IDIAP
Rapport technique

Un interface de recherche documentaire
I de r, version 2.0

Jean-Luc Cochard

Mai 1993

INSTITUT DALLE MOLLE D'INTELLIGENCE ARTIFICIELLE PERCEPTIVE
CASE POSTALE 609 - 1920 MARTIGNY - VALAIS - SUISSE
TELEPHONE : ++41 26 22.76.64 - FAX : ++41 26 22.78.18
E-MAIL : IDIAP@IDIAP.CH

Numéro : 93-04

Un interface de recherche documentaire
I de r, version 2.0

Mode d'emploi

Jean-Luc Cochard

Institut Dalle Molle d'Intelligence Artificielle Perceptive (IDIAP)
Case postale 609, CH-1920 Martigny, Suisse
Adresse électronique : `cochard@idiap.ch`

Résumé

Ce document présente un interface d'utilisation d'un système prototype de recherche documentaire construit dans le cadre du projet *Specification and Prototyping of a System for the Intelligent Management of Information*[1]. Cet interface permet de soumettre des requêtes de recherche de documents et de visualiser les documents sélectionnés par le système.

D'autre part, cet outil illustre les possibilités d'utilisation d'une boîte à outils graphique, **`Tk`**, et son intégration possible avec Prolog via un utilitaire appelé `expect`.

Cet interface fait partie d'une triade d'outils qui comprend un interface d'indexation documentaire [Coc93] et un interface d'administration de documents indexés [Coc].

Mots-clé : interface homme-machine, interface graphique, indexation et recherche documentaire, traitement de la langue naturelle.

[1] Projet N° 4023-26996, PNR 23, FNSRS

Table des Matières

1 Introduction

L'objectif de ce document est de fournir un mode d'emploi des commandes disponibles dans cet outil de recherche documentaire : **I de r 2.0**.

Cet outil permet de faire une recherche dans une base de données de documents préalablement indexés par **I d'i 2.0** [Coc93]. La recherche se fait sur une base multicritère en fournissant des indications sur le contenu de certains champs du document. Comme les documents indexés sont des lettres administratives relativement bien structurées, une requête consiste à décrire schématiquement le contenu de l'*entête* de la lettre.

Le résultat d'une recherche dans la base de données est une liste de documents satisfaisant l'ensemble des critères de recherche. Les documents repérés sont symbolisés par leur cote d'indexation et un icone décrivant le genre du document. Le contenu du document peut être affiché par simple sélection de la cote d'indexation.

Les possibilités de cet outil sont relativement réduites. **I de r 2.0** fait partie d'un environnement prototype en constante évolution. Cet interface sert à la fois à illustrer certains principes novateurs en informatique documentaire, principalement dans son mode de fonctionnement *démo* mais il est aussi un outil de mise au point pour les développeurs qui peuvent aisément vérifier la qualité ou les défauts des index linguistiques.

La suite de ce document décrit les différentes opérations implantées dans cet interface de recherche documentaire. Dans la section 2, nous présentons la structure générale de l'interface ainsi qu'un scénario simple d'utilisation de ses possibilités. La section 4 explique le fonctionnement de la commande de recherche documentaire et la section 5 l'environnement de présentation du contenu des documents repérés. La section 6 présente les particularités du mode de démonstration qui fait apparaître des résultats intermédiaires lors du traitement de la recherche linguistique. La section 7 explique la manière de quitter le programme. La section 8 décrit le format de la commande Unix avec ses paramètres d'exécution, ainsi que les fichiers de description de la langue de l'interface. Et finalement, en conclusion, dans la section 9, nous décrivons quelques améliorations possibles du produit.

2 Le tableau de bord

L'interface de recherche documentaire, **I de r 2.0**, se présente sous la forme d'un "tableau de bord" (cf. Figure 1) constitué de cinq zones avec de haut en bas :

l'entête – cette zone identifie l'application par son nom et son icone, mais n'offre aucune possibilité d'interaction;

la selection du type et la sélection chronologique – cette zone de sélection permet de restreindre la recherche sur des sous-ensembles de documents identiques par leur forme; elle permet aussi de déterminer la "fenêtre" chronologique de la recherche;

les critères textuels – cette zone de sélection qui peut être constituée de plusieurs composantes (cf. la détermination du type à la section 3.1) dans lesquelles il est possible d'introduire du texte libre;

les messages d'exécution – tous les messages à l'intention de l'utilisateur sont centralisés dans cette zone anonyme;

les commandes – elles sont au nombre de quatre et seront décrites en détail dans les sections suivantes.

Un scénario simple d'utilisation de **I de r 2.0** se déroule de la manière suivante :

1. le lancement de l'interface à l'aide d'une commande Unix (cf. section 8.1);

2. la constitution d'une requête en fournissant les critères de recherche (cf. section 3);

3. le lancement d'une commande de recherche documentaire par la commande "**Rechercher**" qui applique séquentiellement les critères de recherche sur la base de données et qui restitue la liste des documents satisfaisant tous les critères (cf. section 4);

4. la présentation des documents respectant les critères de recherche (commande "**Présenter**..." décrite dans la section 5);

5. la fin d'une session qui fait suite à une éventuelle itération des étapes précédentes (commande "**Quitter**..." décrite dans la section 7).

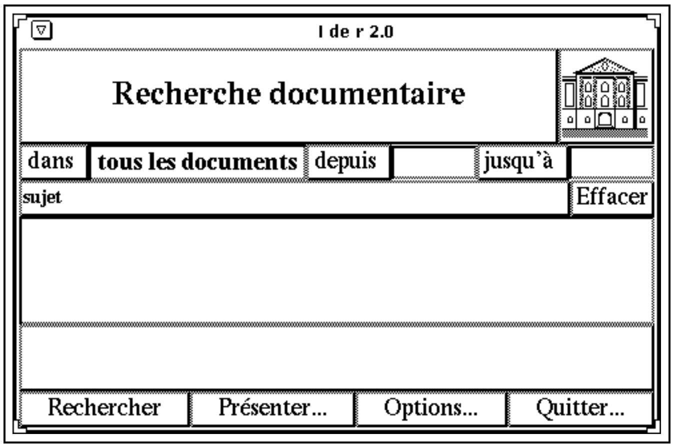

Figure 1 Le tableau de bord de l'interface **I de r 2.0**.

3 La constitution des critères de recherche

La partie centrale du tableau de bord de **I de r 2.0** sert à fournir les critères de recherche dans la base de données des documents indexés. Chacune des zones d'interaction est réservée à la saisie d'un critère précis : le type de lettre, l'adresse du destinataire, la fenêtre chronologique et le sujet du document.

3.1 Le type de document

La sélection du type de document est contrôlée par un menu. Lorsque l'utilisateur sélectionne la valeur courante, le menu apparaît. Lorsqu'il sélectionne une valeur dans ce menu, la valeur courante est remplacée par la nouvelle et le menu disparaît. La Figure 2 présente la liste des types de documents reconnus actuellement.

La sélection d'un type de document a une influence visible sur l'organisation de l'interface. Lorsque l'utilisateur choisit une des possibilités "**les lettres**" ou "**les circulaires**", l'interface fait apparaître un nouveau champ dans les critères textuels permettant d'indiquer le destinaire du document (cf. Figure 3).

Ce critère de sélection est mis en correspondance avec le type de lettre qui a été identifié automatiquement durant l'étape d'indexation des documents. Cette liste va s'étendre au fur et à mesure que la grammaire des structures documentaires décrit de nouvelles formes de documents.

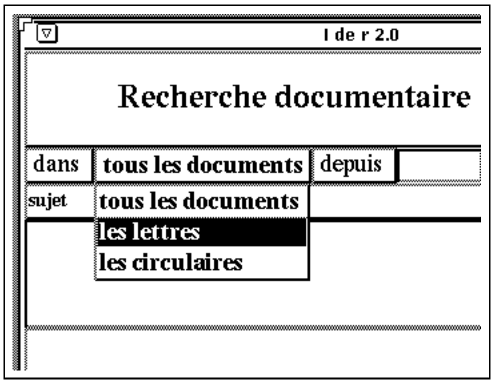

Figure 2 Le menu de sélection du type de document contient deux éléments spécifiques : "**les lettres**" pour les lettres adressées à un destinataire personnel et "**les circulaires**" pour les lettres sans destinataire précis ou adressées à un groupe de personnes. Le troisième élément "**tous les documents**" qui est la valeur par défaut permet des recherches dans l'ensemble complet de la base de données.

Figure 3 Présentation du tableau de bord de **I de r 2.0** lorsque l'utilisateur décide de limiter sa recherche aux seules lettres.

3.2 La fenêtre chronologique

La date précise de rédaction d'un document n'est pas un critère de recherche intéressant. Nous avons plutôt opté pour la spécification d'un intervalle de temps, en années. L'identification des candidats se fait simplement en contrôlant que l'année de rédaction de la lettre fait bien partie de la période mentionnée.

La fenêtre chronologique est définie par les années qui constituent le début et la fin de l'intervalle. La Figure 4 fournit un exemple d'une période valide. Si une des valeurs n'est pas fixée par l'utilisateur, l'intervalle devient infini sur la borne en question. À fortiori, si aucune des deux valeurs n'est fixée par l'utilisateur, la période de recherche va de $-\infty$ à $+\infty$, ce qui revient à dire que ce critère de recherche n'impose plus aucune contrainte sur la sélection de candidats.

L'édition est contrôlée dans les zones de saisie des années si bien que seuls des chiffres peuvent y apparaître. Actuellement, il n'y a aucune limitation sur le nombre de chiffres n'y aucune possibilité de mettre des nombres négatifs, mais nous ne pensons pas que ces faiblesses de l'interface portent à conséquence.

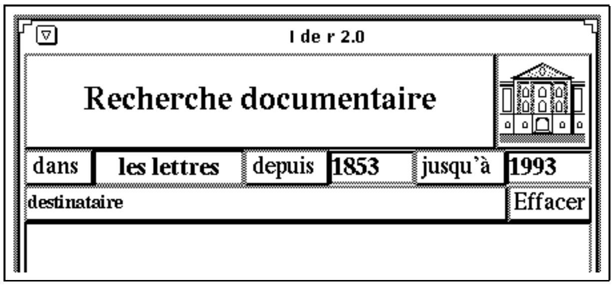

Figure 4 Exemple d'une recherche de lettres sur une période de recherche longue mais finie !

3.3 Les critères de recherche textuels

3.3.1 Le sujet du document

La zone de saisie du sujet du document est primordial dans cet outil de recherche documentaire puisque l'indexation linguistique du sujet des documents a été l'objet principal du projet de recherche mentionné dans l'introduction de ce mode d'emploi. C'est la raison pour laquelle cette zone est toujours présente, quelque soit le type de document sélectionné.

Aucune contrainte syntaxique ou formelle n'est imposée à la saisie de ce critère de recherche. Les quelques commandes d'édition suivantes facilitent la saisie de ce texte :

Positionnement du curseur d'édition – le positionnement se fait par sélection à l'aide de la souris;

Insertion de caractères – l'insertion se fait à la position courante d'édition si le curseur de la souris est positionné dans la zone en question;

Effacement d'un caractère – les touches `Delete` et `Back Space` du clavier effacent le caractère à gauche du curseur d'édition;

Marquage d'une chaîne de caractères – le marquage qui consiste à repérer une chaîne de caractères (affichée en inverse vidéo), est effectué en gardant le bouton de sélection de la souris enfoncée durant un déplacement vers la gauche ou vers la droite;

Effacement d'une marque – la combinaison de touches `Control-D` permet d'effacer une marque;

Effacement du texte – la combinaison de touches `Control-N` permet d'effacer tout le texte de la zone d'édition;

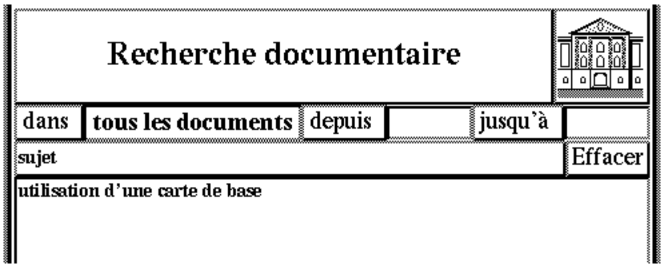

Figure 5 Exemple de sujet accepté et traité correctement par notre système de recherche documentaire.

À noter que cette dernière commande d'effacement du texte est aussi disponible à travers le bouton de commande "**Effacer**". La Figure 5 présente un exemple de sujet constituant un critère de recherche valide pour **I de r 2.0**.

3.3.2 Le destinataire de la lettre

La zone de saisie du destinataire du document permet de fournir tout ou partie de l'adresse. Dans cette version, l'identification des candidats, sur la base de ce critère, est extrêmement simple : le système interprète l'adresse fournie comme une conjonction de mots. Ces mots sont mis en correspondance avec le texte litéral de l'adresse des documents indexés sans que l'ordre d'énumération dans le critère de recherche soit forcément respecté.

Pour l'édition de ce critère de recherche, les commandes disponibles sont identiques à celles décrites ci-dessus (cf. section 3.3.1) et, comme pour le sujet, aucune contrainte formelle et syntaxique n'entrave la liberté de l'utilisateur. Le seul problème d'utilisation de ce critère est la sensibilité aux variations orthographiques. Ainsi la mise en correspondance entre “peuplier” et “peupliers” ou bien “Peuplier” et “peuplier” échouent dans les deux cas.

4 La commande de recherche documentaire

La commande de recherche documentaire “**Rechercher**” lance le processus de sélection de documents répondants aux critères de recherche définis préalablement. La recherche s'effectue en deux étapes :

1. une recherche qui s'appuie sur l'index structurel des documents en vérifiant les contraintes imposées par le type de document, l'adresse du destinataire et la fenêtre de recherche;

2. la recherche linguistique qui prend en considération le sujet et ne travaille que sur les documents sélectionnés dans la première étape.

Deux messages d'exécution distincts fournis durant le processus de recherche précise l'étape en cours de d'exécution (cf. Figure 6). La deuxième étape se décompose en deux sous-étapes : l'analyse linguistique du sujet de la recherche et la mise en correspondance de la clé de recherche linguistique avec les clés d'indexation linguistique.

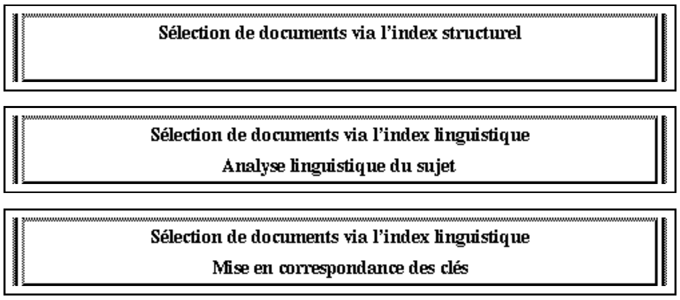

Figure 6 Les trois messages affichés par l'interface lors du processus de recherche.

La première étape se termine par un message de réussite, même lorsqu'aucun document ne répond aux critères structurels de la recherche. La deuxième étape peut se terminer soit sur une réussite, soit sur un échec, soit malheureusement ne pas se terminer du tout. Cette dernière alternative extrêmement fâcheuse est due à un problème dans la mise en correspondance des clés qui n'a pas été corrigé. Pour sortir d'une situation critique de ce genre, veuillez prendre note de la procédure de terminaison forcée décrite dans la section 7.

Remarque : Dans cette version, la réussite des différentes étapes ne signifie pas forcément que des documents ont été sélectionnés. Elle doit plutôt être interprétée comme une indication que le traitement s'est fait sans problème.

5 La présentation des documents

Lorsque la recherche est terminée, il est possible de consulter les documents qui respectent l'ensemble des critères de recherche présents sur le tableau de bord. C'est le rôle du bouton de commande "**Présenter...**". Lorsque cette commande est activée, l'interface affiche la fenêtre de "**Visualisation des documents indexés**". Cette fenêtre est constituée de deux zones : à gauche, la liste des documents avec une indication de la cote d'indexation et un icone précisant le genre de document; à droite, l'éditeur de textes qui affiche le contenu d'un document.

Actuellement, la base de données documentaire manipulée par **I de r 2.0** ne contient que des documents ASCII. Il est cependant prévu d'étendre ce genre de documents à des pseudo-documents. Par exemple, des documents scannés pour lesquels une entête descriptive aurait été fournie manuellement afin d'en permettre l'indexation.

La consultation d'un document se fait par sélection de sa cote d'indexation. Le contenu du document apparaît alors dans l'éditeur de textes (cf. Figure 7).

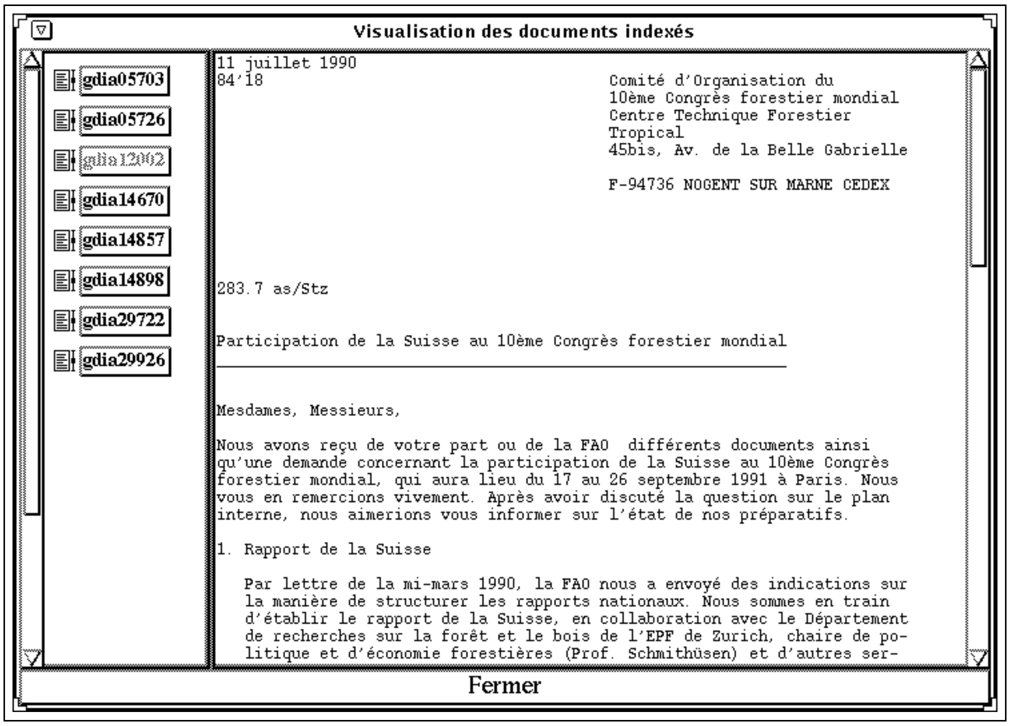

Figure 7 Exemple d'affichage d'un document ayant la cote "**gdia12002**", en réponse à une recherche basée sur le sujet : "le 10ème congrès forestier".

Le bouton de commande "**Fermer**" (cf. Figure 7) permet de faire disparaître la fenêtre de visualisation des documents indexés.

6 Le mode de démonstration

Le mode de fonctionnement de démonstration est une alternative au mode de fonctionnement de la commande "**Rechercher**" décrit dans la section 4, qui est le mode par défaut. Pour activer le mode de démonstration, il faut lancer la commande "**Options...**" et sélectionner le mode "**démo**" (cf. Figure 8).

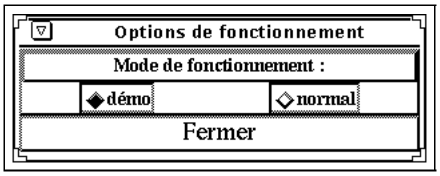

Figure 8 Présentation de la fenêtre d'"**Options de fonctionnement**" avec le mode "**démo**" enclenché.

Dans l'état actuel de cet environnement de démonstration, une seule étape de la recherche est illustrée graphiquement : l'analyse linguistique du sujet. La recherche à travers les index structurels n'est pas illustrée de même que la mise en correspondance des clés linguistiques d'indexation et de recherche.

Remarque : L'illustration de l'analyse linguistique du sujet de la recherche utilise les mêmes outils d'interaction graphiques que ceux décrits dans le mode d'emploi de **I d'i 2.0** [Coc93]. Ce qui suit n'est donc qu'un rappel.

Dans son mode de démonstration, lorsque le processus de recherche est lancé via la commande "**Rechercher**", la première étape de recherche à travers l'index structurel s'exécute normalement. Par contre, lors de la deuxième étape, le message d'exécution est légèrement différent de celui de la Figure 6 : "**Illustration de la sélection de documents via l'index linguistique**". Lorsque l'analyse linguistique est terminée, les résultats sont présentés graphiquement.

Les possibilités d'interaction que doit offrir un interface d'affichage de résultats linguistiques sont en prise directe avec le modèle d'analyse linguistique et plus particulièrement syntaxique mis en œuvre dans ce système. Idéalement, un analyseur syntaxique ne devrait fournir qu'un résultat, celui qui décrit la structure syntaxique sous-jacente de la phrase qui est analysée. Pratiquement, et particulièrement dans notre approche de l'analyse syntaxique, nous nous éloignons de cet idéal selon deux axes distincts : premièrement, la stratégie d'analyse indéterministe fournit plusieurs résultats distincts acceptables compte tenu des connaissances internes du système et, deuxièmement, le processus aboutit souvent à des analyses partielles. Dans ce deuxième cas, le résultat final de l'analyse est une séquence d'arbres syntaxiques.

La fenêtre d'"**Analyse linguistique du sujet**" prend en considération tous ces paramètres. La structure générale de son contenu est la suivante (cf. Figure 9) avec de haut en bas :

le texte à analyser – cette zone affiche pour mémoire le sujet du document qui a été analysé;

les arbres syntaxiques – cette zone est découpée verticalement en autant de fenêtres de résultats qu'il y a de constituants partiels dans le résultat de l'analyse;

les ascenseurs – ils permettent de commander un déplacement horizontal et vertical dans la fenêtre de résultats *courante*;

les commandes d'agrandissement – le bouton de commande "**Taille d'origine**" et le potentiomètre à sa droite, permettent de faire varier la largeur de la fenêtre de résultats *courante*;

les commandes principales – ces commandes permettent d'afficher des résultats d'alternatives d'analyse syntaxique — commandes "**Précédent**" et "**Suivant**" — et de poursuivre le traitement à l'aide des commandes usuelles — "**Continuer**" et "**Continuer+Fermer**".

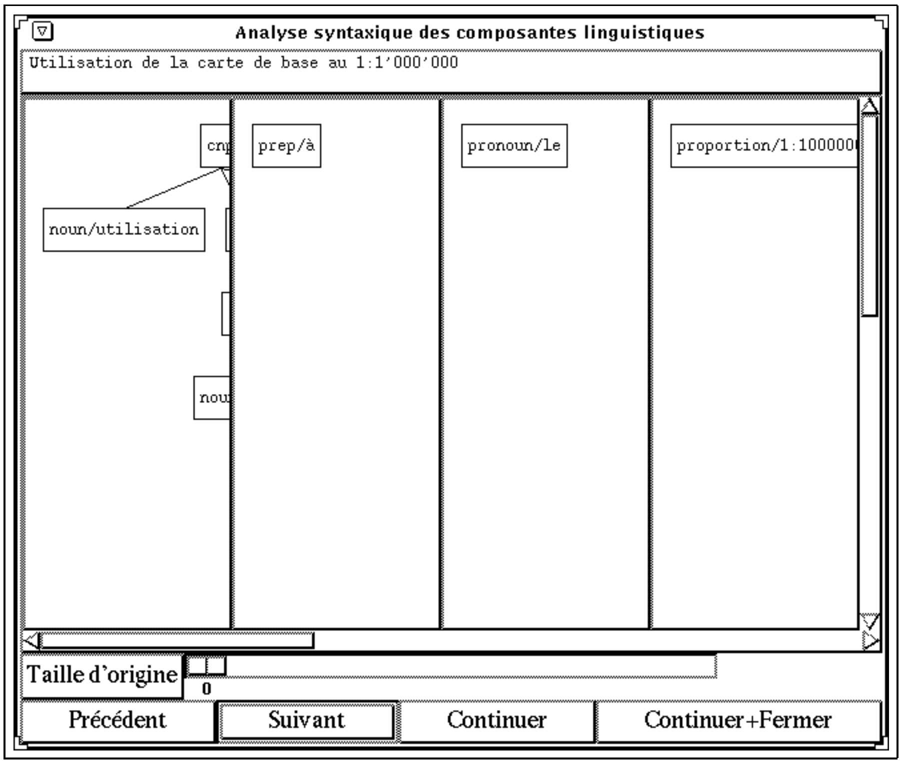

Figure 9 Présentation du résultat de l'analyse de "`Utilisation de la carte de base au 1:1'000'000`". Dans ce cas, l'analyseur n'a pas réussi à produire une analyse syntaxique complète de la phrase. Le résultat est composé de quatre analyses partielles, la première recouvre le texte "`Utilisation de la carte de base`", la deuxième, le texte "`à`", résultat de la décomposition de "`au`" en "`à le`", la troisième, le texte "`le`" et la quatrième, le texte "`1:1'000'000`".

Le contenu des nœuds de l'arbre syntaxique est extrêmement réduit par rapport aux informations utiles à leur constitution. Mais ce choix délibéré reflète très précisément les données que l'analyseur syntaxique transmet à l'analyseur sémantique responsable de la création de l'index linguistique. Il s'agit donc essentiellement de la catégorie morpho-syntaxique et, pour les feuilles de l'arbre, cette information est complétée par la forme de citation du mot.

Les opérations élémentaires qui peuvent être effectuées dans cette fenêtre sont les suivantes : détermination de la fenêtre de résultats courante, déplacement dans cette fenêtre, modification de la taille de cette fenêtre et affichage de solutions alternatives dans cette fenêtre.

La détermination de la fenêtre courante

La détermination se fait simplement en sélectionnant une des fenêtres avec la souris. Les deux conséquences visibles de cette opération sont : la fenêtre sélectionnée se place en retrait par rapport aux autres; sous le curseur du potentiomètre, apparaît la largeur de cette fenêtre.

Le déplacement dans la fenêtre courante

Le déplacement dans cette fenêtre est possible à l'aide des ascenseurs, lorsque la détermination de la fenêtre courante a eu lieu.

La modification de la taille de la fenêtre courante

Les deux éléments d'interaction qui permettent de modifier la taille de la fenêtre courante sont le bouton de commande "**Taille d'origine**" et le potentiomètre sur sa droite. En fait seule la largeur de la fenêtre peut être modifiée (cf Figure 10). Cette largeur peut varier entre 0 et 600 points. La sélection d'une largeur à l'aide du potentiomètre peut se faire, soit en faisant glisser le curseur — le résultat est désagréable car le contenu est réaffiché par à-coups — soit en sélectionnant un point de la zone de déplacement de ce curseur — ce qui ne provoque qu'un seul réaffichage de la fenêtre.

Le bouton de commande "**Taille d'origine**" ne sert qu'à redonner la taille originale à la fenêtre courante qui a été calculée par le programme.

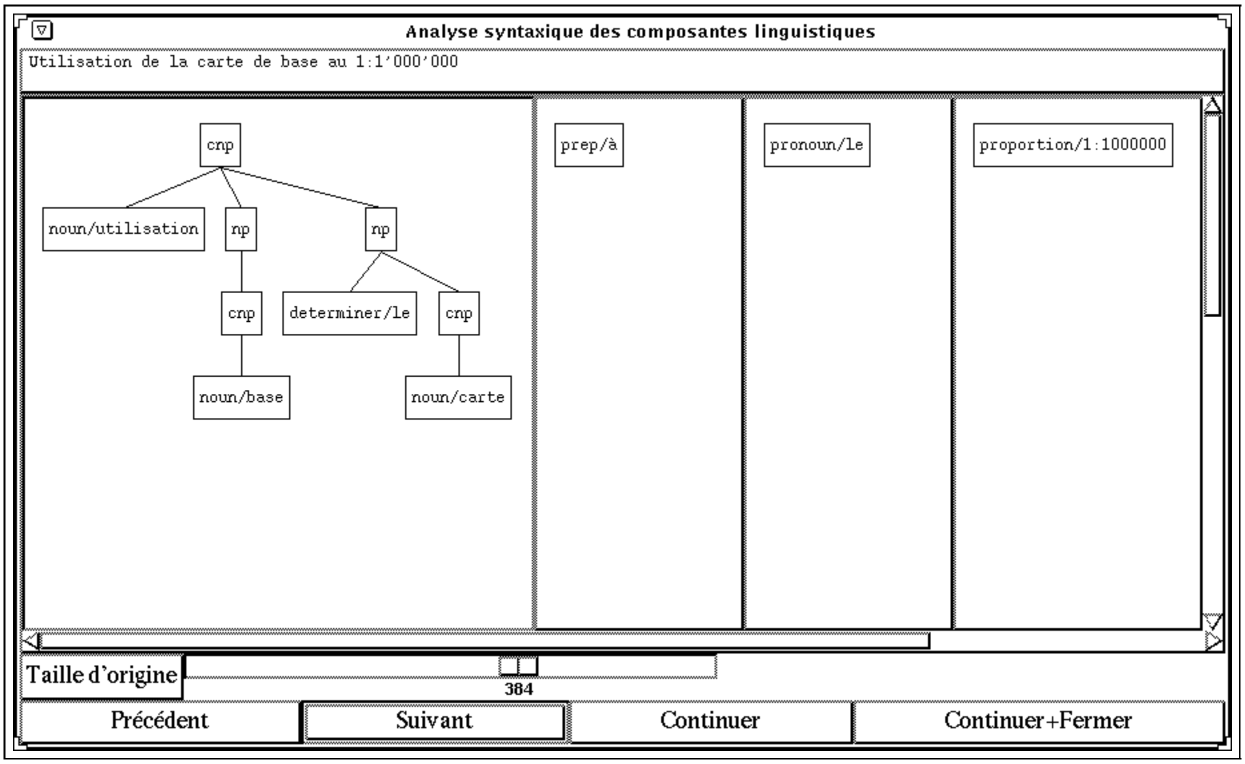

Figure 10 Présentation des résultats partiels d'analyse après édition de la largeur de certaines fenêtres de résultats.

L'affichage des solutions alternatives

Plusieurs résultats d'analyse peuvent être associés à une fenêtre de résultats donnée. Les boutons de commandes "**Précédent**" et "**Suivant**" (commande par défaut) servent précisément à parcourir cette liste dans une direction ou dans l'autre.

Remarque : Dans l'ensemble des interfaces de ce système de gestion des documents informatisés, la mise en évidence des commandes par défaut associées à des boutons est visualisée systématiquement par une mise en retrait du bouton et la touche `Return` est toujours la touche d'activation de cette commande.

7 La terminaison du programme

7.1 La terminaison normale

La commande "**Quitter...**" du tableau de bord permet de mettre fin à une session de travail. Une confirmation est demandée avant de tuer tous les processus impliqués dans ce traitement (cf. Figure 11). La combinaison de touches `Control-C` est une autre solution pour mettre fin à une session. Elle donne lieu à la même demande de confirmation.

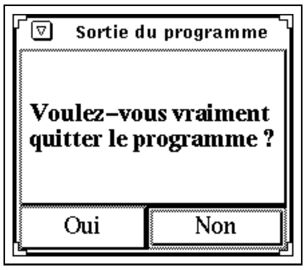

Figure 11 Fenêtre de confirmation de la commande de fin de session. Le dialogue est modal, ce qui oblige l'utilisateur à répondre à la question posée avant de faire autre chose, et la réponse par défaut est "**Non**".

7.2 La terminaison forcée

Lorsque la recherche tarde à fournir des résultats alors que le message "**Mise en correspondance des clés**" est affiché, il y a de fortes chances que le processus soit entré dans une boucle infinie. La seule solution actuellement disponible pour sortir de cette impasse est de tuer le programme que vous avez fait démarrer. À cet effet, vous allez faire un `Control-C` dans la fenêtre qui a servi au démarrage de **I de r 2.0**.

8 L'organisation logicielle

Pour pouvoir utiliser **I de r 2.0**, il y a un certain nombre de prérequis plus ou moins simples à satisfaire :

- disposer d'une SPARCstation de SUN Microsystems;
- disposer du logiciel SICStus-Prolog 0.7 #9[2];
- disposer de l'interprète `wish` du langage Tk[3] [Ous91];
- disposer de l'interprète `expect` (extension de Tcl)[4] [Lib90];
- disposer des programmes `recherche.tcl`, `in_pro` et `ir`[5].

8.1 L'installation de I de r 2.0 et la commande Unix

Pour installer l'interface de recherche documentaire, il faut que les logiciels `prolog`, `expect` et `wish` soient accessibles directement, en complétant éventuellement le contenu de la variable d'environnement `PATH`.

Il faut aussi que le répertoire dans lequel sont centralisés les index et les documents archivés soit repéré par la variable d'environnement `ARCHIVEHOME`.

La commande Unix qui démarre l'environnement démontré dans ce rapport est la suivante :

```
% in_pro recherche.tcl french ir >& /dev/null
```

Comme cet interface est bilingue, français-allemand, il est possible de faire démarrer l'environnement en précisant que les titres et messages doivent être fournis en allemand :

```
% in_pro recherche.tcl german ir >& /dev/null
```

Pour des raisons élémentaires de confort d'utilisation, il est conseillé de créer un "alias" plus concis.

[2] Adresse électronique à contacter : `sicstus_request@sics.se`.
[3] Site pour faire un "ftp" : `sprite.berkeley.edu`.
[4] Site pour faire un "ftp" : `ftp.cme.nist.gov`.
[5] Adresse électronique à contacter : `cochard@idiap.ch`.

8.2 Les opérations internes du lancement de I de r 2.0

Le lancement de l'exécution de **I de r 2.0** est une opération assez longue. Quelques explications sur le chargement et l'interaction des processus de ce système seront utiles pour donner un sens à la commande Unix décrite ci-dessus. En outre, cela permettra à l'utilisateur de patienter en connaissance de cause.

Le programme qui gère le démarrage de tout ce système est `in_pro` qui a deux rôles : le premier, est de lancer l'exécution des programmes qui figurent en paramètres, à savoir `indexation.tcl french` et `ir`; le deuxième, est d'assurer la communication entre ces deux programmes "fils".

L'interface défini dans `recherche.tcl` est un programme de taille modeste qui est assez rapidement chargé. Par contre `ir` est un très gros programme — sa taille actuelle est supérieure à 10 MB — et son chargement prend du temps.

8.3 Les fichiers de langues

L'interface **I de r 2.0** est bilingue, français-allemand. Les noms, titres et messages dans chacune des deux langues sont centralisés dans deux fichiers paramètres : `.RetrieveFrench` et `.RetrieveGerman`. Il s'agit de fichiers texte qui peuvent être édités et modifiés à la convenance de l'utilisateur. Ces fichiers sont des fichiers de ressources dans la terminologie X [QO90] et toute adaptation du contenu doit respecter les règles des fichiers de ressources de X. Pour que **I de r 2.0** trouve ces fichiers paramètres, il est indispensable que chaque utilisateur en ait une copie dans son répertoire `$HOME`.

Le fichier paramètre pour le français a le contenu suivant :

```
*Button.font: *Times-medium-r-*-18-*
*Button.borderwidth: 3

*Label.font: *Times-bold-r-*-14-*
*Menu.font:  *Times-bold-r-*-14-*
*Menubutton.font: *Times-bold-r-*-14-*

*version.text: I de r 2.0
*retrieve_b.text: Rechercher
*display_b.text: Présenter...
*options_b.text: Options...
*quit_b.text: Quitter...
*clean_b.text: Effacer

*yes_b.text: Oui
*no_b.text: Non
*close_b.text: Fermer
*continue_b.text: Continuer
*close_continue_b.text: Continuer+Fermer
*next_b.text:           Suivant
*previous_b.text: Précédent
*default_size_b.text: Taille d'origine


*doc_type_t.text:       dans
*any_document_t.text: tous les documents
*normal_letter_t.text: les lettres
*circular_letter_t.text:les circulaires

*from_year_t.text:      depuis
*to_year_t.text: jusqu'à
```

```
*dest_addr_t.text: destinataire
*subject_t.text: sujet

*main_title.text: Recherche documentaire

*quit_question.text: Sortie du programme
*quit_mess.text: Voulez-vous vraiment quitter le programme ?

*mess.initialisation.text: Chargement du système
*mess.done.text: terminé
*mess.init_failed.text: Erreur d'initialisation !
*mess.start_failed.text: Erreur au démarrage !
*mess.succeeded.text: a réussi !
*mess.failed.text: a échoué !
*mess.structuralQuery.text: Sélection de documents via l'index structurel
*mess.linguisticQuery.text: Sélection de documents via l'index linguistique
*mess.illustrated_linguisticQuery.text: Illustration de la sélection de documents \
via l'index linguistique
*mess.parsing.text: Analyse linguistique du sujet
*mess.querying.text: Mise en correspondance des clés

*gdi_doc_browser.text: Visualisation des documents indexés

*options.text: Options de fonctionnement
*fonctionnement.text: Mode de fonctionnement :
*demo_flag.text: démo
*normal_flag.text: normal
```

9 Commentaires et conclusion

I de r 2.0 a été développé sur des SPARCstations SUN et fonctionne dans l'environnement OpenWindows de SUN. L'interface graphique a été réalisé à l'aide de la boîte à outils graphique Tk; l'indexation est prise en charge par un programme écrit en SICStus-Prolog 0.7 #9 et la communication entre ces deux applications est réalisée à l'aide d'une extension de Tcl, appelée Expect.

Cette application en collaboration avec **I d'a 1.0** et **I d'I 2.0** préfigure ce que pourrait être un environnement d'indexation automatique et de recherche documentaire dans un cadre de travail comme les Archives fédérales suisses ou tout autre institution qui manipule de grosses bases de données textuelles.

9.1 Problèmes connus de l'interface

Dans sa version actuelle, **I de r 2.0** souffre d'un certain nombre de problèmes de jeunesse qui seront corrigés prochainement. Voici pour information la liste des problèmes déjà recensés :

- Rendre plus intelligente la recherche à l'aide du destinataire en introduisant une indexation morphologique du contenu de l'adresse et une recherche comparable;
- La commande de recherche n'est pas assez fiable car elle entre fréquemment dans une boucle infinie. Ce comportement est dû à un bug dans le programme de mise en correspondance des clés;
- Il existe une confusion au niveau des messages d'exécution sur l'interprétation à donner au message "**a réussi**";
- Il manque un indicateur du nombre de documents actuellement sélectionnés par le système.

Toute suggestion d'amélioration ou compte rendu de problèmes sont les bienvenus. Ils peuvent être envoyés à l'adresse électronique figurant dans l'entête de ce document.

Références

[Coc] Jean-Luc Cochard. Un interface d'administration de documents indexés, i d'a 1.0. Rapport technique, IDIAP, Institut Dalle Molle d'Intelligence Artificielle Perceptive. À paraître.

[Coc93] Jean-Luc Cochard. Un interface d'indexation documentaire, i d'i 2.0. Rapport technique RT-93-03, IDIAP, Institut Dalle Molle d'Intelligence Artificielle Perceptive, mai 1993.

[Lib90] Don Libes. The expect user manual – programmatic dialogue with interactive programs. Nist ir, National Institute of Standards and Technology, November 1990.

[Ous91] John K. Ousterhout. An X11 toolkit based on the Tcl language. In *Proceedings of the 1991 Winter USENIX Conference*, 1991.

[QO90] Valerie Quercia and Tim O'Reilly. *X Window System User's Guide for X11 R5*. O'Reilly and Associates, Inc., 1990.