

de connaissances pour répondre à la requête d'un ou de plusieurs utilisateurs simultanément. On accède à *Galaxy* par téléphone ou par ordinateur, avec une connexion au réseau *Internet*, qui transmet les données.

Galaxy assure cinq fonctions : la reconnaissance de la parole, la compréhension du langage, l'extraction de l'information, l'élaboration du langage et la synthèse vocale. Quand on pose une question, un serveur nommé *Summit* analyse ce qu'il entend à l'aide d'une bibliothèque de phonèmes (les plus petites unités sonores qui constituent les mots, dans une langue), puis il dresse une liste ordonnée de phrases possibles. Un deuxième serveur, nommé *Tina*, applique les règles grammaticales, décompose la phrase en éléments (sujet, verbe, complément...) et traduit alors l'énoncé le plus plausible en un bloc sémantique, c'est-à-dire une série de commandes que la machine comprend. Par exemple, *Tina* reformulera la question : « Où se trouve le musée d'Orsay ? » en la commande suivante : « Localiser le musée nommé musée d'Orsay. »

Ensuite, un troisième serveur, *Genesis*, convertit le bloc sémantique en une requête adaptée au format de la banque de données la plus pertinente. *Tina* réorganise alors l'information extraite en un nouveau bloc sémantique, que *Genesis* convertit en une phrase dans la langue de l'utilisateur : « Le musée d'Orsay se trouve au numéro 1 de la rue de la Légion d'honneur, à Paris, dans le VII^e arrondissement. » Enfin, un synthétiseur vocal dit la phrase.

Pour passer d'un système à l'autre, l'utilisateur n'a qu'à demander : « Je veux parler à *Jupiter* » ou : « Passe-moi *Voyager*. » Depuis mai 1997, *Jupiter* a répondu à plus 30 000 appels, et le taux de compréhension des requêtes d'utilisateurs novices atteint 80 pour cent.

Le projet MASK

Dans les halls de gare, des billetteries automatiques à interface tactile sont destinées à faciliter l'accès de la clientèle au système de distribution. Malheureusement, la rigidité et le manque de convivialité de ce système n'ont pas séduit les voyageurs. Depuis 1994, dans le cadre d'un programme de recherche européen, en collaboration avec le CNRS, l'Université de Londres et les Sociétés *Mors* et *Signes particuliers*, la SNCF met au point une borne d'information et de distribution qui, à terme, remplacera les billetteries actuelles.

Il s'agit d'un kiosque de réservation et d'achats de billets de train à interface naturelle multimodale : la machine répond aux paroles ou au toucher de l'utilisateur par une voix de synthèse, des séquences vidéo ou des textes. Le kiosque, nommé *MASK*, est composé de deux ordinateurs : une station de travail *Silicon Graphics* qui gère les différents modules, ou logiciels, du système, et un ordinateur personnel classique, qui se charge de l'interface graphique. Lorsqu'un utilisateur énonce une requête, un premier module de reconnaissance de la parole, qui intègre les hésitations, traduit les signaux acoustiques en signaux numériques. Un deuxième module extrait de ces signaux les informations requises (date de départ, destination, gare de départ...) qu'un troisième module, le gestionnaire de dialogue, recherche dans une banque d'information, nommée *Riho*. Le lexique répertorie aujourd'hui plus de 1 500 mots, dont les noms de 600 villes françaises. Quand des informations sont encore nécessaires, un module de génération de phrase formule une question qu'un synthétiseur vocal prononce. Lorsque la requête est complète, une liste de trains est proposée sur l'écran, et le voyageur n'a plus qu'à choisir. L'achat d'un billet peut durer moins d'une minute.

Le kiosque a été testé pendant un mois à la gare Saint-Lazare par des clients volontaires qui suivaient un scénario préétabli. On analyse ces essais préliminaires afin d'évaluer la machine en conditions réelles, et d'améliorer ses capacités de compréhension et de dialogue.

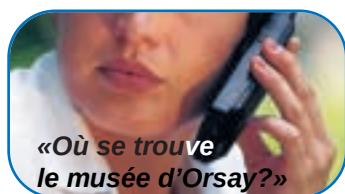


Nous enregistrons les appels, puis nous les analysons, afin d'améliorer les performances du système.

Nos systèmes de reconnaissance de la parole seraient parfaitement adaptés aux appareils portatifs *Handy 21* du projet *Oxygen* : le clavier, notamment, pourrait être supprimé, et le langage parlé permettrait aux utilisateurs de communiquer plus efficacement avec leur appareil. Ainsi, un homme d'affaires en déplacement pourrait demander à son ordinateur : « Dis-moi quand l'action Crédit lyonnais dépassera 210 francs. » La machine deviendra ainsi un assistant qui accomplit des tâches variées à partir d'un minimum d'instructions.

Nous devons encore progresser avant d'atteindre ce but. Notamment, nous devons réaliser des programmes capables de traiter de l'information dans des domaines complexes, nous devons augmenter le nombre de langues comprises par la machine et nous devons obtenir des systèmes qui réunissent des données variées sans y être explicitement invités. L'ordinateur ne doit pas seulement comprendre ce que l'on dit, mais plutôt ce que l'on veut dire.

Victor ZUE est directeur adjoint du Laboratoire d'informatique de l'Institut de technologie du Massachusetts.



« Où se trouve le musée d'Orsay ? »

SUMMIT
comprend
la phrase

TINA
formule
la demande

GENESIS
recherche
l'information

TINA
organise
la réponse



« Le musée d'Orsay est au 1, rue de la Légion d'honneur. »

LES APPLICATIONS À RECONNAISSANCE VOCALE utiliseront l'architecture *Galaxy* pour répondre à la question d'un utilisateur. Le système *Summit* établit une liste d'interprétations de la phrase, puis le

système *Tina* traduit la proposition la plus plausible en instructions qui permettent au système *Genesis* d'extraire l'information d'une banque de données. Enfin, le système *Tina* formule une réponse.

Les caméléons de la communication

JOHN GUTTAG

Nos téléphones mobiles seront bientôt remplacés par des appareils qui seront à la fois des téléviseurs, des messagers de poche, des radios, des ordinateurs et... des téléphones.

Les informaticiens du projet Oxygen cherchent à remplacer nos divers outils de communication par un dispositif unique, portable et polyvalent. Pourquoi en effet utiliser plusieurs appareils qui font presque tous la même chose? Après tout, nous n'utilisons pas une machine particulière pour les traitements de textes, une autre pour les tableurs, une troisième pour le traitement des images et une quatrième pour le courrier électronique. Un seul appareil portable, sans fil, ne pourrait-il, à notre choix, faire office de téléphone cellulaire, de messenger de poche, de radio, de télévision ou d'ordinateur relié au réseau Internet?

Le Handy 21 que nous mettons au point est cet appareil. Au premier abord, le projet semble utopique : comment doter un appareil d'autant de fonctions sans qu'il devienne intransportable? Nous tentons de résoudre ce problème avec le projet Spectrum Ware.

Nous étudions ainsi des systèmes de communication à usages multiples, nommés caméléons de la communication, qui sont programmés pour recevoir et transmettre différents types de signaux. L'utilisateur qui souhaite appeler un correspondant activera le mode téléphone cellulaire ; le même appareil exécutera un programme dif-

férent pour accéder au réseau Internet, pour recevoir des images télévisées, des ondes radio...

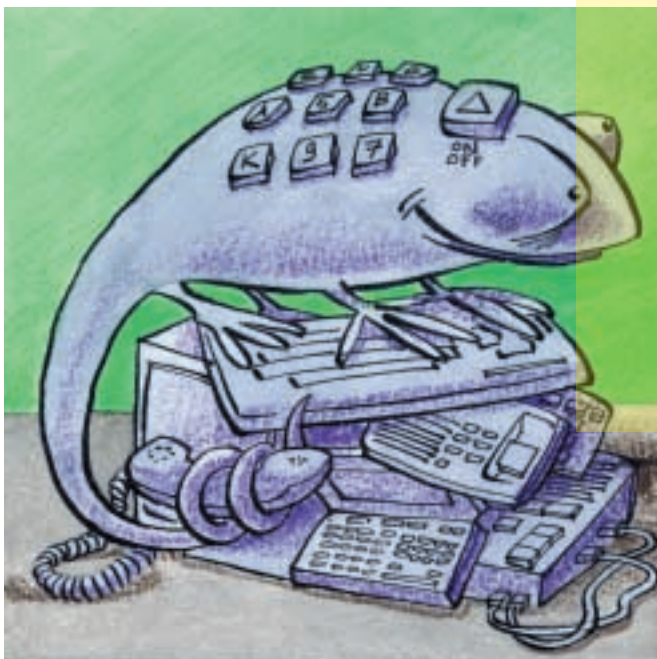
Les programmes, qui remplaceront le matériel électronique trop encombrant et moins flexible, sont les éléments clés de ce projet. Ils feront d'un caméléon de la communication un matériel généraliste, relié à plusieurs applications. Par exemple, dans notre laboratoire, Vanu Bose et Matt Welborn ont mis au point un programme de radio qui fonctionne sur un ordinateur de bureau équipé d'une antenne et d'un amplificateur. Un autre programme, créé par John Ankcorn, fait du même ordinateur une télévision.

Des programmes émettent des signaux

Pour notre radio, nous avons minimisé les traitements analogiques des signaux captés par l'antenne et faisons exécuter par programme le plus de tâches possible. Le matériel capte les ondes électromagnétiques par bandes spectrales d'une largeur de dix mégahertz, et il les transforme en une fréquence intermédiaire, qui est transmise à la mémoire vive d'un ordinateur. Le traitement du signal est effectué par un microprocesseur généraliste, qui utilise un système d'exploitation standard, tel Windows ou Mac OS. Avec un tel système, l'ordinateur assure l'essentiel

de la fonction d'une radio, tout comme il se charge des autres applications. Le combiné, qui contient peu de matériel spécialisé, est facile à moderniser : notre logiciel de radio sera plus rapide sur un ordinateur lui-même plus rapide. De plus, avec une telle architecture, des appareils jugés jusqu'à aujourd'hui incompatibles (tels les téléphones mobiles digitaux et analogiques) ne le seront plus.

Jadis, pour téléphoner, on décrochait le combiné et l'on demandait à une standardiste : «Je voudrais parler à untel.» Avec les nouveaux systèmes, l'utilisateur formulera la même demande à un appareil sans fil, connecté



Jean-Michel Thiriet



Tom Draper Design - Michèle Tcherevkoff, The Image Bank (télévision)

LE HANDY 21 sera à la fois (de gauche à droite) un téléviseur, un messenger de poche, une radio, un téléphone mobile et un système de navigation connecté au réseau *Internet*. L'antenne du dispositif transmettra et recevra des signaux de communication qu'un unique microprocesseur généraliste traitera.

à un réseau. Le système se chargera alors de transférer l'appel vers le récepteur de la personne en question.

Ces systèmes peuvent également transmettre des données : reprogrammé, le même matériel deviendra un appareil médical portable, qui enverra et recevra des échographies ou les signaux d'un électrocardiographe. Les médecins munis de ces appareils auront instantanément accès aux dossiers médicaux et aux résultats des analyses de leurs patients, même loin de leur cabinet.

Les caméléons et Oxygen

Les caméléons de la communication utilisant le système *SpectrumWare* seront sans doute les éléments majeurs de l'infrastructure du projet *Oxygen*. Toutefois, dans un réseau sans fil, tel celui qui relierait les *Handy 21*, l'état des canaux varie de

façon imprévisible. De plus, les diverses applications requièrent des largeurs de bande différentes et tolèrent des taux d'erreurs variables, avec, parfois, des niveaux de sécurité élevés. Par exemple, une application de commerce électronique nécessite un cryptage du signal de communication plus performant qu'une application de radio ou de télévision. Or, les interfaces d'un réseau classique sont fixes : elles sont conçues pour fonctionner dans les pires conditions qui puissent être tolérées, plutôt que pour s'adapter aux conditions réellement rencontrées par le système.

Avec le système *SpectrumWare*, nous voulons exploiter mieux les ressources du réseau en changeant les caractéristiques de la communication pour chacune des applications. Les bornes des réseaux devront alors modifier les canaux en fonction du nombre d'unités mobiles dans la zone

couverte et de leurs besoins spécifiques. Le réseau affectera des canaux personnalisés aux *Handy 21* qui nécessiteront une réponse en temps réel ou des débits d'information élevés, alors qu'il fera partager un même canal à d'autres unités moins exigeantes.

Aujourd'hui, nous construisons les programmes de communication qui équiperont les réseaux du futur. Nous conservons trois idées directrices : nos programmes sont tous généralistes ; nous prévoyons une attribution dynamique des ressources, au lieu de prévoir la pire des situations ; nous faisons exécuter par un programme le maximum de tâches, afin de minimiser la part du matériel.

John GUTTAG est directeur du Département de génie électronique et d'informatique de l'Institut de technologie du Massachusetts.