

Ça vous chatouille?

Le rire provoqué par les chatouilles semble être un réflexe.

Pourquoi rions-nous quand on nous chatouille? Et pourquoi ne rions-nous pas quand nous nous chatouillons nous-mêmes? Ces deux questions ont intéressé les plus grands esprits, de Platon à Darwin, mais des expériences de psychophysiologie font le tri parmi les théories. Notamment des expériences à l'Université de San Diego montrent que les chatouilles déclenchent probablement un réflexe.

Socrate, un des premiers «théoriciens de la chatouille», avait noté le mélange de plaisir et de douleur qu'elles procurent. D'ailleurs, les physiologistes modernes distinguent divers chatouillis. Certains, provoqués par le passage léger d'un corps sur la peau, provoquent l'envie de frotter la peau à l'endroit chatouillé quand la stimulation a cessé. Les chatouilles qui font rire, d'autre part, sont des pressions plus fortes, répétées, sur des zones particulières du corps : sur les côtes, sous les bras, sous les pieds...

Les chatouilles du premier type sont ressenties par les animaux : les chevaux bougent les oreilles quand une mouche les chatouille. En revanche, les chatouilles du second type semblent

l'apanage des primates : la plupart des primatologistes ont vu les chimpanzés et d'autres singes se chatouiller au cours de leurs jeux. Ces animaux ne rient alors pas, parce que leur appareil vocal ne s'y prête pas, mais ils font entendre des sons de respiration qui équivalent au rire humain.

Les chatouilles procurent-elles du plaisir? Cette question aussi est difficile, car les sentiments que nous ressentons sont souvent provoqués par l'expression du visage : quand des personnes reproduisent une émotion, elles ressentent souvent, à un certain degré, l'émotion qu'elles reproduisent. Par exemple, en 1988, des psychologues ont découvert que des personnes jugent des dessins animés plus drôles quand on leur demande de sourire pendant la projection que quand on leur demande de se renfrogner.

Voulant distinguer le rire dû à l'humour et le rire des chatouilles, Christine Harris et Nicholas Christenfeld ont cherché à savoir si celui d'une sorte facilitait celui de l'autre sorte. En 1997, à l'Université de San Diego, ils ont soumis des histoires drôles à des étudiants et observé que, dans une série, les dernières blagues sont plus drôles que les premières. Cet effet de préparation n'a pas lieu avec les chatouilles : les sujets qui ont été préparés par le visionnage de comédies n'ont pas ri plus que des sujets qui n'avaient pas vu les films. Ainsi le rire des chatouilles n'est pas associé à un état d'esprit amusé, de même que nous ne sommes pas

tristes quand nous pleurons en épluchant des oignons.

Si le rire des chatouilles n'a pas de rapport avec la joie ou avec l'amusement, quelle est son origine? Tout d'abord, il ne semble pas apparaître chez les enfants âgés de moins de six mois. Ensuite, il semble pouvoir déclencher un conditionnement. Arthur Koestler, après Darwin, a également signalé que ce rire n'apparaît que si la situation est jugée sans danger par celui qui est chatouillé. Enfin, les psychologues ont observé que si le rire des chatouilles est provoqué par un comportement interpersonnel, on ne pourrait pas se chatouiller soi-même. Les psychologues de San Diego ont testé expérimentalement l'hypothèse : une personne qui saurait qu'elle est chatouillée par une machine rirait-elle?

En 1998, N. Christenfeld et C. Harris ont interrogé un groupe d'étudiants et montré, d'abord, que 50 pour cent d'entre eux pensaient qu'ils ne riraient pas s'ils étaient chatouillés par une machine ; seuls 15 pour cent avaient l'idée opposée. Comme il n'était pas certain que la machine à chatouiller qu'ils utilisaient chatouillerait comme un être humain, les deux psychologues ont fabriqué une machine où seule une main gantée dépasse ; la machine fait entendre divers bruits de respiration, afin qu'on puisse la confondre avec une personne.

On disait aux sujets volontaires (yeux bandés) qu'ils seraient chatouillés deux fois : une fois par un expérimentateur et une fois par la machine. Toutefois, toutes les chatouilles étaient, en réalité, faites par un expérimentateur humain. L'expérience a montré que les sujets riaient autant quand ils croyaient être chatouillés par une machine ou par un être humain. Ainsi le rire des chatouilles ne dépend pas de ce qui chatouille. Ce n'est probablement pas un comportement interpersonnel.

Est-ce un réflexe? Et pourquoi, alors, ne rit-on pas quand on se chatouille soi-même? La question n'est pas résolue, mais on sait qu'on peut provoquer l'extension réflexe de la jambe quand on se tape soi-même sur le genou ; en revanche, les sursauts de peur sont un réflexe que l'on ne peut s'infliger. De fait, les sujets testés à San Diego riaient plus quand ils n'avaient pas les yeux bandés, et des études du cerveau par résonance magnétique nucléaire montrent que les aires du cerveau activées quand on se chatouille soi-même diffèrent de celles qui sont activées quand on est chatouillé par autrui.



D.R.

Si le rire est le propre de l'homme, le phénomène des chatouilles n'en est pas l'apanage. Pourquoi?

La révolution de l'informatique

MICHAEL DERTOUZOS

L'infrastructure informatique nommée Oxygen, en cours de mise au point, facilitera l'accès aux informations réparties dans le monde. Enfin, les hommes pourront en faire plus en en faisant moins.

En 1998, dans l'avion qui nous emmenait à Taiwan, avec quelques collègues, j'essayais de transférer mon agenda sur mon nouvel ordinateur portable : il y avait seulement une carte à insérer et un logiciel à installer. Après trois heures d'énervement, je passais le problème non résolu à mon voisin Tim Berners-Lee, qui m'avait proposé son aide. Après une heure de travail, l'inventeur du réseau *Internet* s'avouait vaincu.

Je sollicitai alors Ronald Rivest, l'inventeur de la cryptographie à clef publique, qui, dans sa grande sagesse, déclina poliment l'offre. Puis, l'un des plus jeunes membres de l'équipe, sûr de lui, s'y risqua... et abandonna bientôt. J'essayai alors n'importe quoi, changeant les paramètres de configuration au hasard... et, par chance, trois heures plus tard, l'installation était faite.

En faire plus en en faisant moins

La majorité des utilisateurs de l'informatique rencontrent ce type de difficultés. C'est la rançon de 40 ans d'une informatique qui s'est surtout préoccupée des ordinateurs : elle conçoit séparément les systèmes et les sous-systèmes, que les utilisateurs doivent ensuite assembler eux-mêmes.

Imaginez que, lorsque vous achetez une voiture, vous deviez régler des dizaines de boutons afin d'assurer l'injection du carburant, l'allumage des bougies, l'écartement des pistons...

Ça suffit ! À l'ère de l'informatique, les utilisateurs ont besoin d'un volant, d'un accélérateur et de freins. Ils doivent – enfin – disposer de techniques qui leur donneront plus alors qu'ils en feront moins.

Ces nouvelles techniques doivent entrer dans nos vies, et pas l'inverse : nous ne devons pas changer nos vies pour plonger dans un cyberspace qui nous est étranger. Lors de la révolution industrielle, les moteurs sont venus à nous dans les réfrigérateurs et les automobiles, sans que nous changions nos mœurs. J'attends des ordinateurs et des systèmes de communication qu'ils participent de la même façon à nos vies et qu'ils n'existent que par des tâches précises qu'ils réalisent pour nous.

L'informatique doit augmenter la productivité humaine, tout en se simplifiant. On devrait pouvoir demander à un ordinateur portable : « Emmène-nous à Athènes ce week-end. » La machine se connecterait au système de réservation d'une compagnie aérienne, elle saurait que « nous » représente deux personnes habituées à la classe économique, aux places côté couloir... En quelques minutes, un vol adapté serait trouvé et confirmé. L'ordre prendrait ainsi trois secondes de temps humain, et l'informatique œuvrerait pendant dix minutes : le gain de productivité serait de 20 000 pour cent.

De tels gains ne seront pas possibles dans toutes les applications, mais je pense qu'au cours du XXI^e siècle on augmentera la productivité humaine de 300 pour cent, en moyenne, en auto-

matissant les activités de routine dans les bureaux et en déchargeant une partie du travail de nos cerveaux et de nos pupilles sur nos bulldozers électroniques (les logiciels), de la même manière que nous nous sommes déchargés des travaux de force sur les bulldozers. La transition n'a pas commencé : nous cliquons encore avec acharnement sur les fenêtres de navigation du réseau *Internet* ou de courrier électronique, nous nous faisons mal aux yeux en scrutant les pages affichées et nous nous donnons des migraines à force de les déchiffrer, sans nous apercevoir que nous sommes encore au stade de la pelle. Nous sommes éblouis parce que ces pelles sont en or et qu'elles sont déclarées « haute technologie ». Nous devons demander aux ordinateurs bien plus que le peu qu'ils nous donnent aujourd'hui.

Le commerce électronique

Les marchands d'ordinateurs abusent de l'expression « facile à utiliser » : ils habillent un singe d'une blouse verte et veulent nous faire croire que c'est un chirurgien ! Je ne parle pas de l'aspect esthétique des ordinateurs ou des logiciels, mais d'une réelle facilité d'utilisation. Les différences entre les logiciels de navigation et les systèmes d'exploitation doivent disparaître rapidement : n'ont-ils pas la même fonction d'accès à l'information, les premiers à distance, les seconds localement ? L'utilisateur, lui, n'a pas à se préoccuper du lieu où se trouve

l'information, et il doit disposer d'un seul type de commandes. Le système actuel est aussi absurde que si le volant d'une voiture guidait le véhicule en ville et le freinait à la campagne.

Enfin, les nouvelles techniques doivent s'adresser à tous les hommes. Aujourd'hui, on se satisfait qu'une centaine de millions d'ordinateurs soient interconnectés : ils ne relient pourtant que 1,7 pour cent de la popu-

lation mondiale ! Les voix de milliards d'êtres humains ne nous parviennent que filtrées par la télévision ou par les propagandes gouvernementales. De plus, sans contrôle, l'informatique creusera le fossé entre les riches et les pauvres, car les riches utiliseront leurs ordinateurs pour devenir plus productifs, donc plus riches, tandis que les pauvres en resteront au même point.

De telles disparités provoquent inévitablement des conflits sanglants. Pourtant, l'informatique peut aider ceux qui en ont besoin : les riches pourraient décider d'acheter des services aux plus démunis, de la même façon que l'on peut décider d'acheter des produits



fabriqués dans des pays en développement (des étudiants de notre laboratoire ont déjà créé un tel système sur le site Internet www.compassioncorps.org). L'aide pourrait être internationale : un médecin du Sri Lanka pourrait donner des consultations à des sans-abri de Boston par l'intermédiaire d'un kiosque équipé d'une caméra vidéo, connecté à des instruments médicaux et administré par une infirmière ; le service, bien qu'imparfait, serait meilleur qu'une absence totale de soins.

Le commerce informatique commence seulement. Dans les dix prochaines années, un demi-milliard de machines et d'innombrables ordinateurs intégrés à des appareils électriques, des détecteurs, à des systèmes de régulation, etc. seront interconnectés. Ces machines et leurs utilisateurs achèteront, vendront et échangeront de l'information et des services informatiques. Aujourd'hui, 300 milliards de francs sont échangés sur l'Internet chaque année ; en 2030, j'estime que les sommes s'élèveront à 25 000 milliards de nos francs actuels, soit un quart des capitaux de l'industrie mondiale. Elles concerneront principalement le secteur des services et de l'informatique, qui représente la moitié de cette économie.

L'échange d'informations via les systèmes informatiques devrait également s'imposer dans nos vies pri-

vées : dans les familles, les activités associatives, l'enseignement, la culture, la politique...

Amener la technique au cœur de nos vies, accroître la productivité humaine et la simplicité d'utilisation des machines, rendre ces gains accessibles à tous : quel système informatique concrétisera une telle vision ?

La conception d'Oxygen

Notre laboratoire étudie un système, radicalement nouveau, de matériels et de logiciels adapté aux hommes et à leurs besoins. En collaboration avec le Laboratoire d'intelligence artificielle de l'Institut de technologie du Massachusetts, 30 membres de notre équipe participent à ce projet, nommé *Oxygen*.

Le *Handy 21* est au cœur du projet. Il ressemble à un téléphone portable, mais il est muni, en outre, d'un écran, d'une caméra, de détecteurs infrarouges et d'un ordinateur. Toujours près de l'utilisateur, n'importe où, il est entièrement configurable par logiciels : une commande suffit pour en faire un téléphone ou un émetteur-récepteur en relation avec les autres *Handy 21*, un nœud de réseau près d'un réseau sans fil à grand débit, ou encore une simple radio (voir *Les caméléons de la communication*, page 40, et *Une puce universelle*, page 42).

La deuxième pièce du système *Oxygen* est l'*Enviro 21*, qui, plus rapide et de plus grande capacité, reste relié à son environnement, tel un bureau, un appartement ou une voiture. L'*Enviro 21* est au *Handy 21* ce que la prise de courant est à la pile. Il fait la même

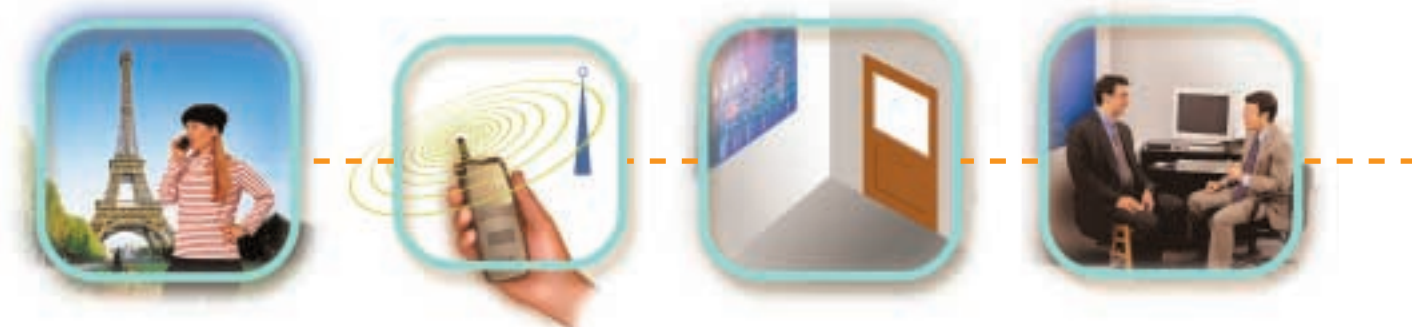
chose que le *Handy 21*, mais il est plus puissant et plus rapide. Il peut notamment commander des appareils variés, tels des détecteurs, des téléphones, des télécopieurs, des réseaux de caméras ou de microphones.

Le système *Oxygen* interagit avec le monde physique grâce à ces appareils contrôlables et à des détecteurs infrarouges placés dans les *Handy 21*. Pour créer un lien entre la porte et vos machines, vous collez sur cette porte un détecteur infrarouge : lorsque quelqu'un pointera son *Handy 21* vers la porte, la machine identifiera cette porte et indiquera ce qui est derrière. Les utilisateurs seront en relation avec les objets de leur environnement.

Le *Handy 21* et l'*Enviro 21* seront reliés grâce à un nouveau réseau, le *Net 21*, qui créera des régions de collaborations sécurisées pour les utilisateurs d'*Oxygen* qui souhaitent communiquer. Le *Net 21* doit donc s'émanciper du réseau Internet, trop encombré et trop vaste : il gèrera les fluctuations (créations et destructions) des nœuds de connexion ; il doit pouvoir vous trouver où que vous soyez ; et il se connectera à de nombreux appareils et aux réseaux du monde entier. Pour mettre au point ce réseau, nous devons construire des protocoles de connexion originaux, intégrant des concepts d'auto-organisation et d'auto-adaptation.

Le système *Oxygen* doit également être un système avec lequel nous interagissons simplement : nous ne naissons pas avec des souris et des claviers, mais avec une bouche, des oreilles et des yeux. Toutes les composantes du système *Oxygen* et toutes

UNE JOURNÉE des utilisateurs du système *Oxygen* est imaginée au travers de ces vignettes : cinq collaborateurs prennent une décision rapide grâce au système qui les trouve, qui assure leurs communications et effectue pour eux des recherches.



1 – À PARIS, JEANNE a repéré un site intéressant pour les bureaux de sa société. L'affaire est à saisir rapidement. Elle demande à son *Handy 21* de joindre son patron, Robert.

2 – LE HANDY 21 DE JEANNE se connecte au réseau local après l'avoir détecté et établit instantanément la connexion avec l'*Enviro 21* situé dans le mur du bureau de Robert.

3 – L'ENVIRO 21 DE ROBERT, dans le mur, reconnaît le *Handy 21* de Jeanne, comprend l'urgence et fait suivre l'appel à Château-Roux, où Robert bavarde avec le représentant local.

4 – LE BUREAU DE CHÂTEAU-ROUX est également équipé d'un *Enviro 21*, qui réceptionne l'appel. Il détecte automatiquement qu'il peut interrompre la conversation.

les applications reconnaîtront la parole et l'utiliseront pour transmettre des informations. Aujourd'hui, les systèmes créés par l'équipe de Victor Zue gèrent une demande de renseignements sur des domaines restreints, telles la météo ou les compagnies aériennes (voir *Une conversation avec votre ordinateur*, page 38). Nous travaillons à réunir ces différents domaines, à incorporer des images et des graphiques pour élar-

gir le champ de la communication entre l'homme et la machine.

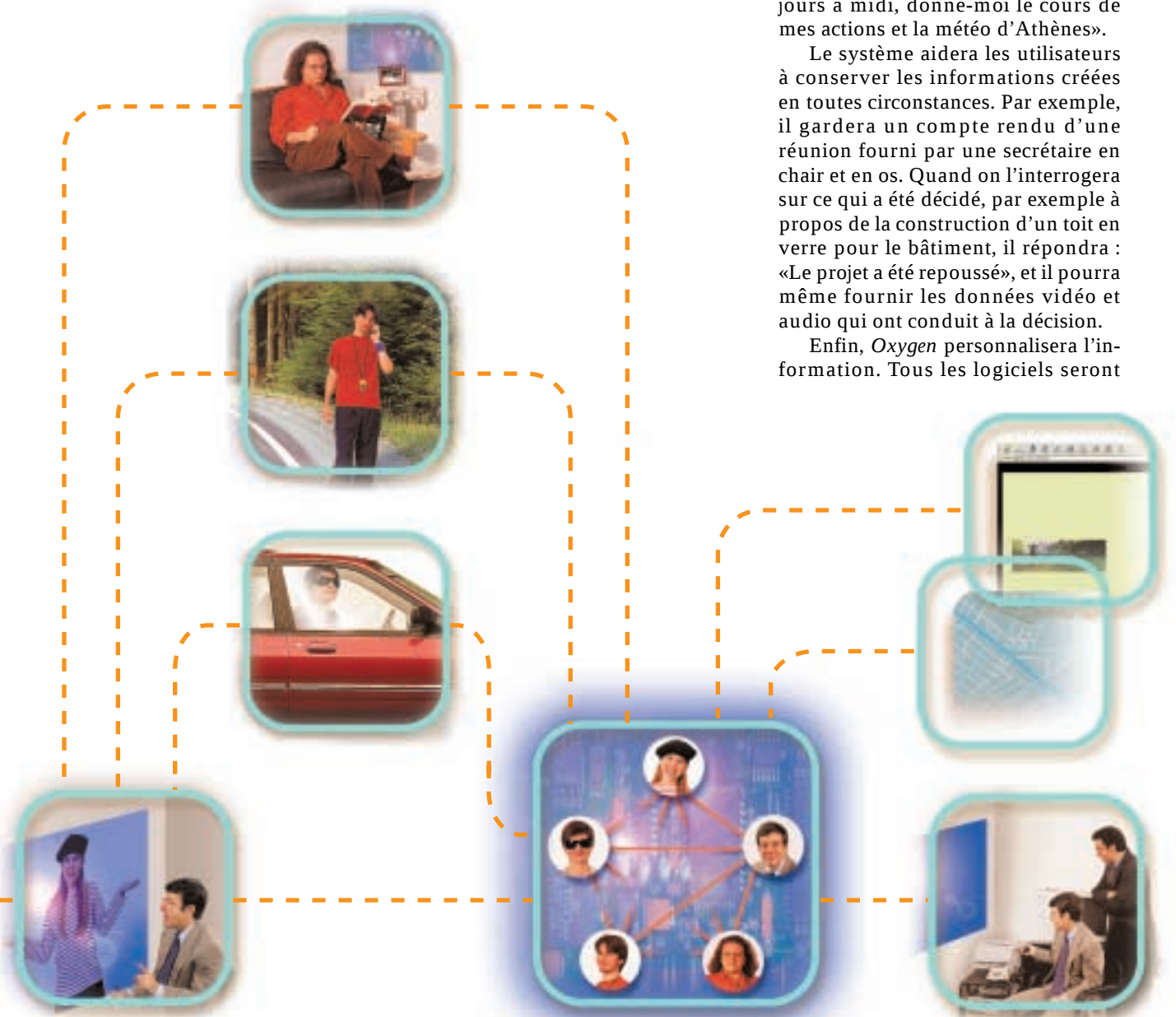
Le cinquième composant d'*Oxygen* permettra d'utiliser des archives personnelles de manière naturelle. On pourra demander au système : «Trouve-moi le gros dossier rouge qui est arrivé il y a un mois», sans fournir de référence, ni d'indexation, et *Oxygen* cherchera également les informations dans les archives des amis ou collègues d'accord pour partager

les données. *Oxygen* fera également des recherches dans le réseau *Internet*, en sélectionnant les réponses selon ce que les utilisateurs ou leurs partenaires ont déjà dans leurs bases de connaissances.

Oxygen se chargera des travaux les plus simples ou les plus répétitifs, et il exécutera des commandes qui automatiseront des tâches, tout en surveillant les nombreux appareils connectés aux *Enviro 21* : «Augmente le chauffage», «Imprime ça», «Tous les jours à midi, donne-moi le cours de mes actions et la météo d'Athènes».

Le système aidera les utilisateurs à conserver les informations créées en toutes circonstances. Par exemple, il gardera un compte rendu d'une réunion fourni par une secrétaire en chair et en os. Quand on l'interrogera sur ce qui a été décidé, par exemple à propos de la construction d'un toit en verre pour le bâtiment, il répondra : «Le projet a été repoussé», et il pourra même fournir les données vidéo et audio qui ont conduit à la décision.

Enfin, *Oxygen* personnalisera l'information. Tous les logiciels seront



5 - L'IMAGE DE JEANNE apparaît sur le mur du bureau de Châteauroux. Elle présente le site à Robert, qui demande : «*Oxygen*, trouve-moi vite Roger, Michel et Marie !»

6 - OXYGEN trouve Roger courant dans les bois, Michel tranquillement installé dans son canapé, un livre à la main, et Marie dans sa voiture, en route pour rejoindre sa famille.

7 - LE NET 21 organise une conférence à cinq en quelques secondes. Lors de leurs échanges, les cinq personnes demandent à *Oxygen* des informations sur ce nouveau site sur le réseau.

8 - «C'EST D'ACCORD», conclut Robert. Il pointe son *Handy 21* vers l'imprimante munie d'un détecteur infrarouge, et donne l'instruction : «*Oxygen*, imprime les documents que nous avons examinés.»

Parler n'est pas suffisant

La communication naturelle ne se limite pas aux échanges oraux. De ce fait, une interaction homme-machine n'est idéale que si la communication est «multimodale», c'est-à-dire si elle intègre la parole, mais aussi les gestes de désignation et la perception visuelle. Lorsque nous parlons, nous faisons souvent référence à des objets de notre environnement, que nous voyons et désignons, par des déterminants ou pronoms particuliers, tels «une», «la», «cette», «celle-ci»...

Naturelle chez les êtres humains, cette multimodalité pose à la machine des problèmes d'interprétation et d'influence réciproque entre les modes d'interaction. Par exemple, avec une interaction associée à deux désignations, telle «place cette figure au-dessus de ce texte», la parole, le geste et la perception ne peuvent pas s'interpréter indépendamment.

Depuis dix ans, notre équipe, composée d'informaticiens et de linguistes, étudie la mise en place d'une telle communication : à partir d'études linguistiques, cognitives et informatiques, nous élaborons tout d'abord des modèles informatiques de représentation des connaissances, puis des algorithmes pour traiter un problème particulier, dont nous testons ensuite les capacités de reconnaissance, de compréhension et d'interprétation sur des applications concrètes. Quand l'algorithme est performant, nous l'intégrons enfin à un logiciel ou à une interface.

Ainsi, un de nos modèles linguistiques intègre et comprend les hésitations, les reprises et les corrections, nombreuses à l'oral et qui, bien que diverses, respectent une structure prévisible. Un autre modèle est spécialisé pour interpréter les gestes de désignation naturelle de l'utilisateur, et un troisième modèle, dit de représentation mentale, unifie l'ensemble des informations connues sur chaque composant d'une tâche. Un autre modèle tient compte des présupposés inclus dans les énoncés d'un utilisateur afin de mieux adapter les réponses de la machine et de gérer le dialogue.

Par ailleurs, nous cherchons des algorithmes qui exploitent les connaissances de ces modèles pour comprendre les énoncés multimodaux. Ainsi un de nos algorithmes interprète les différentes références incluses dans une tâche, notamment quand celle-ci est visualisée sur un écran, et intègre le fonctionnement des déterminants en français. S'appuyant sur un pavage de l'espace cognitif, par exemple, de l'écran, l'algorithme lève les ambiguïtés d'énoncés, tels «la figure», «cette figure», «une de ces figures». Nos algorithmes sont écrits en Java, le langage en vigueur sur le réseau *Internet*, assurant ainsi leur adéquation aux divers systèmes informatiques.

Enfin, nous avons proposé une nouvelle architecture, modulable et répartie, pour les systèmes de dialogues multimodaux, qui s'adapte d'une part au modèle de l'application à gérer, c'est-à-dire à la définition des objets à traiter, à leurs propriétés et aux fonctions qui leur sont applicables. D'autre part, l'architecture s'adapte aux spécificités lexicale, syntaxique, sémantique et d'usage des différents modes de communication ; ainsi, pour le langage, nous traitons la langue soutenue, technique, écrite et orale.

Bien que nos recherches soient encore fondamentales, des applications industrielles sont déjà en cours. Ainsi, avec la Société *Alcatel*, nous avons mis au point une interface de communication, répondant aux gestes et à la voix, qui équipe un terminal de télécommunication composé d'un écran tactile et d'un combiné téléphonique. Nous mettons également au point un environnement multimodal de consultation d'annuaires complexes pour ce type de terminal de télécommunication qui pourrait bientôt apparaître sur le marché.

À l'avenir, seule une communication multimodale nous permettra d'accéder de façon naturelle aux nombreuses applications interactives qu'offrent aujourd'hui le réseau *Internet* et la télévision numérique.

Jean-Marie PIERREL,
Université de Nancy et Équipe
«Langue et Dialogue» du LORIA

transférés directement du *Net 21* sur les *Handy 21* et *Enviro 21*, en fonction des demandes, des erreurs et des mises à jour des utilisateurs. Les utilisateurs adapteront les machines qui les entourent à leurs besoins et à leurs habitudes.

Une conviction et un souhait

Oxygen nécessite des ordinateurs portables, des ordinateurs muraux et des ordinateurs embarqués, un réseau innovant, des systèmes de compréhension de la parole et de synthèse vocale, l'accès aux connaissances, la collaboration, l'automatisation et la personnalisation. *Oxygen* devra sa puissance à l'ensemble de ses éléments, qui sont tous tournés vers l'utilisateur. La communication avec les ordinateurs a commencé avec des nombres binaires. Puis on a utilisé des textes et, récemment, des icônes ; il est temps d'améliorer encore les interfaces homme-machine.

Je crois que l'informatique du futur dépendra essentiellement des échanges vocaux (ou à l'aide d'autres capacités sensorielles), de l'accès aux connaissances, de l'automatisation, de la collaboration et de la personnalisation. Ces cinq aspects, à mon avis incontournables, sont le volant, l'accélérateur et les freins que l'informatique utilisera pour nous en donner plus.

Je propose que tous ceux qui souhaiteront tirer profit du nouveau monde de l'information exploitent les potentialités du système *Oxygen*. Les médecins, par exemple, seraient plus efficaces s'ils disposaient à tout instant des données de *Medline* (qui regroupent des articles médicaux de très nombreuses revues), des dossiers médicaux de leurs patients et des avis de collègues spécialisés.

L'informatique, avec *Oxygen* et d'autres systèmes de même nature, doit maintenant nous libérer des détails techniques que nous supportons depuis 40 ans et nous faire entrer dans une nouvelle ère : celle des ordinateurs au service de l'utilisateur. Après la charrue, le moteur et l'ordinateur, la quatrième révolution concerne la ressource terrestre la plus précieuse : nous-mêmes.

Michael DERTOUZOS est directeur du Laboratoire d'informatique de l'Institut de technologie du Massachusetts.

Conversation avec votre ordinateur

VICTOR ZUE

Comment commander les ordinateurs sans lever le petit doigt.

Les auteurs de science-fiction savent depuis longtemps que le langage parlé est, pour l'homme, le moyen le plus pratique de communiquer avec les machines. Croire qu'ils donnent simplement leur version du mythe de Pygmalion est insuffisant. La parole est surtout naturelle : nous savons parler avant de savoir lire ou écrire. La parole est rapide : nous parlons environ cinq fois plus vite que nous ne tapons à la machine, et peut-être dix fois plus vite que nous n'écrivons. Enfin, le langage parlé est flexible : nous n'avons besoin d'aucun autre support que l'oral pour mener une conversation, les yeux ou les mains sont libres.

La première génération de systèmes intégrant des techniques vocales apparaît aujourd'hui. Certains d'entre eux reconnaissent des dizaines de milliers de mots. Des logiciels de dictée sont déjà fabriqués par plusieurs sociétés : IBM, Dragon Systems, Lernout & Hauspie ou Philips. D'autres systèmes, initialement étudiés par les Laboratoires AT&T Bell, puis mis au point par les Sociétés Nuance, Philips et Speech Works, traitent des discours transmis par téléphone sans étalonnage préalable. Ces techniques, intégrées dans des services d'assistant virtuel, permettent aux utilisateurs d'accéder par téléphone à des informations de presse, à des cotes en Bourse ou au courrier électronique. Cependant, le projet Oxygen requiert des systèmes de reconnaissance de la parole plus efficaces.

Avec les systèmes de la nouvelle génération, les hommes communiqueront avec les ordinateurs presque aussi facilement qu'entre eux. Les ordinateurs ne se limiteront pas à comprendre ce qu'on leur dit, mais devront tenir des conversations : les dispositifs classiques de reconnaissance vocale, qui convertissent des signaux sonores en symboles numériques, seront associés à des logiciels de compréhension du langage, afin que l'ordinateur saisisse la signification des mots prononcés.



L'ordinateur sera capable de parler, c'est-à-dire de trouver des documents sur le réseau Internet, d'en extraire l'information appropriée et de la présenter sous la forme de phrases correctement construites. À chaque étape du processus, la machine dialoguera avec l'utilisateur pour clarifier certains points et rectifier d'éventuelles erreurs. Il posera des questions : «Vous avez dit Rodez, dans l'Aveyron, ou Rodes, dans les Pyrénées-Orientales?»

Galaxy vous parle

Nous étudions de tels systèmes depuis dix ans au Laboratoire d'informatique de l'Institut de technologie du Massachusetts. Jusqu'à présent, les machines qui ont vu le jour ont des domaines de compétence très limités, mais elles sont utiles parce qu'elles donnent une information à jour, par téléphone. Par exemple, le système Jupiter fournit des prévisions météorologiques de 500 villes à travers le monde (ce service gratuit est disponible au 00 1 881 573 8255) ;

le système Pegasus donne les horaires de 4 000 vols aux États-Unis ; le système Voyager informe sur le trafic routier. Aujourd'hui, ces systèmes parlent l'anglais, l'espagnol et le chinois. Si l'on ne tient pas compte du temps nécessaire à la récupération des informations sur le réseau Internet, nos systèmes répondent à la vitesse d'une conversation normale. En France, la SNCF met au point le système ARISE, qui permettra de réserver des billets de train

par téléphone, et le système MASK, qui sera intégré dans un kiosque de vente de billets (voir l'encadré).

Nos applications de reconnaissance et de synthèse vocale sont fondées sur une architecture, nommée Galaxy, que nous avons construite il y a cinq ans. C'est une architecture répartie, c'est-à-dire que différents calculs sont exécutés par plusieurs ordinateurs. Galaxy extrait de l'information de différents domaines

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.

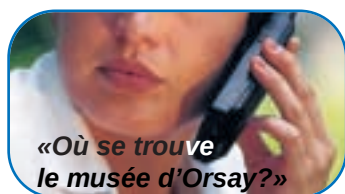


Nous enregistrons les appels, puis nous les analysons, afin d'améliorer les performances du système.

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...

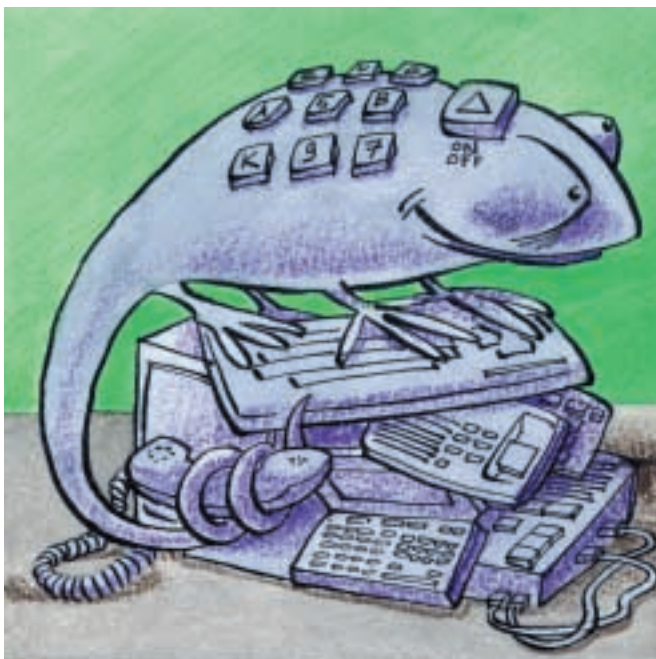
Les programmes, qui remplaceront le matériel électronique trop encombrant et moins flexible, sont les éléments clés de ce projet. Ils feront d'un caméléon de la communication un matériel généraliste, relié à plusieurs applications. Par exemple, dans notre laboratoire, Vanu Bose et Matt Welborn ont mis 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.
