

NICHOLAS NEGROPONTE

L'HOMME NUMÉRIQUE

Traduit de l'américain par Michèle Garène



ROBERT LAFFONT

7.

Des bits et des humains

Réaction fatale

Voilà des années que je passe au minimum trois heures par jour devant un ordinateur, et il m'arrive encore de trouver cela assez frustrant. Comprendre les ordinateurs est à peu près aussi simple que de déchiffrer un relevé de banque. Est-il bien utile que les ordinateurs (et les relevés de banque) soient aussi compliqués ? Pourquoi le numérique est-il aussi ardu ?

En fait, ils ne le sont pas et ils n'ont pas besoin de l'être. L'évolution de l'informatique a été si rapide que cela fait très peu de temps que nous disposons de suffisamment de puissance de traitement bon marché pour la consacrer sans compter à l'amélioration de l'interaction entre votre ordinateur et vous. Avant, dépenser du temps et de l'argent pour améliorer l'interface avec l'utilisateur faisait figure de gaspillage inutile, parce que les cycles informatiques étaient trop précieux pour servir à autre chose qu'à résoudre un problème.

Les scientifiques avaient toutes sortes de bonnes raisons pour justifier l'interface muette. Au début des

années 70, par exemple, on a vu paraître quelques articles « savants » sur les raisons de la « supériorité » des écrans noir et blanc sur les écrans couleur. La couleur n'a rien de condamnable. La communauté de la recherche voulait seulement justifier son incapacité à mettre au point une bonne interface à un coût raisonnable, ou, pour être un peu plus cynique, de faire fonctionner ses méninges.

Ceux d'entre nous qui travaillaient sur l'interface homme-ordinateur à la fin des années 60 et dans les années 70 étaient considérés comme les mauviettes de l'informatique, dignes du plus parfait mépris. Nous n'avions pas choisi la bonne orientation, même si notre domaine de recherche commençait à être accepté. Pour mesurer l'importance de la détection, de l'effet et du feed-back, rappelez-vous la dernière fois que le bouton de l'ascenseur ne s'est pas allumé quand vous avez appuyé dessus (vraisemblablement parce que l'ampoule était morte). Votre frustration devait être énorme : mon ordre a-t-il été compris ? La conception et la fonction de l'interface sont essentielles.

En 1972, on dénombrait 150 000 ordinateurs dans le monde, alors que dans cinq ans le fabricant de circuits intégrés Intel pense en vendre à lui seul 100 millions chaque année (et selon moi, ils sont très loin du compte). Il y a trente ans, se servir d'un ordinateur, comme piloter un véhicule lunaire, était le privilège de quelques initiés au jargon nécessaire pour conduire ces machines dotées dans le meilleur des cas de langages primitifs (des interrupteurs à bascule et

des lumières clignotantes). Selon moi, on faisait inconsciemment tout pour préserver le mystère.

Nous payons encore la note aujourd'hui.

Quand les gens parlent du *look and feel* des ordinateurs, ils font allusion à l'interface graphique (GUI pour les professionnels). Celle-ci a commencé à faire d'énormes progrès vers 1971 grâce aux travaux de Xerox et, peu après, à ceux du MIT et de quelques autres laboratoires, et elle est devenue un véritable produit dix ans plus tard quand Steve Jobs a eu la bonne idée de lancer le Macintosh. Le Mac a été un pas de géant, et, en comparaison, il ne s'est presque rien produit depuis sur le marché. Il a fallu encore cinq années aux autres fabricants pour copier Apple et, dans certains cas, avec des résultats inférieurs, même aujourd'hui.

Chaque fois que l'homme a cherché à faciliter l'usage de la machine par l'homme, il s'est presque exclusivement efforcé d'améliorer les points de contact sensoriels et d'élaborer une meilleure conception physique. L'interface a le plus souvent été traitée comme un problème traditionnel de conception industrielle. Quand on dessine la poignée d'une théière ou le manche d'un râteau, on pense forme, transfert de chaleur et moyen de prévenir la formation d'ampoules.

La conception d'un cockpit est un véritable défi, non seulement à cause du nombre de commandes, de boutons, de cadrans et de jauges, mais également parce que deux ou trois données sensorielles proches peuvent interférer les unes avec les autres. En 1972, un L-1011 d'Eastern Airlines s'est écrasé au sol parce

que son train d'atterrissage n'était pas sorti. Entre la voix du contrôleur du ciel et le bip de l'ordinateur de bord, l'équipage n'avait pas entendu le message d'avertissement. Conception d'interface fatale.

J'ai eu un magnétoscope très intelligent doté d'une reconnaissance vocale presque parfaite qui me connaissait très bien. Je pouvais lui demander d'enregistrer tel et tel programme en le nommant et, dans certains cas, même supposer qu'il le ferait automatiquement, sans que j'aie rien à dire. Puis, tout à coup, mon fils est entré au lycée.

Voilà plus de six ans que je n'ai pas enregistré une émission de télévision. Non que je ne puisse pas le faire, mais l'intérêt est trop faible pour l'effort demandé. C'est inutilement difficile. On a réduit l'emploi du magnétoscope et des télécommandes en général à un problème de boutons à presser. De la même manière, on a traité l'interface avec les PC comme un problème de conception physique. Mais l'interface ne se résume pas à l'aspect et au toucher d'un ordinateur. C'est un problème de création de personnalité, de conception de l'intelligence, et de construction de machines capables de reconnaître l'expression humaine.

Un chien est capable de vous reconnaître à votre démarche à plus de cent mètres de distance, alors qu'un ordinateur ne sait même pas que vous êtes là. N'importe quel animal familier ou presque peut sentir que vous êtes en colère, mais un ordinateur n'en a pas la moindre idée. Même un chiot comprend qu'il a fait une bêtise ; l'ordinateur, non.

Le défi de la prochaine décennie n'est pas de met-

tre au point de plus grands écrans ni de proposer une meilleure qualité sonore et des dispositifs graphiques plus simples à utiliser. C'est de fabriquer des ordinateurs qui vous connaissent, n'ignorent rien de vos besoins, et comprennent les langages verbal et gestuel. Un ordinateur devrait être capable de faire la différence entre « Veil » et « veille », non parce qu'il peut déceler la petite différence acoustique, mais parce qu'il est capable de comprendre le sens. Voilà une conception d'interface digne de ce nom.

Aujourd'hui, le poids de l'interface repose totalement sur les épaules de l'homme. Une chose aussi banale que l'impression d'un document peut être un exercice débilitant plus proche du vaudou que d'un comportement humain respectable. De ce fait, de nombreux adultes baissent les bras, dégoûtés, et prétendent qu'ils sont nuls devant un ordinateur.

Cela va changer.

Odyssées

En 1968, Arthur Clarke a été nommé aux Oscars avec Stanley Kubrick pour le film *2001 : l'Odyssée de l'espace*. Bizarrement, le film est sorti avant le livre. Cela a permis à Clarke de revoir son manuscrit après avoir visionné les rushes (fondés sur une version antérieure de l'histoire). Au sens le plus strict, Clarke a été en mesure de simuler son histoire et de raffiner ses concepts. Il a pu voir et entendre ses idées avant de les confier à l'impression.

Cela explique peut-être pourquoi HAL, l'ordinateur vedette du film, était une vision aussi géniale (bien que mortelle) de l'interface homme/ordinateur de l'avenir. HAL (dont le nom ne vient pas des lettres précédant celles du signe IBM) avait une parfaite maîtrise du langage (compréhension et énonciation), une vision excellente et de l'humour, ce qui est le test suprême de l'intelligence.

Il a fallu près d'un quart de siècle pour qu'apparaisse un autre exemple d'excellente interface : *The Knowledge Navigator* (le navigateur du savoir). Cette vidéo, également une production théâtrale, ce que l'on appelle un prototype vidéo, était une commande de John Sculley, alors PDG d'Apple, dont le propre livre avait aussi pour titre *Odyssée*. Le livre de Sculley se terminait sur l'idée d'un navigateur du savoir, qui allait devenir une vidéo. Il voulait donner un exemple d'une interface de l'avenir, au-delà des souris et des menus. Il a fait de l'excellent travail.

The Knowledge Navigator décrit un appareil plat ressemblant à un livre ouvert sur le bureau d'un professeur BCBG. Dans un coin de l'écran, on voit un homme à nœud papillon qui n'est autre que la personnalité de la machine. Le professeur demande à cet agent de l'aider à préparer une conférence, lui délègue plusieurs tâches et, de temps à autre, se voit rafraîchir la mémoire sur d'autres sujets. L'agent peut voir, entendre et réagir intelligemment, comme un assistant humain.

HAL et le *Knowledge Navigator* ont cela de commun qu'ils font preuve d'un tel degré d'intelligence que l'interface physique disparaît presque. Voilà le secret

de la conception d'une interface : la faire disparaître. Quand vous rencontrez quelqu'un pour la première fois, vous serez d'abord très conscient de son apparence physique, de sa voix et de ses gestes. Mais rapidement le contenu de votre dialogue va dominer, même si les intonations et les expressions faciales jouent un grand rôle. Une bonne interface devrait se comporter ainsi. Le problème est moins de concevoir un tableau de bord qu'un être humain.

Pourtant, la plupart des concepteurs d'interface s'obstinent à vouloir à rendre facile l'usage de machines stupides pour des êtres intelligents. Ils s'inspirent d'un domaine qu'on appelle « facteurs humains » aux États-Unis et « ergonomie » en Europe, qui s'intéresse à la manière dont l'homme se sert de ses sens et de son corps dans ses rapports avec les outils de son environnement immédiat.

Le combiné téléphonique est probablement l'application qui a été la plus redessinée, revue et corrigée sans que cela donne de résultats probants. Avec leur interface inutilisable, les téléphones cellulaires sont encore pires dans le genre que les magnétoscopes. Un téléphone Bang & Olufsen relève de la sculpture, pas de la téléphonie, et il est encore plus difficile à utiliser que le bon vieil appareil à cadran en bakélite.

En outre, on n'arrête pas d'ajouter des caractéristiques aux téléphones. Mémoire d'appel, rappel du dernier numéro composé, gestion de carte de crédit, mise en attente d'appels, renvoi d'appels et filtrage d'appels : on multiplie les fonctions d'un appareil que sa taille minuscule rend pratiquement inutilisable.

Mais peu m'importent toutes ces caractéristiques !

Ce que je veux, c'est de ne pas avoir à toucher mon appareil. Pourquoi les concepteurs d'appareils téléphoniques ne comprennent-ils pas que nous ne voulons pas avoir à composer un numéro ? Nous voulons joindre des gens par téléphone !

Si c'était possible, on déléguerait cette tâche, ce qui me fait dire qu'au lieu de réfléchir au design des combinés téléphoniques on ferait mieux de songer à fabriquer un secrétaire robot qui tiendrait dans notre poche.

Au-delà du simple trait

On a commencé à s'intéresser à l'interface avec l'ordinateur en mars 1960 après la publication de l'article de J.C.R. Licklider sur « la symbiose homme-ordinateur ». Psychologue expérimental et acousticien de formation, converti et messie de l'informatique, « Lick » a dirigé les premières recherches de l'ARPA (*Advanced Research Projects Agency*) dans ce domaine. Au milieu des années 60, on lui a demandé d'écrire une annexe à un rapport de la Commission Carnegie sur l'avenir de la télévision. C'est là que Lick a lancé le terme de « diffusion ciblée » (*narrow casting*). Il était loin de se douter que ses deux contributions, la symbiose homme-ordinateur et la diffusion ciblée, seraient appelées à converger dans les années 90.

Au début des années 60, la recherche balbutiante en matière d'interface homme-ordinateur allait prendre deux orientations qui mettraient vingt ans à se

rejoindre. L'une s'intéressait à l'interactivité, l'autre se concentrait sur la richesse sensorielle.

On s'est attaqué à l'interactivité en résolvant le problème du partage d'un ordinateur, une ressource alors onéreuse et monolithique. Dans les années 50 et au début des années 60, un ordinateur était tellement précieux qu'on s'arrangeait pour qu'il fonctionne sans arrêt. Il était impensable de se connecter, de faire poser une question par l'ordinateur et de le laisser sans rien faire pendant qu'un homme lisait la question, réfléchissait et y répondait. On inventa alors une méthode appelée « le temps partagé », qui permettait à de multiples utilisateurs installés de partager une seule machine à partir de lieux éloignés. En divisant la ressource entre dix personnes, par exemple, non seulement chacune pouvait disposer d'un dixième de la machine, mais encore son temps de réflexion pouvait être utilisé par une autre pour se servir de l'ordinateur.

Ce partage du gâteau numérique ne posait pas de problème tant qu'aucun utilisateur ne monopolisait la machine, en ayant besoin de grandes quantités de calcul ou de largeur de bande. Les premiers terminaux fonctionnaient à 110 bauds. Je me souviens très bien du jour où ils sont passés à 300 bauds, cela semblait tenir du miracle.

En revanche, on s'est intéressé à la richesse sensorielle dans le domaine de l'interaction graphique demandant une très grande largeur de bande. Les premiers graphiques informatiques nécessitaient une machine exclusivement vouée à la production d'images. Sur le principe, l'ordinateur n'était pas différent

du micro actuel, mais il remplissait une grande pièce et coûtait des millions de dollars. Au début, l'infographie permettait de tracer des lignes, ce qui nécessitait une énorme puissance de calcul pour contrôler directement le faisceau du tube cathodique.

Ce n'est que dix ans plus tard que l'infographie a commencé à tracer non plus seulement des lignes, mais des formes et des images. Ces nouveaux écrans, à balayage de type télévision, nécessitaient beaucoup de mémoire pour stocker l'image point par point. Ils sont tellement répandus aujourd'hui que la plupart des gens ignorent qu'au départ ils faisaient figure d'hérésie. (Personne ou presque ne croyait en 1970 que la mémoire informatique serait un jour suffisamment abordable pour que l'on puisse en consacrer autant au graphisme.)

Le temps partagé et l'infographie n'ont pas fait très bon ménage au cours des deux décennies suivantes. Des systèmes de temps partagé peu évolués sur le plan de l'interface sont apparus comme un outil acceptable pour l'informatique des affaires et de l'université, donnant naissance aux systèmes bancaires informatisés et aux systèmes de réservations des compagnies aériennes, que l'on trouve parfaitement normal aujourd'hui. Le temps partagé commercial a marché main dans la main avec une conception d'interface très parcimonieuse, généralement une sortie machine à écrire, cherchant presque volontairement à rendre le système suffisamment lent pour tout utilisateur, afin que les autres obtiennent leur juste part.

En revanche, l'infographie s'est développée en grande partie avec l'informatique individuelle. En

1968, les mini-ordinateurs dans les 20 000 dollars ont commencé à apparaître, notamment parce que l'automatisation des machines et des usines nécessitait des contrôles très précis en temps réel. Ce qui était aussi le cas de l'infographie. Associés à des écrans de visualisation, ces systèmes infographiques ont été les précurseurs de ce que nous connaissons aujourd'hui sous le nom de « stations de travail », qui ne sont rien d'autre que des micro-ordinateurs en pantalon long.

L'interface multimodale

On considère généralement la redondance comme un mauvais symptôme, l'indice d'une verbosité inutile ou d'un manque d'attention. Au début des recherches sur l'interface, on a étudié des techniques d'interaction en tentant de sélectionner judicieusement un moyen ou un autre, selon les circonstances. Un crayon optique était-il mieux qu'une tablette graphique ? Cette mentalité de l'alternative était mue par la fausse croyance qu'il existait une solution universelle idéale pour toute situation donnée ; c'est faux parce que les gens sont différents, que les situations changent, et que le contexte d'une interaction donnée dépend souvent du canal disponible. Il n'y a pas d'idéal en matière de conception d'interface.

Au milieu des années '70, j'ai rencontré un amiral qui disposait d'un des systèmes de commande et de contrôle les plus évolués. Il aboyait des ordres à un deuxième classe qui entrait consciencieusement les bonnes instructions dans l'ordinateur. En ce sens, le

système avait une interface géniale : non seulement il était capable de reconnaître la voix, mais il avait de la patience à revendre. L'amiral pouvait parler et gesticuler en arpentant la pièce. Il pouvait être lui-même.

Néanmoins, il lui aurait été impossible de mettre sur pied un plan d'attaque par le biais d'une interface aussi indirecte. Il savait que le marin regardait la situation par le petit trou de la serrure de l'écran de l'ordinateur. L'amiral préférait avoir un contact direct avec une grande carte du « théâtre des opérations » sur laquelle il pouvait punaiser des petits bateaux bleus et rouges de la forme voulue. (À l'époque, le fait que les Russes utilisent les mêmes couleurs nous faisait beaucoup rire.)

L'amiral était à l'aise avec la carte, non pas parce que c'était un objet traditionnel doté d'une très haute résolution, mais parce que son corps tout entier participait. Le seul fait de pouvoir déplacer des bateaux l'aidait à mémoriser la situation. Il était profondément impliqué dans son rapport avec la carte, jusqu'aux muscles de son cou. En l'occurrence, il n'était pas obligé de choisir entre un type d'interface ou un autre.

Cette façon de voir a créé une révolution, parce qu'on a commencé à se dire que la redondance avait du bon. On a compris que la meilleure interface devait faire appel à plusieurs canaux de communication différents et concomitants, par le biais desquels un utilisateur pourrait exprimer quelque chose en se servant de différents sens ou appareils sensoriels (ceux de l'utilisateur et ceux de la machine). En outre, et c'est tout aussi important, on a compris qu'un canal de

communication pourrait fournir l'information manquant dans l'autre.

Mettons que nous soyons plusieurs dans une pièce et que je demande à quelqu'un comment il s'appelle, la question n'a aucun sens pour mes interlocuteurs à moins qu'ils ne puissent voir à qui je m'adresse.

Cela a été superbement illustré par un programme appelé Put-That-There (« Posez-ça là ») mis au point au MIT par Dick Bolt et Chris Schmandt. Le premier prototype du programme, en 1980, vous permettait de parler et de gesticuler devant un écran de la taille d'un mur et de déplacer des objets simples (qui sont devenus des bateaux) sur un écran blanc (qui est devenu la mer des Caraïbes). Dans une démonstration filmée de Put-That-There, le programme a mal interprété une commande. Le « Oh ! merde » spontané de Schmandt n'a pas été coupé au montage, pour bien rappeler aux nombreux futurs spectateurs l'ampleur du travail qui restait à faire.

L'idée est simple : il devrait être possible de parler, de montrer du doigt et de se servir de ses yeux dans un rapport avec l'ordinateur : voilà l'interface multimodale ; il s'agit moins d'échanger des messages (la base du temps partagé) que d'instaurer un dialogue, comme entre deux humains. À l'époque, toutes ces tentatives de mettre au point une conception de l'interface « et/ou » avaient des allures de science bâclée. J'ai peu de respect pour les tests et l'évaluation dans la recherche en interface. Mon argument, peut-être arrogant, est que, s'il est nécessaire de multiplier les tests pour mesurer le progrès apporté par une découverte, pour trouver ce qui fait la petite diffé-

rence, cela veut dire qu'elle ne représente pas un progrès suffisant.

La différence perceptible

Quand j'étais petit, ma mère avait un placard à linge au fond duquel il y avait un « mur secret ». Ce n'était pas un grand secret : une série de traits de crayon que nous faisions périodiquement pour mesurer ma croissance. Tous les traits étaient consciencieusement datés, et ils étaient plus ou moins distants les uns des autres, parce que, par exemple, nous avions passé l'été ailleurs. Utiliser deux placards ne paraissait pas logique.

Ce mur au fond du placard était une affaire privée qui mesurait, d'une certaine manière, ma consommation de lait, d'épinards, entre autres aliments sains.

En revanche, ma croissance avait un aspect beaucoup plus spectaculaire. Un oncle pouvait dire : « Combien de centimètres as-tu pris, Nicky ? » (parce qu'il ne m'avait pas vu pendant deux ans). Mais j'étais incapable de comprendre le changement. Je ne voyais que les traits de crayon au fond du placard.

La différence juste perceptible est une unité de mesure en psychophysique. Son nom seul a influencé la conception de l'interface humaine. S'il s'agit d'une différence juste perceptible, pourquoi s'embêter avec ? S'il faut mesurer soigneusement pour déceler une différence, peut-être ne travaillons-nous pas aux choses qui ont vraiment de l'importance.

Par exemple, des études savantes ont suggéré que la parole et le langage naturel ne sont pas des canaux de communication appropriés entre l'homme et l'ordinateur dans la plupart des applications. Ces rapports techniques regorgent de tableaux, de groupes témoins, etc., prouvant que le langage naturel est déroutant pour la communication homme-ordinateur.

Je n'attends certainement pas du pilote d'un 747 qu'il décolle en chantant à tue-tête : « Et on est partis ! », mais je ne vois pas pourquoi on n'exploiterait pas la richesse du discours et des gestes, même dans un cockpit. Où que se trouve l'ordinateur, si l'on veut que l'interface soit efficace, il faut associer la richesse sensorielle et l'intelligence de la machine.

Quand cela se produira, la différence sera perceptible. Nous verrons ce que voyait mon oncle, au lieu des traits de crayon au fond du placard.

Des interfaces intelligentes

En matière d'interface, mon rêve est de voir des ordinateurs ressembler davantage à des hommes. On reprochera à cette idée d'être trop romantique, vague ou irréalisable. Je lui reprocherais de manquer d'ambition. Peut-être existe-t-il des canaux de communication complètement exotiques auxquels nous n'avons pas encore songé. (Marié à une vraie jumelle et frère de vrais jumeaux, je suis tout prêt à croire par expérience que la communication extrasensorielle n'est pas impossible.)

Au milieu des années 60, je me suis fixé comme objectif d'essayer d'imiter la communication en face à face, avec toute la richesse de la gestuelle, des expressions du visage, et des mouvements du corps. J'ai pris l'amiral comme modèle.

Le système de gestion des données spatiales (vers 1976) avait pour objectif de fournir une interface humaine qui « amènerait les ordinateurs directement aux généraux, aux PDG d'entreprises et aux enfants de six ans. » Le système était conçu pour que l'on puisse apprendre à l'utiliser en trente secondes. Il permettait d'explorer et de manipuler du son, de l'image et des données complexes avec autant de facilité que l'on se sert d'un ordinateur de bureau ou d'une bibliothèque.

C'était révolutionnaire pour l'époque — la fin des années 70 —, mais c'était encore très loin de la simplicité de la communication entre l'amiral et son marin. L'interface homme-ordinateur future s'ancrera dans la délégation, pas dans le langage de la manipulation directe — baisser une manette, presser un bouton, cliquer — ni dans les interfaces avec les souris. La « facilité d'utilisation » a été un objectif tellement irrésistible que nous en oublions parfois que beaucoup de gens se refusent à se servir de la machine. Ils veulent obtenir un résultat, point.

Ce que nous appelons aujourd'hui les « interfaces orientées agents » va devenir le moyen dominant par lequel les ordinateurs et les gens dialogueront. Il y aura des points précis dans l'espace et dans le temps où des bits seront convertis en atomes, et vice versa. Que cela se fasse par le biais de la transmission d'un

cristal liquide ou par la réverbération d'un synthétiseur vocal, l'interface devra avoir une taille, une forme, une couleur, un ton de voix, et tout l'attirail sensoriel.