

MATIS : un système multimodal d'information sur les transports aériens

Laurence Nigay
CLIPS-IMAG

B.P. 53 - 38041 Grenoble Cedex 9
Laurence.Nigay@imag.fr

RÉSUMÉ :

MATIS est un système multimodal d'information sur les transports aériens. Il fournit, en réponse à des requêtes de l'utilisateur, des informations sur les vols entre deux villes. Le système a été développé sur machine NeXT avec le système de reconnaissance de la parole Sphinx. MATIS autorise des énoncés de commandes tels que la phrase dite "*I would like a Usair flight from this city to this city*" conjuguée à deux sélections-souris pour spécifier les villes de départ et d'arrivée. MATIS a permis une étude sur l'architecture logicielle des systèmes multimodaux ainsi que sur l'intégration des modalités : mécanisme de fusion des événements et références multimodaux.

MOTS CLÉS : Modalité, fusion, parallélisme.

PRÉSENTATION DE MATIS

MATIS [3] (Multimodal Airline Travel Information System) est un système d'information sur les transports aériens qui a été développé sur machine NeXT en utilisant Interface Builder pour la partie graphique et Sphinx [2], un système de reconnaissance de la parole continue et multilocuteur. MATIS fournit, en réponse à des requêtes de l'utilisateur, des informations sur les vols entre deux villes. La base de données sous-jacente contient actuellement neuf villes américaines, neuf compagnies aériennes et des informations sur chaque vol telles que le numéro du vol, les repas servis à bord etc. L'utilisateur exprime des requêtes au système afin de planifier son voyage. Pour cela, l'interface en entrée de MATIS offre les deux métaphores introduites par D. Norman [5] : la métaphore conversationnelle et la métaphore du monde réel.

Dans le cas de la métaphore conversationnelle, l'utilisateur a le choix entre :

- le quasi langage naturel oral ou écrit , comme "*I would like to know the flights from Pittsburgh to Boston serving a meal*" ; il est à noter que MATIS offre un dialogue finalisé restreint à un domaine sémantique délimité pour lequel l'utilisateur utilise la langue naturelle restreinte : le domaine couvre toutes les informations sur des transports aériens. Ce dialogue est de type question/réponse simple où les demandes et les informations transmises sur un vol sont exprimées en langue naturelle mais où l'enchaînement des requêtes et de négociation n'existent pas ;
- ou le langage artificiel écrit directement dans les champs d'une requête, comme "PIT" qui est le code de la ville Pittsburgh.

D'autre part, l'utilisateur peut spécifier une requête par manipulation directe avec la souris (métaphore du monde réel). En effet MATIS offre une interface graphique accessible via la souris qui représente l'univers de l'action par une métaphore du monde réel. Chaque champ d'une requête fait l'objet d'un formulaire accessible par une fenêtre à l'écran qui présente les outils ; un formulaire permet la spécification par manipulation directe d'une valeur d'un champ particulier d'une requête.

Enfin l'utilisateur peut combiner les langages offerts et la manipulation directe. Par exemple l'utilisateur énonce la phrase "*Flights from Pittsburgh to this city*" tout en sélectionnant avec la souris la ville de Boston dans la liste des villes ; il peut alors préciser sa demande en saisissant au clavier "1500" dans le champ de l'heure de départ de la requête.

Nous présentons l'interface en sortie et en entrée de MATIS dans les deux paragraphes suivants. Ceci nous permet d'introduire les possibilités offertes à l'utilisateur par le système. Le développement de MATIS a permis une étude sur l'architecture logicielle des systèmes multimodaux ainsi que sur l'intégration des modalités. Les résultats de cette étude font l'objet de notre conclusion.

INTERFACE EN SORTIE DE MATIS

Pour mieux appréhender les possibilités offertes à l'utilisateur, nous décrivons ici l'interface en sortie de MATIS. Celle-ci est uniquement graphique et se compose de plusieurs fenêtres. La fenêtre "Office Manager" met en évidence les systèmes connectés (c.-à-d., MATIS) au système de reconnaissance de la parole. Cette fenêtre affiche aussi la dernière phrase reconnue par le système de reconnaissance de la parole, ou les messages d'erreur en cas d'échec de la reconnaissance. C'est aussi la zone d'édition des phrases écrites en langage naturel. Le système offre un ensemble d'outils ("Requests Tools") pour la spécification d'une requête par manipulation directe, tel que la liste des villes. L'usage de MATIS a montré que ces outils étaient aussi utilisés comme compléments à la parole, l'utilisateur ne sachant pas quelles étaient les informations disponibles dans la base de données. Des outils d'utilité générale sont aussi fournis : un bloc-notes et un historique des requêtes qui permet en particulier de réitérer une requête au moyen des opérations Copier/Coller. La fenêtre "Results of request" affiche sous forme de table le résultat d'une requête soumise à la base de données. Les informations contenues dans la table peuvent être sélectionnées et copiées dans une autre fenêtre, comme celle du bloc-notes.

La fenêtre "Request" contient la requête en cours de spécification. Comme les champs d'une requête sont nombreux, nous les avons divisés en deux catégories selon leurs fréquences d'utilisation. Les champs obligatoires ou très utilisés sont affichés par défaut à l'ouverture d'une nouvelle requête (c.-à-d., Ville de départ, Ville d'arrivée, Heure de départ, Heure d'arrivée).

Une requête vide est affichée par défaut au lancement de MATIS. MATIS offre un dialogue à plusieurs fils d'activité : il est possible de commencer une nouvelle requête sans achever la requête courante. Deux requêtes sont alors en cours de spécification. Pour rendre active une requête, il suffit de sélectionner la fenêtre correspondante. Celle-ci est alors affichée au premier plan et définit le nouveau contexte de l'interaction.

INTERFACE EN ENTRÉE DE MATIS

A l'opposé de son interface en sortie assez pauvre, MATIS a été conçu de façon à offrir de nombreuses possibilités à l'utilisateur dans sa tâche de spécification d'une requête. Au travers des exemples ci-dessous nous tentons de montrer la richesse des possibilités. Conjugué à un dialogue à plusieurs fils d'activité (plusieurs requêtes en cours de spécification) MATIS offre :

1- L'usage synergique des langages et dispositifs que nous illustrons par l'exemple suivant : la phrase orale "*Flight from Pittsburgh to this city*" accompagnée de la sélection de "*Boston*" avec la souris.

2- L'usage synergique des dispositifs uniquement, comme l'exemple suivant : la phrase orale "*Flight from Pittsburgh*" accompagnée de la saisie au clavier de la phrase "*servng a meal and arriving in the afternoon*". Le langage utilisé est ici le même.

3- L'usage alterné des langages et dispositifs : ceci correspond au cas 1 avec séquençement des actions. Par exemple l'utilisateur articule une phrase puis sélectionne avec la souris.

4- L'usage alterné des dispositifs : ceci correspond au cas 2 avec séquençement des actions. Par exemple l'utilisateur commence par articuler le début d'une phrase et la finit par saisie au clavier.

5- L'usage concurrent des langages et dispositifs, comme l'exemple suivant : l'utilisateur spécifie une requête en la dictant en langage naturel, tout en complétant une autre requête par saisie directe dans le formulaire. L'usage concurrent permet la spécification de plusieurs requêtes en parallèle.

6- L'usage concurrent des dispositifs uniquement, comme l'exemple suivant : l'utilisateur spécifie une requête en la dictant en langage naturel, tout en saisissant au clavier une autre requête en langage naturel.

7- L'usage exclusif des dispositifs et langages : typiquement MATIS peut être utilisé comme un système vocal uniquement.

L'interface en entrée de MATIS a été complètement spécifiée de façon formelle en utilisant les langages UAN (User Action Notation) et LOTOS dans [1].

CONCLUSION

L'expérience de MATIS a montré que le modèle d'architecture PAC-Amodeus [3] fournit un cadre de

réalisation pour le traitement en parallèle et la fusion de données en entrée. La fusion des données est effectuée au niveau du composant Contrôleur du Dialogue et s'appuie sur la hiérarchie d'agents PAC qui organise ce composant. Le moteur de fusion [4] manipule des informations représentées dans un formalisme unique, le creuset. Ce formalisme est indépendant du domaine de l'application. De même le moteur de fusion s'appuie sur des critères tels que le temps et la structure de l'information à fusionner ; il est donc indépendant du domaine considéré. Son comportement peut être modifié en injectant de nouveaux critères de fusion. Afin d'obtenir un retour d'information le plus immédiat possible, sa stratégie est précoce et implique parfois de dé-fusionner. Mais le moteur de fusion, et par conséquent le système, ne se bloquent jamais en attente d'un événement complémentaire. Cette stratégie permet en particulier la gestion d'un dialogue à plusieurs fils d'activité.

Le modèle PAC-Amodeus offre aussi une grande flexibilité. Par exemple, il est possible de changer ou de rajouter une modalité d'interaction au système MATIS en ne modifiant que deux composants logiciels, Niveau Bas d'interaction et Techniques de Présentation, le Contrôleur de Dialogue étant indépendant des modalités.

Notre travail s'oriente vers la modélisation de l'incertitude au sein du mécanisme de fusion pour résoudre complètement les expressions déictiques. De plus nous avons validé le modèle PAC-Amodeus et notre moteur de fusion générique pour la réalisation des systèmes multimodaux en entrée. Le problème de la fusion ou de la fission des informations en sortie par exemple n'a pas été abordé. Notre étude actuelle [3] sur les interfaces en sortie est restée à un niveau conceptuel et doit être approfondie. Notre support pour ces travaux sera le système MATIS.

REMERCIEMENTS

MATIS a été développé à Carnegie Mellon University (CMU) à Pittsburgh. Je remercie A. Rudnicky, de CMU pour m'avoir accueillie dans son équipe à deux reprises.

BIBLIOGRAPHIE

1. Coutaz, J., Paterno, F., Faconti, G., Nigay, L. A Comparison of Approaches for Specifying MultiModal Interactive Systems. *Workshop ERCIM on Multimodal Human-Computer Interaction*, (Nov, Nancy), 1993.
2. Lunati J.M. et Rudnicky A.I. Spoken Language Interfaces: The OM system. *Actes de CHI'91 Human Factors on Computing Systems*, (27 Avril - 2 Mai, New Orleans), ACM Press, 1991, pp. 453-454.
3. Nigay, L. Conception et modélisation logicielles des systèmes interactifs. Thèse de l'Université Joseph Fourier, Grenoble, 1994, 350 pages.
4. Nigay, L. Réalisation logicielle des systèmes multimodaux : un moteur de fusion générique. *Actes de IHM'94 (8 et 9 Déc., Lille)*, 1994, pp. 15-21.
5. Norman, D. A., Draper S. W. User Centered System Design. *Lawrence Erlbaum Associates Publishers*, 1986, 526 pages.

