ajoutées pour contourner ces obstacles on propose d'éviter les nœuds à forte connectivité. Il s'agit de nœuds représentant des concepts très généraux et par conséquent liés à de très nombreux autres concepts.

3.2. Analogie entre résonance et pertinence

Nous présentons un panorama des différentes formalisations de la pertinence proposées en RI en mettant en évidence la convergence autour de l'idée de double implication entre document et requête. Dans un second temps, nous établissons de façon plus précise l'analogie entre cette double implication et la résonance au sens où elle apparaît dans le modèle ART.

3.2.1. Formalisation de la notion de pertinence en recherche d'information

L'étude de la notion de pertinence renvoie à l'identification de la relation qui unit l'objet qualifié de pertinent (par exemple un document) à l'objet par rapport auquel on définit cette pertinence (par exemple un besoin d'information). Un préalable à l'étude de la relation concerne la mise en évidence des caractéristiques des objets intervenant dans la relation. Suivant ce schéma, on peut distinguer deux grands types de recherche portant sur la formalisation de la notion de pertinence en RI. Le

premier type concerne donc l'identification de la relation qui unit les deux objets. Les recherches de Green et Bean [GREE 95a, b] qui analysent la relation liant les thèmes d'un guide thématique aux passages référés dans le guide sont des exemples de ce type de recherche. Le second grand type de recherche porte sur la mise en évidence de ces critères d'évaluation (dans le premier type d'étude, on se focalise exclusivement sur le contenu du document, le critère thématique). En procédant à l'interrogation des utilisateurs sur leurs jugements de pertinence et en analysant leurs réponses, Barry [BARR 94] a par exemple distingué 7 grandes catégories de critères dont, le contenu du document, les expériences de l'utilisateur (connaissances par exemple), les croyances et préférences de l'utilisateur, les autres informations et autres sources (consensus, vérification externe), les sources des documents (qualité de la source), les documents comme entité physique (coût, facilité pour se le procurer), le contexte de l'utilisateur (contraintes de temps par exemple).

Nous nous plaçons ici dans le cadre du premier type de recherche, c'est-à-dire dans la problématique de l'identification de la relation liant un document pertinent à une requête. Un certain nombre d'études ont porté sur une formalisation mathématique de cette relation. La logique et les probabilités sont deux cadres dans lesquels des propositions ont été faites.

3.2.1.1. Formalisations logiques de la pertinence

Dans un cadre logique, déjà Cooper [COOP 71] insiste explicitement sur le lien étroit entre pertinence et conséquence logique " One might, on first reflection, suppose that logical consequence could be distantly related to relevance, but it will be argued shortly that when problem is posed in terms of declarative sentences, logical consequence and relevance are very intimately connected " ([COOP 71] p. 22). Plus récemment, dans le même ordre d'idée, van Rijsbergen [VAN 86] propose aussi de considérer la pertinence comme une implication logique entre un document et une requête, soit $D \rightarrow Q$. Pour illustrer cette proposition, considérons trois documents D1, D2 et D3, respectivement représentés par A, B&C et A&B&C (où A, B et C sont des mots clefs et "&" désigne une conjonction). Considérons également une requête Q représentée par A&B. Le document pertinent est D3, car D1 est trop général, D2 ne porte pas sur A et D3 est bien le seul document tel que $D \rightarrow Q$ ($A \& B \& C \rightarrow A \& B$). Nie [NIE 88] ajoute le rôle important joué par l'implication inverse ($Q \rightarrow D$) et propose une agrégation de l'évaluation des deux implications, soit $F(P(D \rightarrow Q), P(Q \rightarrow D))$. Pour illustrer cette idée, en reprenant l'exemple précédent, considérons le document D4 représenté par A&B. Il est manifestement plus en adéquation avec la requête que D3, et le seul moyen de départager D4 de D3 est de considérer $D \rightarrow Q$.

3.2.1.2. Formalisations probabilistes de la pertinence

La formulation logique de la pertinence n'est pas déconnectée des modèles probabilistes utilisés en RI. Van Rijsbergen [VAN 86] suggère que $P(D \rightarrow Q)$ soit évalué comme une probabilité conditionnelle $P(Q|D)$ et non comme une implication

matérielle¹. Les modèles probabilistes quant à eux sont basés sur le PRP [ROBE 77], principe selon lequel un système de recherche d'information obtient des performances optimales s'il est en mesure de classer les documents dans leur ordre de probabilité de pertinence $P(R|D,Q)$ (probabilité de pertinence sachant un document et une requête). Le modèle BIR (Binary Independence Retrieval Model), considère une requête fixe et des documents variables et s'intéresse à l'évaluation de la probabilité de pertinence de ces différents documents vis-à-vis de la requête fixée. Pour l'estimation de cette probabilité de pertinence, une application de la formule de bayes nous ramène à l'évaluation de $P(D|R)$ (la probabilité d'observer le document D sachant que le document observé est pertinent vis-à-vis de la requête). Le modèle BII (Binary Independence Indexing Model), considère un document fixe et des requêtes variables et s'intéresse à l'évaluation de la probabilité de pertinence du document vis-à-vis des différentes requêtes. Dans ce modèle, la formule de bayes est utilisée différemment (en considérant le document fixe cette fois-ci) pour faire apparaître $P(R|D)$ (probabilité de pertinence du document pour une requête arbitraire). On retrouve dans nombre de recherches [FUHR89], [KWOK90], [ROBE82], [WONG89], [WONG95] le souci d'une approche double mêlant l'approche "orientée requête" (dans laquelle on va considérer une requête fixe et estimer la pertinence d'un document en utilisant les jugements sur la pertinence des autres documents vis-à-vis de la même requête) et l'approche "orientée document" (dans laquelle on va considérer un document fixe et estimer la pertinence du document en utilisant les jugements de pertinence du document vis-à-vis des autres requêtes). Par exemple, Kwok [KWOK 90] montre que la prise en compte de ces deux aspects permet une augmentation des performances.

3.2.1.3. Synthèse

Cette problématique (approche orientée requête/document), n'apparaît pas déconnectée de la double implication que l'on retrouve dans les modèles logiques. En effet, en identifiant la pertinence R avec la requête Q , notions qui d'un point de vue abstrait correspondent toutes deux à l'objectif de la recherche (ce qui est pertinent correspond à la requête et vice-versa), les modèles BIR et BII nous ramènent respectivement à l'évaluation de $P(D|Q)$ et $P(Q|D)$ qui expriment l'entraînement de D par Q et inversement et qui renvoie donc à l'évaluation de $D \rightarrow Q$ et $Q \rightarrow D$ du modèle logique. Les systèmes connexionistes PIRC et MERCURE proposés respectivement par Kwok [KWOK95] et Boughanem [BOUG99] qui

¹ L'implication matérielle qui est vraie lorsque la prémisse est fausse quelle que soit la valeur de vérité de la conclusion, est souvent remise en cause pour l'évaluation de la pertinence. En particulier, une de ses conséquences est que lorsqu'un document est incohérent, il devient pertinent pour toutes les requêtes. Pour faire face à ce problème, certaines logiques comme la logique de la pertinence [ANDE75] ou les logiques conditionnelles [STAL 68] proposent d'autres représentations de l'implication dans lesquelles le désavantage de l'implication matérielle disparaît. Pour des exemples de l'utilisation de ces logiques dans le cadre de la RI, on pourra respectivement se reporter à [MEGH96] et [NIE96].

implémentent des modèles probabilistes renvoient de même à l'évaluation des 2 implications.

Pour compléter ce panorama, on peut aussi citer les modèles utilisant des réseaux bayésiens [TURT 91] qui calculent $P(D|Q)$ pour évaluer la pertinence des documents. On peut aussi remarquer que même le schéma de pondération $tf*idf$ utilisé dans le modèle de l'espace vectoriel [BAEZ199] est une forme de ces implications. La fréquence d'un terme dans un document donné (tf) est une forme de $P(t|D)$ et un composant de $P(Q|D)$ (puisque ce sont les termes de la requête que l'on considère par la suite). L'indice idf qui est basé sur la proportion de documents contenant un terme donné est une forme de $P(D|t)$, un composant de $P(D|Q)$.

On est donc amené à émettre l'hypothèse selon laquelle, la double implication entre D et Q formerait l'essence de la notion de pertinence. Même si ces deux implications ont été mesurées et combinées de différentes façons dans différents modèles.

3.2.2. L'analogie

Il est intéressant de noter l'analogie qui existe entre l'idée de résonance et les formalisations de la notion de pertinence proposées en RI. En effet, il est naturel de considérer que le document D correspond à l'objet provenant de l'environnement, c'est-à-dire au stimulus et que la requête Q correspond à l'objectif du système, c'est-à-dire à une représentation interne. Dans ces conditions, l'implication $D \rightarrow Q$ correspond à la propagation d'activation de $F1$ vers $F2$ et l'implication inverse à la propagation de $F2$ vers $F1$. Cette dernière analogie n'est pas une analogie de surface. La propagation est réalisée en tenant compte du poids des connexions liant les nœuds des deux couches. Or, ces connexions sont calculées selon une règle d'association qui, en quelque sorte, mesure une "implication statistique" [SMOL192], forme subsymbolique de l'implication logique.

Plus précisément, considérons un nœud de $F1$ pour chaque terme du document. Considérons un nœud de $F2$ pour représenter une requête particulière. Au cours de l'apprentissage, la règle d'association permet de calculer pour chaque mot la force d'une règle du type « \square le mot est présent alors le document est pertinent par rapport à la requête \square . Cette force est matérialisée par les poids des connexions des nœuds de $F1$ vers le nœud de $F2$ représentant la requête. La règle d'association permet aussi de calculer la force de règles du type « \square le document est pertinent par rapport à la requête alors le mot est présent \square . Cette force est matérialisée par les poids des connexions du nœud de $F2$ représentant la requête vers les nœuds de $F1$. Si l'on reprend l'analogie, un document est un stimulus. Il est représenté par un pattern d'activation dans $F1$, les nœuds correspondant aux termes présents dans le documents sont activés. L'activation propagée au nœud requête de $F2$ dépend du poids des connexions menant des nœuds activés de $F1$ vers ce nœud. L'activation propagée dépend donc de la force des règles du type « \square le terme est présent alors le document est pertinent \square pour les termes du documents. L'activation propagée

renvoie donc bien à une mesure de $D \rightarrow Q$. L'activation propagée est d'autant plus retro-propagée au nœuds initialement activés dans F1 que les connexions orientées du nœud requête vers ces nœuds qui expriment l'entraînement « \square le document est pertinent alors le terme est présent \square sont fortes. On retrouve bien une expression de $Q \rightarrow D$. La rétro-propagation aux nœuds initialement activés qui se traduit par un phénomène de résonance est donc bien l'analogie en terme de propagation d'activation de $D \rightarrow Q$ et $Q \rightarrow D$ qui correspondent aux mesures traditionnellement considérées pour la formalisation de la pertinence en RI.

3.3. De grands principes cognitifs pour un système de RI

Sur la base de l'analyse des modèles cognitifs que nous venons de présenter et leur cohérence avec les modèles traditionnellement utilisés en RI, nous proposons d'extraire différents grands principes mis en évidence. Des modèles de la mémoire sémantique, nous retenons l'idée de sélection d'information par propagation d'activation dans un réseau de concept de nature associative. Du modèle ART, nous retenons le principe de sélection par résonance.

Nous proposons donc un système de recherche d'information basé sur la mise à jour et l'exploitation d'un réseau de concepts. Les concepts correspondront dans le cadre de la recherche de documents textuels à des termes ou des requêtes. Nous proposons que ce réseau soit construit sur la base d'une règle d'association similaire à la règle de Hebb. Nous proposons aussi que la sélection d'information, c'est-à-dire la sélection de documents pertinents relativement à une requête ou la sélection de mots pour l'extension d'une requête soit basé sur un principe de résonance.

4. Description du système RELIEFS

Avec la description du système RELIEFS, nous proposons maintenant une mise en œuvre des principes énoncés précédemment. Il ne s'agit que d'une mise en œuvre possible, d'autres choix cohérents avec les principes d'association et de résonance auraient probablement pu être faits.

4.1. Construction et Exploitation d'un réseau associatif

Dans le cadre de la sélection de documents textuels, nous considérerons que les stimuli correspondent aux différents mots des documents et nous considérerons une unité abstraite (que nous représenteront par le « \square nœud requête \square ») désignant la pertinence d'un document représentant l'objectif du système c'est-à-dire la requête. On notera qu'un nœud requête distinct est créé pour chaque requête différente. Nous allons tenir à jour un réseau liant les nœuds représentant les différents mots rencontrés dans les documents et le nœud requête (figure 3).

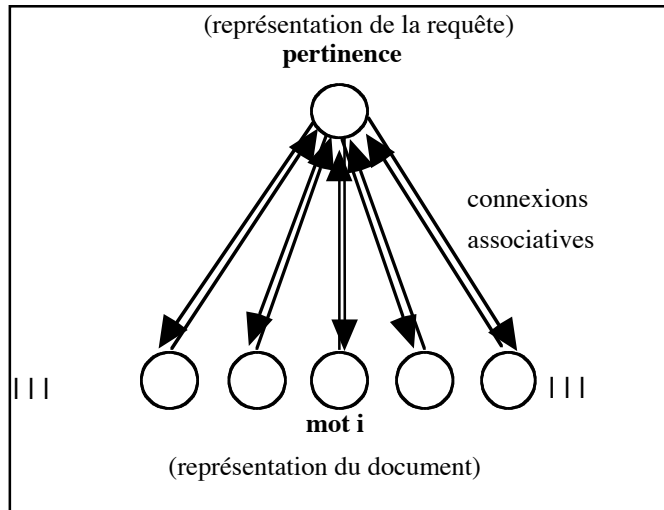


Figure 3. Réseau construit et mis à jour par RELIEFS.

On peut résumer le fonctionnement global du système en quelques lignes. Lorsqu'un document se présente, les mots présents dans ce document vont activer les nœuds correspondant dans le réseau et agir comme des indices de la pertinence du document. Un bon indice est un mot dont la résonance avec le nœud requête est importante (les deux implications sont fortes). La pertinence globale du document est calculée sur la base de la résonance de l'ensemble de ses mots vis-à-vis du nœud requête. Enfin, pour adapter le système au besoin de l'utilisateur, le système doit apprendre à partir des retours de pertinence fournis par l'utilisateur pour les documents sélectionnés. Cet apprentissage consiste à modifier la valeur des connexions entre les mots et le nœud requête.

4.2. Evaluation de la pertinence d'un document

La pertinence d'un document est évaluée comme une résonance. Dans le but d'évaluer une résonance entre un document et une requête, nous proposons d'évaluer la résonance de chacun des mots présents dans le document. La qualité de résonance d'un mot correspond à sa capacité à propager l'activation vers le nœud requête et à recevoir une activation de ce même nœud. Soit W_{iR} , le poids de la connexion orientée du mot i vers le nœud requête. Soit W_{Ri} et le poids de la connexion inverse. La résonance entre le mot i et la requête est évaluée par le produit $W_{iR} \cdot W_{Ri}$. L'absence d'un mot est donc d'autant plus pénalisante que ce mot est résonant avec la requête c'est-à-dire que $W_{iR} \cdot W_{Ri}$ est grand. Elle correspond à l'idée que l'absence

d'un relais de l'activation entraînerait une dissipation de l'activation et ne favoriserait pas l'émergence d'un état de résonance. La fonction d'évaluation retenue correspond à la somme des produits pour chacun des mots présents dans le document. De plus, de façon à normaliser cette somme nous l'avons rapportée à la même somme pour l'ensemble des mots (que ces mots soient présents ou pas dans le document).

L'évaluation de la pertinence d'un document D vis-à-vis d'une requête Q s'écrit donc finalement :

$$R(D,Q) = \frac{\sum_{i^*} W_{Ri} \cdot W_{iR}}{\sum_i W_{Ri} \cdot W_{iR}} \quad [1]$$

où i^* sont les indices des mots présents dans le document.

Le produit $W_{iR} \cdot W_{Ri}$ peut s'interpréter, comme nous venons de le faire en terme de propagation d'activation. Si l'on tient compte du fait qu'il s'agit d'un produit de fréquences relatives, il peut aussi s'interpréter plus empiriquement. Ainsi, un mot important (un mot dont la présence influe de manière significative sur la décision de prendre ou non le document qui le contient) est un mot qui est à la fois un bon prédicteur (quand le mot est présent, en général, le document est pertinent) et un mot qui est assez fréquemment présent lorsqu'un document est pertinent. La prise en compte de W_{iR} permet de sélectionner de bons prédicteurs. La prise en compte de W_{Ri} permet de sélectionner parmi ces prédicteurs, ceux qui surviennent le plus fréquemment c'est-à-dire ceux pour lesquels la qualité de prédiction a le moins de chance d'être liée à la présence fortuite du mot dans quelques documents pertinents et ceux qui seront le plus susceptibles d'être présents dans les prochains documents pertinents et donc le plus susceptibles d'être utiles aux futures prises de décision concernant à la pertinence des documents. On notera que la prise en compte de W_{Ri} permet aussi d'éviter les mots qui sont simplement très présents (fréquence marginale élevée) comme les mots vides qui sont donc plus présents que les autres mots dans les documents pertinents sans pour autant être "réellement" liés au thème sur lequel porte la requête. Il s'agit des mots qui sont écartés dans le modèle vectoriel sur la base de l'indice idf. Ces mots correspondent aussi, dans le cadre des méthodes de propagation d'activation, aux nœuds à forte connectivité, nœuds à éviter si l'on ne souhaite pas activer tout le réseau.

L'évaluation de la correspondance proposée en (1) entre le document et la requête est assez proche de celle considérée dans le modèle vectoriel. En effet, le produit scalaire des vecteurs document et requête revient aussi à faire la somme des poids des termes qui représentent la requête et qui sont aussi présents dans le document. Le dénominateur n'est quant à lui pas très éloigné de la norme du vecteur requête qui apparaît dans le modèle vectoriel. Par contre, nous n'avons pas tenu compte au niveau de la formule de correspondance de la taille du document

(exprimé dans le modèle vectoriel par la norme du vecteur document) ni de l'indice tf. Ce sont des aspects que nous souhaiterions intégrer par la suite tout en restant dans une interprétation de la correspondance en terme de résonance. Par ailleurs, les systèmes de filtrage adaptatif utilisant le modèle vectoriel considèrent aussi outre la pondération tf*idf, une mise à jour du poids des termes tenant compte des retours de pertinence des utilisateurs. C'est la formule de Rocchio [ROCC 71] qui est la plus souvent utilisée pour cette mise à jour. Or, dans cette dernière, on trouve explicitement la prise en compte de W_{Ri} et indirectement celle de W_{iR} . On peut ainsi vérifier comme nous le mentionnions dans la partie concernant les formalisations de la pertinence que l'on retrouve dans notre approche la plupart des aspects présents dans les modèles traditionnellement utilisés en RI et en FI bien qu'ils soient présentés sous un nouvel éclairage.

4.3. Règle d'apprentissage

Etant donné un document et un jugement de pertinence, nous souhaitons mettre à jour les connexions présentes dans la structure. Soient $W_{AB(k-1)}$ et W_{ABk} les évaluations de la connexion orientée de le nœud A vers le nœud B, avant et après la prise en compte de la k-ième observation (un couple document-jugement de pertinence, que l'on note O_k). Soit $\mu_A(O_k)$ une valeur indiquant la présence ou l'absence de l'objet A dans O_k (1 si présent et 0 si absent). La règle d'apprentissage proposée est :

$$W_{ABk} = \frac{\alpha W_{AB(k-1)} + \mu_A(O_k) \cdot \mu_B(O_k)}{\alpha + \mu_A(O_k)} \quad \text{avec } \alpha = \sum_{i=1}^{k-1} \mu_A(O_i) \quad [2]$$

Dans notre cas, A et B correspondent respectivement à un mot (représenté par un nœud de F1) et à la "pertinence" (représentée par le nœud requête) ou inversement. La règle fonctionne de la façon suivante :

- Si A et B sont présents dans l'observation, la connexion de A vers B et la connexion inverse sont renforcées. On prend ainsi en compte le jugement de l'utilisateur en augmentant le poids des mots apparaissant dans le document jugé pertinent.

- Si A est présent et B ne l'est pas, la connexion de A vers B est affaiblie et la connexion inverse reste inchangée. Si A est un mot et B le nœud requête ceci revient à affaiblir le poids attribué à un mot qui apparaît dans un document jugé comme non pertinent par l'utilisateur. Si A est le nœud requête et B est un mot, ceci revient à affaiblir le poids attribué à un mot qui n'apparaît pas dans un document jugé comme pertinent par l'utilisateur.

Le numérateur de la règle (2) correspond précisément à la forme générale de la règle de Hebb [HEBB 49] dont le principe est associatif (si deux nœuds sont

activées simultanément leurs connexions sont renforcées). Le dénominateur permet de normaliser le poids. On peut d'ailleurs aisément montrer que W_{ABn} représente la fréquence de l'objet représenté par B sachant la présence de l'objet représenté par A pour les n premières observations [BROU 00b]. La règle associative présentée permet donc simplement de calculer de façon incrémentale la fréquence relative de documents pertinents sachant la présence d'un mot ainsi que la fréquence d'un mot dans les documents pertinents.

Conformément à notre objectif, nous avons proposé une règle basée sur un principe d'association fidèle à la formulation de Hebb qui conduit à la mesure d'une implication. La mesure d'implication considérée est une simple fréquence relative. D'autres mesures plus complexes auraient pu être considérées [CHER 02] (on pourra aussi se reporter à [BROU 00b] pour une discussion sur le problème de l'évaluation de l'implication), la difficulté étant alors d'en fournir une formulation "Hebbienne".

5. Expérimentations

Nous avons expérimenté ce système dans la tâche de filtrage adaptatif dans le cadre de la conférence TREC-9. La particularité de cette sous-tâche de filtrage est de partir d'un très petit nombre de documents pertinents et de tenter de trouver de nouveaux exemples de documents pertinents tout en essayant de minimiser les erreurs. L'ensemble des documents considéré pour la tâche de filtrage de TREC-9 était la collection OHSUMED extraite de MEDLINE (articles de médecine). Les documents des années 1988-1991 (environ 300.000 documents) constituaient l'ensemble test et seulement deux documents de 1987 par requête servaient à un pré-apprentissage. Dans le but de pouvoir comparer les systèmes, une fonction d'utilité et une fonction de précision étaient fournies par les organisateurs. Cette année, la fonction d'utilité avait la forme : $T9U = 2 * R - N$, où R représente le nombre de documents pertinents sélectionnés et N représente le nombre de documents non pertinents sélectionnés. La fonction de précision avait la forme $T9R = (R / \max((R+N), 50))$. Nous avons essentiellement focalisé sur l'optimisation de T9U. Les systèmes étaient testés sur 63 requêtes fournies avec la collection.

5.1. Adaptation de RELIEFS à la tâche de filtrage

Les principes de RELIEFS tels que nous venons de les décrire ont constitué l'essentiel du système qui a été utilisé dans la tâche de filtrage adaptatif de TREC-9. Néanmoins, d'autres composants spécifiques à cette tâche devaient être ajoutés. Il s'agit de la gestion des seuils, dont l'objectif est de définir le score à partir duquel le système décide de sélectionner le document et d'un module servant à la sélection des mots du texte pour l'évaluation ainsi que pour l'apprentissage (on ne considère donc pas tous les mots du document).

5.1.1. Gestion des seuils

Le seuil optimal varie selon la requête et l'apprentissage réalisé sur les précédents documents sélectionnés. Bien que des considérations théoriques puisse être faites sur ce sujet [ARAM 00], les systèmes de filtrage adoptent généralement des règles assez empiriques en ce qui concerne la gestion des seuils. Pour notre part, nous avons considéré une mise à jour incrémentale considérant les deux règles suivantes:

- Si un document sélectionné n'est pas pertinent le seuil est augmenté (on considère que le système est trop tolérant).
- Si un document (qu'il soit pertinent ou pas) n'est pas sélectionné le seuil est abaissé (on considère que le système est trop sélectif).

L'amplitude de l'augmentation est plus importante que celle de la diminution car cette dernière est réalisée beaucoup plus fréquemment (la majorité des documents ne sont pas sélectionnés). Nous considérons cette diminution de façon à éviter que le système reste silencieux dans le cas où le seuil initial fixé (que nous basons sur le score réalisé par les documents donnés en apprentissage) est trop haut. Dans les deux cas, nous avons considéré différents critères pour adapter l'amplitude de la variation à la situation. Il s'agit des critères locaux suivants :

- le nombre de documents non pertinents consécutifs sélectionnés. Plus ce nombre est important, plus l'augmentation est forte et la diminution faible.
- le nombre de documents pertinents consécutifs (uniquement pour la diminution). Plus ce nombre est important, plus la diminution est forte.
- le nombre de documents traités (on prend plus de risque au début qu'à la fin). Plus ce nombre est important moins l'amplitude de la variation est forte.
- la densité de documents pertinents estimée c'est-à-dire le nombre de documents pertinents trouvés sur le nombre de documents traités.

Nous avons aussi considéré un critère plus global portant sur les probabilités de pertinence estimées pour chaque seuil. Un tableau indiquant pour chaque score une fréquence de pertinence est tenue à jour. Le score correspondant à une probabilité de 0.33 (qui correspond à la probabilité optimisant la mesure d'utilité) est une valeur de seuil vers laquelle le système tend à se régler.

La mise à jour des seuils a ainsi été réalisée de la façon suivante□

$$\text{nouveau_seuil} = \text{seuil} \pm (cg \cdot c1 \cdot c2 \cdot c3 \cdot c4 \cdot \max[0.1, |\text{Score}(0.33) - \text{seuil}|])$$

où cg est un coefficient qui permet de régler globalement l'amplitude des variations (dans nos expériences pour l'augmentation $cg=0.1$, pour la diminution $cg=0.00001$), $c1$, $c2$, $c3$, $c4$ sont des coefficients variant entre 0.5 et 3 qui dépendent respectivement du nombre de documents pertinents consécutifs ($c1$ vaut toujours 1 dans le cas d'une diminution), du nombre de documents non pertinents consécutifs,

du nombre de documents traités, de la densité estimée de documents pertinents. la valeur $Score(0.33)$ correspond au score pour lequel la fréquence de documents pertinents est de 0.33. L'opérateur max permet simplement de garantir une variation minimale. Tous ces coefficients ont été réglés manuellement. Il n'y a pas eu de véritable optimisation. Nous nous sommes bornés à considérer une plage de valeurs acceptables. L'impact de ces différents paramètres pourrait faire l'objet d'une étude à part entière et les coefficients pourraient être appris (évidemment sur une autre base que celle des tests).

De plus, de façon à ce que chaque score soit associé à une probabilité même si aucune observation ne correspondait au score, nous avons considéré un principe de diffusion [BROU 00b] qui permet de généraliser une observation à son voisinage. Enfin, le seuil a été modifié pour prendre en compte de la longueur des documents. Cela a été facilité par le fait que les différents documents étaient tous soit des résumés soit des titres et qu'il n'y avait donc globalement que deux longueurs. On a donc simplement choisi un seuil un peu plus petit pour les titres.

5.1.2. Module de sélection des mots

Dans le but d'alléger le traitement réalisé pour chaque document (évaluation et apprentissage) et dans le but d'éliminer le bruit que constitue la présence fortuite d'un mot dans un document pertinent, nous avons d'abord considéré les meilleurs mots du document (toujours sur le critère du produit des fréquences relatives). Si moins de N mots pouvaient être sélectionnés de cette façon, cette sélection était complétée par les mots du document les mieux reliés aux mots de la requête sur la base d'un thésaurus construit automatiquement à partir de l'ensemble d'apprentissage (documents de 1987). Finalement, lorsque moins de N mots restaient sélectionnés à la suite de cette seconde phase les mots du document étaient pris dans leur ordre de lecture. Dans nos expériences nous avons pris $N=20$.

5.2. Premiers résultats

Pour l'évaluation de TREC-9, nous avons soumis deux résultats, l'un pour l'utilité et l'autre pour la précision. Après un premier examen, la comparaison sur le critère d'utilité avec les autres systèmes est favorable puisque environ 80 % de nos scores sont supérieurs ou égaux à la valeur médiane (60 % sont supérieurs 20% égaux) et le système figure parmi les cinq systèmes à avoir une utilité positive (+1.1).

Depuis les tous premiers résultats nous avons éliminé le bruit correspondant à la présence fortuite de mots dans les documents pertinents en ne considérant que les 50 meilleurs mots pour le calcul du dénominateur de la fonction d'évaluation. Cette simple amélioration nous a permis d'obtenir un score d'utilité de +8.3 (pour une précision de 0,29 et un rappel de 0,24) très proche des résultats obtenus par les meilleurs systèmes.

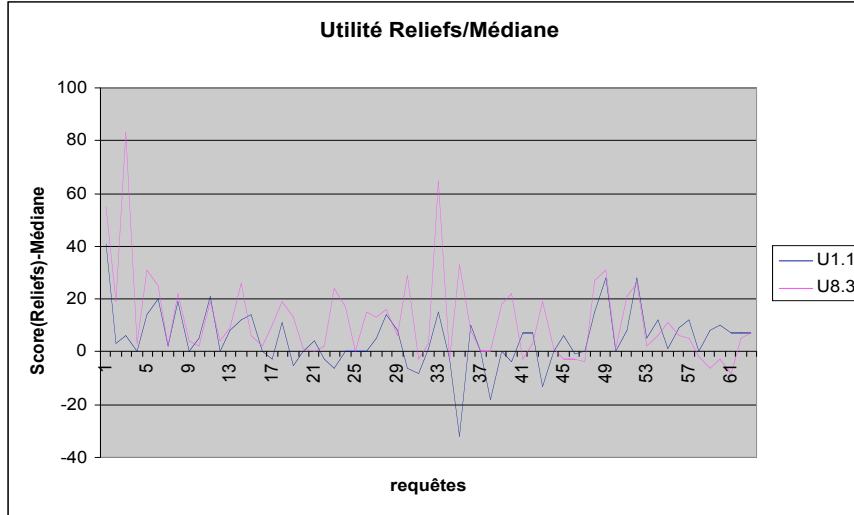


Figure 4. Comparaison de Reliefs dans sa première version (U1.1) et après amélioration (U8.3) sur le critère d'utilité avec la valeur médiane des différents systèmes présents à TREC9.

La figure 4 présente la comparaison avec la valeur médiane des autres systèmes optimisés pour la mesure d'utilité. Aucune véritable optimisation des paramètres du système (coefficient de variation des seuils, nombre de mots considérés,...) n'a été réalisée. Certains aspects comme nous l'avons mentionné précédemment ont été peu ou pas pris en compte (longueur du document, tf). Les bons résultats néanmoins obtenus sur un corpus de grande taille nous laissent penser que l'approche est viable.

5.3 Pondération du rôle relatif des deux implications

Dans les tests précédents, les deux implications sont considérées avec la même importance : elles sont simplement multipliées. Mais ont-elles réellement la même importance? Pour tenter de répondre à cette question nous avons fait varier l'importance d'une implication en remplaçant dans (1) le poids W_{iR} associé à la connexion du terme i vers le nœud requête par $(W_{iR})^\rho$. La formule d'évaluation de la pertinence s'écrit alors :

$$R(D,Q) = \frac{\sum_i W_{Ri} (W_{iR})^\rho}{\sum_i W_{Ri} (W_{iR})^\rho}$$

Le coefficient ρ a un impact direct sur la stabilité/plasticité du système. Plus ρ est grand, plus le système est plastique puisqu'alors nous favorisons D->Q qui selon les termes de ART correspond au processus ascendant (prise en compte de l'environnement). Au contraire, plus ρ est petit plus le système est stable puisque nous favorisons alors Q->D, ce qui selon les termes de ART correspond au processus descendant (prise en compte des connaissances acquises lors des précédentes expériences).

Si l'on considère plus précisément les poids, plus ρ est petit, plus on privilégie les poids W_{Ri} (qui favorisent les mots qui sont les plus fréquemment rencontrés dans des documents pertinents) au détriment des poids W_{iR} et donc plus on exige que les documents sélectionnés contiennent les mots les plus fréquemment rencontrés dans les documents pertinents précédents. Or, puisque l'apprentissage est réalisé sur la base des documents sélectionnés, cela ne permet pas l'émergence des mots encore peu rencontrés dans les documents pertinents. Le système restera dans ce cas très stable sélectionnant toujours les documents sur la base des mêmes mots et ne s'adaptera pas.

ρ	Utilité	ρ	Utilité
0.0	2.00	1.0	8.19
0.1	4.43	1.1	7.25
0.2	5.20	1.2	5.79
0.3	6.57	1.3	6.79
0.4	6.80	1.4	6.12
0.5	6.78	1.5	5.59
0.6	8.30	1.6	5.09
0.7	8.00	1.7	5.51
0.8	7.51	1.8	4.83
0.9	7.49	1.9	1.97

Table 1. Les valeurs d'utilité en fonction des valeurs de ρ .

5.4 Extension de requêtes

Lorsque moins de N mots indicateurs d'une éventuelle pertinence (parce qu'étant déjà apparus dans les documents pertinents précédents) pouvaient être sélectionnés, la sélection était complétée par les mots du document les mieux reliés aux mots de la requête. Cette sélection a été réalisée sur la base d'un thésaurus construit automatiquement à partir des documents de 1987. Ce thésaurus correspond simplement à un réseau dont les nœuds représentent les mots et les connexions sont pondérées par les fréquences relatives (calculées à partir de (2)) liant ces différents mots.

L'évaluation des mots les mieux reliés à la requête a été calculée en se basant sur la mesure de leur double implication vis-à-vis de l'ensemble des mots de la requête (plus le nombre de mots de la requête auquel le mot candidat est spécifiquement relié est grand, plus la relation est considérée comme importante). La double implication qui correspond ici au produit de la fréquence du mot sachant la présence du mot de la requête avec la fréquence inverse a donc une nouvelle fois été utilisée pour sélectionner les mots "pertinents". Le calcul des fréquences relatives liant les mots des documents de 1987 correspond à la construction d'un thésaurus flou [RADE 76], [MIYA 90] et la complétion par les mots reliés à la requête est une forme d'extension de requête. Les résultats montrent une légère amélioration avec l'utilisation du thésaurus (table2). Ces résultats permettent de penser que l'idée de résonance est aussi viable dans le cadre de l'extension de requête. D'autres tests seront réalisés en vue de comparer cette forme d'extension à d'autres formes d'extension.

ρ	AVEC THESAURUS	SANS THESAURUS
0.6	8.30	7.12
0.7	8.00	6.56
0.8	7.51	7.03
0.9	7.49	7.06
1.0	8.19	6.76
1.1	7.25	6.14

Table 2. Résultats avec et sans thésaurus pour les différentes valeurs de ρ .

6. Discussion

Pour la conception de ce système, nous nous sommes inspirés de principes cognitifs dans un sens semblable à certaines recherches qui ont tenté de reproduire les mécanismes de propagation d'activation pour la RI [CRES 97]. D'une part, la règle d'apprentissage que nous utilisons est de type associatif et reprend un principe dont la réalité neurophysiologique a été mise en évidence par Hebb. Par ailleurs, nous avons repris l'idée de résonance développée dans la théorie ART dont la plausibilité psychologique est appuyée par de nombreuses données neurologiques et de nombreux succès dans la modélisation des différents processus cognitifs [GROS 99]. Ce système apparaît donc cohérent avec les principes cognitifs d'association et de résonance. Cet aspect nous paraît important dans le sens où le rôle d'un tel système est de sélectionner des informations pour un utilisateur et que l'intégration des principes de fonctionnement de l'utilisateur (même s'ils sont de très bas niveau) dans le système peut être profitable.

Par ailleurs, il semble qu'en identifiant résonance et pertinence, on apporte un éclairage nouveau sur la notion de pertinence. La sélection d'une information provenant de l'environnement sur la base de sa résonance avec une représentation interne garantit au système un compromis entre stabilité et plasticité. Si l'on identifie résonance et pertinence, une information pertinente devient donc une information qui garantit plasticité et stabilité au système et donc qui garantit une "bonne" évolution du système dans son environnement. La qualité de pertinence serait donc intimement liée au problème de l'interaction d'un système avec son environnement. L'étroite relation qui existe entre pertinence et interaction a déjà été discutée par différents auteurs particulièrement lorsque l'on réduit la notion d'interaction à celle de communication. Ainsi, pour Saracevic [SARA 75] (p.321) «*In the most fundamental sense, relevance has to do with effectiveness of communication*», idée reprise dans la théorie de la communication proposée par Sperber & Wilson [SPER 95].

Notre approche permet d'une certaine façon de réconcilier les approches formelles de la pertinence (pertinence système) avec des approches tournées vers l'utilisateur (pertinence utilisateur) puisque nous montrons qu'un mécanisme cognitif de sélection de l'information (la résonance) est cohérent avec les formalisations de la pertinence en RI. Ce point de vue cognitif insiste sur l'aspect subjectif de la pertinence. Il la définit relativement à l'identité du système auquel l'information est présentée, cette identité évoluant elle-même avec l'apprentissage. En ce sens, il apporte un élément de réflexion par rapport au débat opposant "pertinence système" et "pertinence utilisateur" en déplaçant la question de l'objectivité vers celle de la différence d'identité des systèmes.

Pour résumer, d'un point de vue pratique, nous avons présenté un nouveau système pour le filtrage adaptatif de documents textuels. Ce système a l'avantage de la simplicité puisqu'il se limite à l'évaluation de fréquences relatives et qu'il les compose très simplement. Les résultats obtenus sont concluants. Cette recherche avait cependant des objectifs plus théoriques. Ainsi, nous avons présenté une

approche nouvelle du filtrage de documents qui permet de poser un nouveau regard sur la notion de pertinence et d'en donner une nouvelle interprétation. Notre inspiration est d'ordre cognitive et l'approche que nous avons prise est conceptuelle. Nous nous limitons pour l'apprentissage à considérer un principe d'association. Nous nous limitons pour l'évaluation à considérer un principe de résonance. L'objectif avec RELIEFS est d'isoler des principes généraux et de montrer leur intérêt pour la formalisation de la notion de pertinence. Les bons résultats obtenus par le système viennent confirmer la validité de l'approche.

Bibliographie

- [ANDE 75] Anderson, A. R., & Belnap, N. D. Entailment. *The logic of relevance and necessity*. Princeton: Princeton University Press, 1975.
- [ANDE 83a] Anderson, J. R. *The architecture of cognition*. Cambridge, MA: Harvard University Press, 1983.
- [ANDE 83b] Anderson, J. R. A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22, p. 261-295, 1983.
- [ARAM 00] Arampatiz, A., Beney, J. Koster, C.H.A., van der Weide, T.P. Incrementally, Half-life and Threshold Optimization for Adaptive Document Filtering, *Proceedings of the Text Retrieval Conference (TREC-9)*, 2000, Gaithersburg, p. 589-600.
- [BARR 94] Barry, C. L. User-defined relevance criteria: An exploratory study. *Journal of the American Society for Information Science*, 45(3), p. 149-159, 1994.
- [BAEZ 99] Baeza-Yates, R. & Ribeiro-Neto, B. *Modern Information Retrieval*. New York, NY: ACM Press, 1999.
- [BOUG 99] Boughanem M., Christment, C., Soulé-Dupuy, C. Query Modification Based on Relevance Back-Propagation in an Ad hoc Environment, *Information Processing and Management*, 35(2), p. 121-139.
- [BROU 00a] Brouard, C. & Nie, J-Y. The system RELIEFS: a new approach for information filtering. *Proceedings of the Text Retrieval Conference (TREC-9)*, 2000, Gaithersburg, p. 513-517.
- [BROU 00b] Brouard, C. Une méthode incrémentale et associative pour l'extraction de règles floues. *Revue d'Intelligence Artificielle*, 14 (3-4), 2000, p. 397-426, Hermès, Paris.
- [BROU 01] Brouard, C. & Nie, J-Y. RELIEFS: Un système pour le filtrage adaptatif de documents textuels basé sur la notion de résonance. *Actes de la conférence INFORSID*, Genève, 2001, p. 267-278.
- [BRUC 94] Bruce, H.B. A cognitive view of the situational dynamism of user-centered relevance estimation. *Journal of the American Society for Information Science*, 1994, 45(3), p. 142-148.
- [CAMU 96] Camus, J.-F. *La psychologie cognitive de l'attention*. Armand Colin, Paris, 1996.

RSTI- ISI-7/2002. Recherche et filtrage d'information

- [CHER 02] Cherfi H. & Toussaint Y. Interprétation des règles d'association extraites par un processus de fouille de textes, In *Actes du 13ème Congrès francophone AFRIF-AFIA de Reconnaissance des Formes et d'Intelligence Artificielle*, Angers, 2002, vol 3, p. 975-983.
- [CLAN 97] Clancey, W. *Situated Cognition*. Cambridge: Cambridge University Press, 1997.
- [COHE 87] Cohen, P. R., & Kjeldsen, R. Information retrieval by constrained spreading activation in semantic networks. *Information Processing and Management*, 1987, 23(4), p. 255-268.
- [COLL 75] Collins, A. M., & Loftus, E. F. A spreading activation theory of semantic processing. *Psychological Review*, 1975, 82, p. 407-428.
- [COLL 69] Collins, A. M., & Quillian, M. R. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 1969, 8, p. 240-247.
- [COOP 71] Cooper, W. S. A definition of relevance for information retrieval. *Information Storage and Retrieval*, 1971, 7, p. 19-37.
- [CRES 97] Crestani, F. Application of spreading activation techniques in information retrieval. *Artificial Intelligence Review*, 1997, 11, p. 453-498.
- [DENO 97] Denos, N. Modélisation de la pertinence en recherche d'information : modèle conceptuel, formalisation et application. Thèse de doctorat, Joseph Fourier - Grenoble I, 1997.
- [FELL 98] Fellbaum, C. *WordNet: an electronic lexical database*. Cambridge, Massachusetts: The MIT Press, 1998.
- [FROE 94] Froehlich, T. J. Relevance reconsidered-toward an agenda for the 21st century: Introduction to special issue on relevance research. *Journal of the American Society for Information Science*, 1994, 45(3), p. 124-134.
- [FUHR 89] Fuhr, N. Models for retrieval with probabilistic indexing. *Information Processing and Management*, 1989, 25(1), p. 55-72.
- [GREE 95a] Green, R. Topical relevance relationships. I. Why topic matching fails. *Journal of the American Society for Information Science*, 1995, 46(9), p. 646-653.
- [GREE 95b] Green, R., & Bean, C. A. Topological relevance relationships. II. An exploratory study and preliminary typology. *Journal of the American Society for Information Science*, 1995, 46(9), p. 654-662.
- [GROS 76] Grossberg, S. Adaptive pattern classification and universal recoding: I. Parallel development and coding of neural feature detectors. *Biological Cybernetics*, 1976, 23, p. 121-134.
- [GROS 99] Grossberg, S. The link between brain learning, attention and consciousness. *Consciousness and Cognition*, 1999, 8, p. 1-44.
- [HEBB 49] Hebb, D. O. *The organisation of behaviour*. New York: John Wiley and Sons Inc, 1949.
- [HOW 94] Howard, D. L. Pertinence as reflected in personal constructs. *Journal of the American Society for Information Science*, 1994, 5(3), p. 172-185.

- [KWOK 90] Kwok, K.L. Experiments with a component theory of probabilistic information retrieval based on single terms as documents components. *ACM Transactions on Information Systems*, 1990, 8, p. 363-386.
- [KWOK 95] Kwok, K.L. A Network Approach to Probabilistic Information Retrieval. *ACM Transactions on Information Systems*, 1995, 13(3), p. 324-353.
- [LABE 95] Laberge, D. *Attentionnal Processing*. Cambridge, Mass: Havard University Press, 1995.
- [MEGH 96] Meghini, C., & Straccia, U. A relevance terminological logic for information retrieval. *Proceedings of the 19th Annual ACM Conference on Research and Development in Information Retrieval*, 1996, Zurich.
- [MIYA 90] Miyamoto, S. Information retrieval based on fuzzy associations. *Fuzzy Sets and Systems*, 1990, 38, p. 191-205.
- [NIE 89] Nie, J.Y. An information retrieval model based on modal logic. *Information Processing and Management*, 1989, 25(5), p. 477-491.
- [NIE 96] Nie, J. Y., & Brisebois, M. An inferential approach to information retrieval and its implementation using a manual thesaurus. *Artificial Intelligence Review*, 1996, 10, p. 409-439.
- [NEEL 91] Neely, J. H. Semantic priming effects in visual words recognition : A selective review of current findings and theories. In D. B. G. W. Humphreys (Ed.), *Basic processes in reading : visual word recognition*, (pp. 264-336). Hillsdale, NJ: Erlbaum, 1991.
- [PAVL 27] Pavlov, I. P. *Conditioned reflexes*. Londres: Oxford University Press, 1927.
- [QUIL 68] Quillian, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing*, (pp. 216-270). Cambridge, MA: MIT Press, 1968.
- [RADE 76] Radecki, T. Mathematical model of information retrieval system based on the concept of fuzzy thesaurus. *Information Processing and Management*, 1976, 12, p. 313-318.
- [ROBE 77] Robertson, S.E. The probability ranking principle in IR. *Journal of Documentation*, 1977, 33, p. 294-304.
- [ROBE 82] Robertson, S.E., Maron, M.E. & Copper, W.S. Probability of relevance: A unification of two competing models for document retrieval. *Inf. Tech. Res. Devel.*, 1982, 1, p. 1-21.
- [ROCC 71] Rocchio, J.J., *Relevance feedback in information retrieval*. In Salton Gerard (Ed.) *The SMART Retrieval System Experiments in Automatic Document Processing*, 313-323. Prentice Hall Inc, 1971.
- [SARA 75] Saracevic, T. Relevance: a review of the literature and a framework for thinking on the notion in information science. *Journal of the American Society for Information Science*, 1975, p. 321-343.
- [SMOL 92] Smolensky, P. *IA connexionniste, IA symbolique et cerveau*. In D. Andler (Ed.), *Introduction aux sciences cognitives*, p. 79-106, Paris: Gallimard, 1992.

RSTI- ISI-7/2002. Recherche et filtrage d'information

- [SPER 95] Sperber, D. Wilson, D. *Relevance: Communication and Cognition* (2nd ed.) Oxford : Blackwell, 1995.
- [STAL 68] Stalnaker, R. *A theory of conditionals*. In N. Rescher (Ed.), *Studies in Logical Theory*. Oxford, 1968.
- [TREIS 93] Treisman, A. *L'attention, les traits et la perception des objets*. In D. Andler (Ed.), *Introduction aux sciences cognitives*, (p. 152-191). Paris: Gallimard, 1993.
- [TURT 91] Turtle, H. & Croft, W.B. Evaluation of an inference network-based retrieval model. *ACM Transactions on Information Systems*, 1991, 9, p. 187-222.
- [VAN 86] van Rijsbergen, C. J. A non-classical logic for information retrieval. *The Computer Journal*, 1986, 29(6), p. 481-485.
- [WONG 89] Wong, S.K.M. & Yao, Y.Y. A probability distribution model for information retrieval. *Information Processing and Management*, 1989, 35, 1, p. 39-53.
- [WONG 95] Wong, S.K.M. & Yao, Y.Y. On modeling information retrieval with probabilistic inference. *ACM Transactions on Information Systems*, 1995, 13(1), p. 38-68.