LE PROJET NEIMO : UNE PLATE-FORME MAGICIEN D'OZ MULTIMODALE

Daniel Salber

Laboratoire de Génie Informatique, IMAG
B.P. 53 X, 38041 Grenoble Cedex
Tél.: +33 76 51 44 40 — E-mail: salber@imag.fr

Résumé

Le Magicien d'Oz est une technique d'évaluation expérimentale qui consiste à faire simuler par un "compère" humain dissimulé des services manquants dans un système informatique. Jusqu'à présent, cette technique a principalement été utilisée pour l'étude des interfaces en langage naturel. Nous proposons d'étendre cette technique à l'étude de l'interaction multimodale. La simulation de l'interaction multimodale par Magicien d'Oz pose un certain nombre de difficultés spécifiques. Nous exposons ces difficultés et montrons quelles solutions nous nous proposons d'y apporter. Nous décrivons aussi Neimo, une plate-forme Magicien d'Oz multimodale en cours de développement.

1. Introduction

Le système Neimo (Nouvelle Evaluation des Interfaces par Magicien d'Oz) est une plate-forme d'expérimentation Magicien d'Oz multimodale, générique et extensible. L'objectif de ce projet est de fournir aux concepteurs d'interfaces multimodales un outil d'évaluation permettant aussi bien l'observation du comportement des utilisateurs, qu'une analyse a posteriori d'une session d'expérimentation.

Une expérimentation Magicien d'Oz consiste à faire simuler par un "compère" humain les fonctions non implémentées d'un logiciel, utilisé par un opérateur qui ignore la présence du compère. Jusqu'à présent, les systèmes Magicien d'Oz ont généralement été utilisés pour simuler la reconnaissance du langage naturel, parlé ou écrit, le plus souvent pour des applications de consultation de bases de données [1, 2, 3]. Avec Neimo, nous proposons une extension du Magicien d'Oz intégrant l'interaction multimodale.
L'intégration de l'interaction multimodale pose des difficultés spécifiques. Nous détaillons ces difficultés, puis nous montrons les solutions apportées dans Neimo. Nous décrivons ensuite l'architecture logicielle du système Neimo.

2. Spécificités d'un Magicien d'Oz multimodal

2.1. Le pôle compère

La simulation du langage naturel, donc d'une seule modalité (cas d'un système Magicien d'Oz classique) se heurte à des difficultés lorsqu'il n'y a qu'un seul compère pour assurer la simulation [5]. En particulier, les temps de réponse deviennent vite rédhibitoires pour

l'opérateur qui utilise l'application simulée. Dans le cadre de l'interaction multimodale, l'opérateur peut utiliser plusieurs dispositifs d'entrée de données, éventuellement simultanément. Le volume d'informations à traiter pour un compère qui doit répondre aux actions de l'utilisateur peut vite devenir trop important. Il faut donc que plusieurs compères puissent travailler ensemble pour effectuer la simulation. Nous parlons alors de "pôle compère". On peut envisager par exemple que chacun des compères simule soit l'usage d'une modalité non implémentée dans l'application étudiée (par exemple la reconnaissance de la parole), soit une fonction manquante de l'application étudiée. Le système doit être suffisamment souple pour permettre une reconfiguration facile des rôles des compères. Ceci permettra d'étudier différentes configurations possibles du pôle compère, et d'en tirer des règles pour organiser le pôle compère de façon optimale, suivant l'expérience à effectuer (nombre et type de fonctions et de modalités à simuler, combinaison de modalités, …). Cette reconfiguration doit aussi pouvoir être effectuée en cours de session pour répartir dynamiquement la charge de travail entre les différents compères.

2.2. Outils d'analyse

Lors d'une session d'expérimentation Magicien d'Oz, le comportement de l'opérateur observé est enregistré pour analyse ultérieure. Ces données représentent un volume d'informations très important [4]. Leur analyse a posteriori est donc un travail long et difficile. Nous souhaitons fournir un outil facilitant cette tâche. Notre objectif à long terme est l'automatisation partielle ou totale des activités d'analyse.

2.3. Généricité

La plupart des systèmes Magicien d'Oz développés jusqu'à présent répondent à un besoin d'étude particulier et ne sont pas réutilisables. De notre point de vue, un Magicien d'Oz multimodal doit être ouvert et générique, à la fois vis-à-vis des modalités utilisées et des logiciels étudiés. Nous voulons pouvoir ajouter facilement une nouvelle modalité et greffer sur la plate-forme Magicien d'Oz n'importe quel logiciel.

3. Description de Neimo

Nous avons exposé les spécificités d'un Magicien d'Oz multimodal et les principales difficultés qui y sont liées. Neimo est une plate-forme Magicien d'Oz multimodale qui a été conçue en tenant compte de ces spécificités [6].

Neimo permet d'observer un utilisateur placé en situation d'interaction multimodale, quelles que soient le nombre et la nature des modalités utilisées. Plusieurs compères peuvent intervenir simultanément, chacun traitant par exemple une modalité distincte. Le système est constitué d'un "pôle compère" (plusieurs compères, chacun disposant d'une station), et d'un "pôle opérateur" (opérateur utilisant un logiciel multimodal) [figure 1].

Neimo est aussi un outil d'analyse a posteriori ; des informations sont enregistrées dans un historique lors de la session et peuvent être exploitées par la suite.

Enfin, Neimo est générique ; tout logiciel dont le code source est accessible, moyennant une adaptation au système Neimo, peut se greffer sur la plate-forme Magicien d'Oz. Du côté compère, des applications dédiées permettent l'observation de l'opérateur et la simulation des fonctions absentes. Ces applications sont modulaires et extensibles ; des modules spécifiques à l'étude d'une modalité donnée ou d'une application précise peuvent être facilement ajoutés.

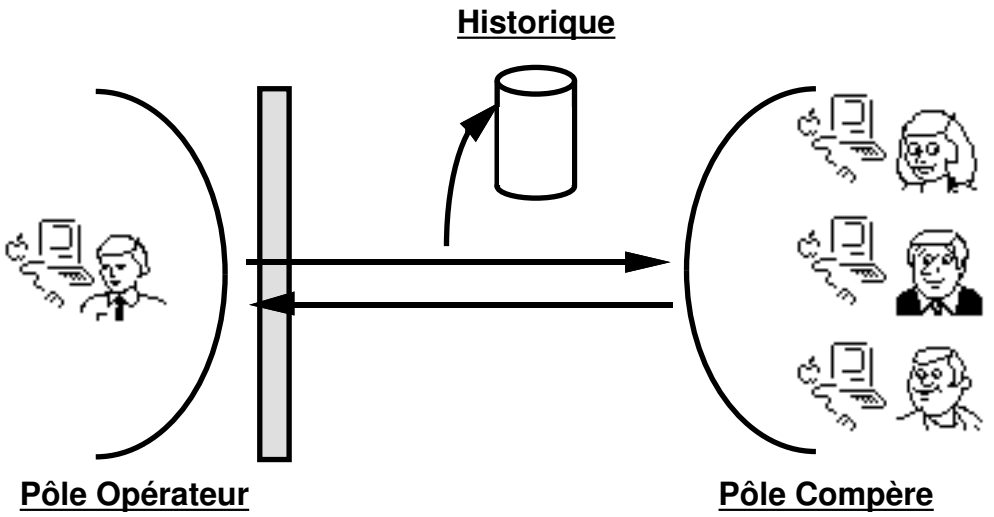

Figure 1. Dispositif expérimental Neimo.
L'opérateur est situé dans une pièce distincte de celle des compères et ignore leur présence. Les compères peuvent être tous situés dans la même pièce ou dans des pièces distinctes.

La configuration actuelle du système Neimo est composée d'un Macintosh Quadra 950 et de deux Quadra 700; Neimo peut fonctionner sur tout Macintosh sous Système 7. Les modalités supportées sont : clavier, souris, geste 2D, parole, expression faciale. La première version du système Neimo est actuellement en fin de développement, en C et C++ avec MPW et MacApp. Un portage sur d'autres plate-formes matérielles et logicielles est à l'étude.

4. Architecture logicielle

Afin de gérer un nombre quelconque de compères, Neimo est architecturé autour d'un noyau de communication, **Neimo Com**, par lequel transitent tous les échanges entre les pôles opérateur et compère. Du fait de sa position centrale, Neimo Com est aussi la base de temps du système et gère la sauvegarde des informations pertinentes dans l'historique. Le processus Neimo Com fonctionne sur n'importe quelle station du pôle opérateur ou du pôle compère, voire même sur une station dédiée. La figure 2 montre l'architecture logicielle du système Neimo.

Les deux pôles communiquent en échangeant des messages structurés porteurs d'informations ; une partie de ces informations est librement déterminée par les applications (application observée et applications compères). La définition de la sémantique de ces informations est à la charge du concepteur des applications. Les informations échangées peuvent donc être de n'importe quel niveau d'abstraction : l'application opérateur peut envoyer dans un message des événements de bas niveau (par exemple des clics souris et des commandes prononcées verbalement par l'opérateur, mais que l'application ne sait pas reconnaître ni interpréter) et demander aux compères de les interpréter ; l'application peut aussi envoyer des informations de niveau plus élevé (par exemple, si l'on souhaite tester un reconnaisseur de parole, l'application observée peut envoyer la phrase reconnue et l'interprétation qui en a été faite, et demander confirmation

aux compères ; en cas de mauvaise interprétation, les compères peuvent pallier la défaillance de l'application observée en simulant l'interprétation correcte).

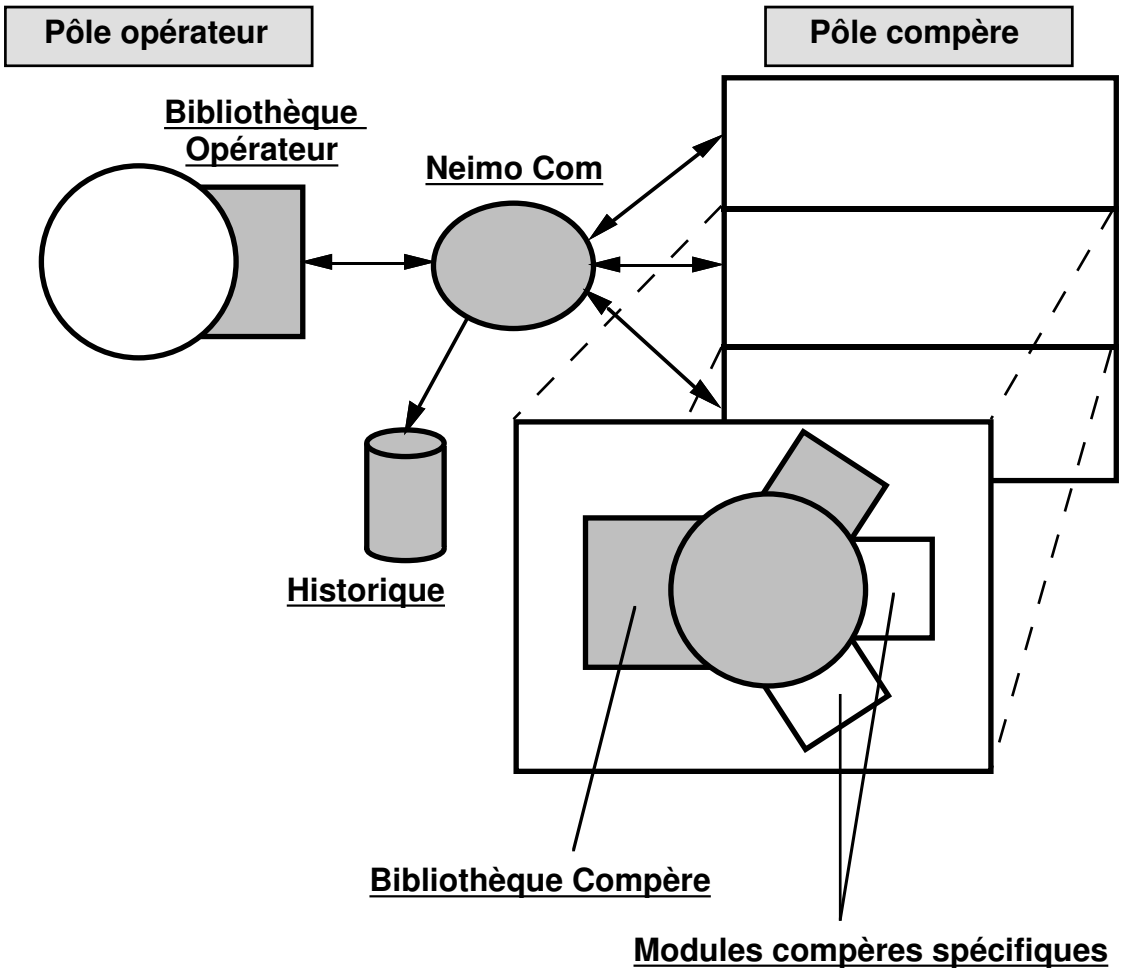

Figure 2. Architecture logicielle du système Neimo.
Les parties grisées sont les composants logiciels de Neimo ; les parties blanches sont des modules spécifiques externes à Neimo.

Afin de garantir la généricité et l'extensibilité du système vis-à-vis des modalités et des applications observées, les applications opérateur et compère communiquent avec Neimo Com à travers une bibliothèque de communication. Le système Neimo proprement dit n'a pas connaissance de la sémantique des informations échangées par les applications.

La **bibliothèque opérateur** fournit au développeur de l'application observée des fonctions permettant :

- l'initialisation et la terminaison d'une session,
- l'envoi d'un message en direction du pôle compère. Lors de l'envoi d'un message, le compère destinataire n'est pas désigné explicitement, seule la compétence requise pour traiter le message est spécifiée ; Neimo Com aiguille ensuite le message vers le compère adéquat,
- l'enregistrement dans l'historique d'une information ou d'un message destiné aux compères,
- la réception de messages en provenance du pôle compère.

La **bibliothèque compère** offre des services symétriques pour le pôle compère ; elle permet en plus aux compères de s'abonner dynamiquement à un type de messages. Cet abonnement est dynamique afin de permettre une reconfiguration du pôle compère en cours de session ; en effet, un compère donné n'est pas forcément utile pendant toute la session, et les compères doivent éventuellement pouvoir changer de rôle en cours de session. La bibliothèque compère permet également aux compères d'échanger des informations, par exemple pour traiter des commandes multimodales combinées.

Par ailleurs, les applications compère sont architecturées de façon modulaire : chaque compère dispose des modules qui lui sont nécessaires pour remplir son rôle. De nouveaux modules peuvent être développés et facilement intégrés à une application compère, par exemple pour observer une nouvelle modalité, ou bien pour simuler une fonction manquante spécifique à une application opérateur donnée. D'autre part, des services communs sont accessibles à tous les compères (informations sur l'état du système, trace des messages provenant du pôle opérateur, …).

Un autre aspect du système est l'exploitation de l'historique pour analyse après la session. Lors de la session, les informations qui transitent par Neimo Com sont sauvegardées à la demande de l'application observée (par programme) ou des compères. Neimo inclut une **bibliothèque d'accès à l'historique** permettant la relecture de ces informations. Cette bibliothèque est destinée à être utilisée par une application d'analyse de la session a posteriori. Pour permettre une analyse efficace du fichier d'historique, la bibliothèque contient un mécanisme de filtrage complet, qui permet une lecture sélective des données. Il est possible de filtrer les informations de l'historique suivant l'une ou plusieurs de leurs caractéristiques : date de l'information, provenance (pôle opérateur ou pôle compère), type de l'information, … Ce mécanisme permet de choisir facilement les informations intéressantes, et ainsi de faciliter le travail d'analyse.

5. Conclusion et perspectives

Nous proposons un outil d'étude de l'interaction multimodale, suffisamment générique et extensible pour pouvoir s'adapter à divers cas d'étude. Cependant, la complexité inhérente à la simulation de l'interaction multimodale rendra probablement délicate la configuration de Neimo, et en particulier le choix du nombre et du rôle des compères ; aussi nous nous proposons d'étudier en détail le fonctionnement du pôle compère, lors d'une première phase d'expérimentations avec Neimo.

6. Remerciements

Nous tenons à remercier Gilles Ambone et Bernard Noz, qui ont participé à la conception et au développement de Neimo. Ce travail n'aurait pu être mené à bien sans la collaboration de Jean Caelen de l'Institut de la Communication Parlée, Grenoble, de Jean-Marc Francony et Evert Kuijpers du CRISS, Grenoble, ainsi que de René Amalberti et Claude Valot du CERMA.

Bibliographie

1. C. Baber et R.B. Stammers: "Is it natural to talk to a computer ? An experiment using the Wizard of Oz technique", proceedings of the Ergonomics society's Annual Conference: Ergonomics-designing Progress 1989.

2. Nils Dahlbäck et Arne Jönsson: "Talking to a computer is not like talking to your best friend", SCAI-88, Scandinavian Conference on Artificial Intelligence, proceedings pp. 53-68.

3. Nils Dahlbäck et Arne Jönsson: "Empirical studies of discourse representations for natural language interfaces", Fourth Conference of the European Chapter of the ACL, proceedings 291-8 1989.

4. Dan Diaper: "The Wizard's Apprentice: A Program to Help Analyse Natural Languages Dialogues", proceedings of the fifth conference of the British Computer Society, Human-Computer Interaction Specialist Group, University of Nottingham, 5-8 September 1989, pp. 231-243.

5. Yolla Polity, Jean-Marc Francony, Rosalba Palermiti, Pierre Falzon, Salma Kazma: "Recueil de dialogues homme-machine en langue naturelle écrite", les cahiers du Criss, n° 17, Décembre 1990.

6. G. Ambone, B. Noz, D. Salber: "Projet Neimo, Spécifications Externes", rapport technique interne, équipe IHM, LGI-IMAG, 1992.