

Université Paris I Panthéon-Sorbonne
UFR d'Arts Plastiques et Sciences de l'Art

Mémoire de recherche préparé sous la direction de Jacinto Lageira
(et sous la co-direction de Marc Jimenez)

Master 2 Esthétique
Adeline Gasnier
2011-2012
adeliga@hotmail.com

ARCHITECTURE DE L'ÉMOTION : LES AGENTS CONVERSATIONNELS ET LA FABRIQUE DU SUJET AFFECTIF

Introduction	p. 3
I. Les agents conversationnels	p. 5
Typologie réduite	p. 5
Environnements propitiatoires, culturels et artistiques	p. 10
Statuts de l'expérience avec des agents conversationnels	p. 22
Les agents conversationnels et l'art : prémices	p. 31
Quels esthètes sont ces agents?	p. 36
Les agents conversationnels sont-ils des œuvres d'art?	p. 43
II. Architecture de l'émotion	p. 57
<i>Diurnisme</i> (Philippe Rahm, Centre Pompidou, 2007)	p. 58
Les formes de l'émotion	p. 64
Rationalisation et «émotionalisation»	p. 73
Neurobiologie émotionnelle	p. 80
III. La fabrique du sujet	p. 86
Vers une culture de l'affectif?	p. 89
<i>Affective Computing</i> (informatique affective)	p. 99
Le «test de Picard»	p. 107
Pathologie de l'interprète	p. 111
Fluidifier les opinions individuelles	p. 121
Conclusion	p. 130
Ressources mentionnées & consultées	p. 133
Annexe (DVD joint)	

INTRODUCTION

« *O veram mortuam vitam sine philosophia!*¹ » [Kepler, 2010, p. 94]. Cette phrase extraite de *L'étréenne ou la neige sexangulaire* de Johannes Kepler, rédigé à la faveur d'un hiver pragois neigeux, entre 1610 et 1611 à la cour de l'Empereur Rudolf II, pourrait peut-être aujourd'hui être retournée : « Ô vraiment que la vie est morte *avec* la philosophie ». En effet des ouvrages philosophiques tels que *La conscience expliquée* (1991) de Daniel C. Dennett, par exemple, nous permettent de mieux comprendre les mécanismes de la vie, de ces phénomènes vitaux et humains réputés inexplicables (à l'instar de la conscience ou des émotions). Nous avons la chance de vivre à cette époque où la science-fiction n'est plus tout à fait fiction au sens courant du terme (en tant qu'imaginaire éloigné et virtuel) dans la mesure où nous pouvons déjà discuter avec des entités « mécaniques » plutôt étonnantes, des agents conversationnels, mais sans pour autant savoir encore dans quel sens leur généralisation pourra aller, ni jusqu'où l'augmentation de l'être humain nous emmènera. Et nous nous situons au seuil d'une époque qui va peut-être devoir revisiter ses deux concepts discriminants les plus basiques et solides : la vie, et la mort. Il ne s'agit pas de déplorer une quelconque perte de valeurs, mais plutôt de noter ce trouble, ce frémissement dans les certitudes.

Voici une façon un peu grandiloquente d'introduire les choses. Il est tout autant possible d'opter pour une version plus légère, plus prompte à faire directement connaissance avec quelques-uns des modèles informatiques tirant partie des recherches (actuelles et passées) en termes de mécanismes humains. Plutôt que de parier sur la fracture, la présente recherche propose de développer quelques aspects soulevés par ces rencontres, en les contextualisant et en ciblant un trait particulier : l'architecture de l'émotion. Les raisons de ces choix feront aussi l'objet des développements à venir. Le travail va se décomposer comme suit : dans la première partie il sera question des agents conversationnels, dans la deuxième nous nous intéresserons à l'architecture des émotions (architecture au sens artistique, mais aussi au sens plus large d'organisme, d'organisation), et enfin la troisième fera le lien entre les deux précédentes en se

¹ « Ô vraiment que la vie est morte sans la philosophie! » [Kepler, 2010, p. 94]. (Ma traduction, d'après la traduction anglaise proposée dans l'ouvrage cité.)

penchant sur la rencontre entre informatique et émotions.

Remarque – Les crochets indiquent que les références complètes de l’ouvrage sont situées dans la liste des ressources mentionnées et consultées, en fin de mémoire (la date est celle de la publication citée, et non celle de la primo-parution en langue originale). Ce système a été adopté afin de réduire les appels de note par ailleurs assez fréquents. Les liens internet se réfèrent tous à des ressources consultées et accessibles en ligne en date du 8 mai 2012. Une courte annexe contenant des images supplémentaires est fournie sur support digital, en plus du présent document imprimé. Les ressources qu’elle contient ne seront pas appelées explicitement dans le corps du texte dans la mesure où ce qui devait être accessible en cours de lecture y a été reproduit directement.

I. LES AGENTS CONVERSATIONNELS

Cette première partie, directement consacrée aux agents conversationnels, va se déployer de la façon suivante : il sera d'abord proposé une sorte de typologie afin de permettre aux personnes les moins familières avec ces entités de se faire une idée de ce dont il est question ; ensuite nous élargirons le champ afin de les situer dans un contexte lui-même spécifique. Ce sera l'occasion de se demander si le contexte suffit à aiguiller vers une compréhension intuitivement correcte de ces agents. Puis, nous recentrerons sur l'échange en tant que tel pour tenter de cerner s'il y a là quelque chose de singulier. La notion d'expérience servira de point de repère. Suite à cela il sera proposé de se tourner du côté de l'expérience de réception artistique, afin d'y glaner quelque appui pour s'orienter dans ce dédale expérimental. Ce rapprochement avec l'art et l'esthétique permettra de se demander si ces agents pourraient avoir quelques prédispositions en termes de sensibilité et d'esthétique. Enfin, la question sera posée de savoir si les agents conversationnels sont des œuvres d'art.

TYPLOGIE RÉDUITE

Qu'est-ce qu'un agent conversationnel? Prenons une porte d'entrée factuelle et commençons par les contextes et les fonctions. Si l'on n'a pas accès aux laboratoires de recherche et développement en informatique (tels le MIT Media Lab, pour ne citer que le plus connu), on rencontre alors ces agents sur Internet. Soit dans ce qui est couramment nommé espace virtuel et qui en soi n'est déjà pas un espace-temps uniforme. Il y a ainsi différents contextes dans lesquels il est aujourd'hui possible en tant « qu'utilisateurs » de rencontrer ces entités. Le paragraphe suivant va tenter de proposer une brève typologie « récapitulative », en vue de délimiter un tant soit peu ce vaste sujet.

Ces agents peuvent être dédiés à la relation client. En ce cas leur tâche consiste notamment à répondre en langage naturel aux questions des utilisateurs. En tant que programmes informatiques ils ont la charge d'opérer la médiation entre les bases de données des serveurs de l'entreprise qu'ils représentent, et les utilisateurs souhaitant obtenir de l'aide ou

des informations à propos des offres et produits de l'entreprise. De telles entités s'appellent par exemple Anna (IKEA) ; Clara (FNAC) ; Léa (EDF) ; Louise (eBay) ; Sarah (PayPal) ; Eva (Scandinavian Airlines), etc., et le secteur d'activité de ces descendantes d'Olympia [cf. Freud, 2001] est celui des services et de l'aide individualisés à la personne. Dans ce contexte, les agents opèrent en tant que serviteurs, aussi bien de l'entreprise qu'ils représentent que des utilisateurs avec lesquels ils interagissent.

On peut aussi les rencontrer sur des canaux IRC (*Internet Relay Chat*) ou des réseaux sociaux. SmarterChild, par exemple, a été présent sur des dispositifs de messagerie instantanée tels qu'AOL Instant Messenger et MSN de 2000 à 2009. Cet agent y a eu jusqu'à dix millions d'amis-utilisateurs avec lesquels il aura pu converser — leur donner des conseils, leur raconter des blagues, ou simplement enregistrer leurs confidences, etc. — allant ainsi jusqu'à recevoir un milliard de messages par jour, c'est-à-dire allant jusqu'à proposer un milliard de réponses par jour. Pour la petite histoire, en 2006 SmarterChild a été racheté par Microsoft avant de disparaître des réseaux en 2009 — un certain flou subsiste quant à l'utilisation pouvant ou ayant pu être faite de « sa mémoire ». Bien sûr la blague usuelle par rapport à ces agents officiant sur des réseaux de *chat*, de rencontre ou de séduction en ligne est celle du programme flouant les individus les plus entreprenants (ou les plus rêveurs) en se faisant passer pour une femme disponible et intéressée par une relation [cf. Epstein, 2007]. Mais de façon bien plus courante ces agents conversationnels (*chatbots*) opèrent sur IRC en tant que relais et modérateurs du canal via la reconnaissance textuelle — ce qui a d'ailleurs stimulé la créativité langagière afin de contourner les interdits sémantiques : le terme « porn » de pornographie étant ainsi devenu « pr0n » (avec un zéro à la place du 'o'), pour ne citer qu'une des occurrences les plus emblématiques du *leetspeak* détourné. Dans ce rôle de modération on retrouve des fonctions de surveillance, de détection, de filtrage, et par là même de critique. L'historien d'art Otto K. Werckmeister, dans son livre *Der Medusa-Effekt* (« L'effet Méduse », 2005, non traduit), évoque à propos de la vidéosurveillance l'idée d'une « vue sans sujet » (*subjektloses Sehen*) [Werckmeister, 2005, p.7] : on peut aussi entendre cette expression dans le sens d'un appareillage capable d'une vue, d'un voir objectif, sans pour autant être doté d'une subjectivité ou d'un Soi. Devrait-on ici parler à propos des agents conversationnels d'un « pouvoir

décisionnaire sans sujet» comme on pourrait le dire du hasard? Cette vaste question fait partie de ce qui motive la présente recherche.

Les agents conversationnels gravitent également sur certaines plateformes de mondes persistants. On peut ainsi faire la connaissance de Suzette sur Blue Mars (plateforme ou unité de simulation s'apparentant à Second Life et développée par la société Avatar Reality) où, à l'instar de n'importe quels autres participants, elle y a au moins un avatar. Le développement de Suzette va plutôt dans le sens d'une autonomisation au sein de ce monde et non dans le sens d'une fusion avec l'environnement. En d'autres termes, elle n'est pas le moteur artificiellement intelligent de cet environnement informatique (*AI Game Engine*) : elle n'est pas «l'esprit du jeu», même si de façon métaphorique elle en est aussi la justification. Son rôle est plutôt présenté comme étant celui d'un apprenti : elle est là pour apprendre de ses conversations avec les autres participants, et ce notamment en tant que candidate au Loebner Prize, une compétition en forme de vitrine promotionnelle pour les agents conversationnels et le test de Turing (même si la «version en ligne» de Suzette n'est pas tout à fait la même que sa «version compétitive»).

À ce stade du tour d'horizon nous pouvons noter que les agents conversationnels se rencontrent dans des contextes divers et qu'ils disposent de degrés de liberté variables. Il y a oscillation entre une utilisation marketing d'un côté et l'expérimentation R&D (Recherche et Développement) de l'autre. La plupart des agents virtuels actuellement utilisés par des sociétés de vente sont «rudimentaires» (ou plutôt monomaniaques) et ne sortent pas de leur mission marketing, tandis que les agents conversationnels principalement orientés par la recherche sont en ligne pour apprendre des utilisateurs. C'est-à-dire pour essayer d'apprendre tout ce qu'il est possible d'apprendre à l'aide du (et sur le) langage naturel. À propos de cette «R&D en ligne», nous pouvons remarquer qu'elle n'est pas tant officielle et subventionnée en conséquence que basée sur des initiatives personnelles (programmeurs et petites entreprises de divertissement, tel Existor Ltd). Cela dit, il serait peut-être hypocrite d'exclure du domaine de la recherche ces initiatives apparemment motivées par l'appât du gain (récompense du Loebner Prize et futures commercialisations, entre autres) au motif qu'aucune institution ne semble actuellement développer de ces entités *online*, ni même s'y intéresser : les choses sont plus nuancées et ces

agents participent de la construction / compréhension du « sens commun ».

Notons également que le flou quant à ces agents (leurs rôles, leurs méthodes, leurs visées, etc.) n'est pas tant du côté des entreprises les utilisant pour leur service client que du côté de la recherche en ligne, c'est-à-dire du côté du développement participatif de ces entités apprenantes. Ou pour le dire autrement : les utilisateurs ne participent pas ici au développement du code et des algorithmes, mais participent au développement de la base de données de ces programmes — manière abrupte de qualifier le phénomène d'élaboration mémorielle, d'éducation réciproque et d'apprentissage mutuel à l'œuvre. En cela il y a bien formation et information, mais sans pour autant qu'il soit simple de définir dans quels sens et à quels niveaux cela s'effectue, sans qu'il soit simple de définir qui forme qui. Et quel que soit le sens de cet échange, il n'est pas sans impact. Par ailleurs, l'indécision et le trouble participent de l'attrait exercé par ces agents : quand bien même nous serions blasés et pétris de science-fiction, l'hypothèse est qu'il peut se produire là quelque chose qui repondère nos représentations du monde. Peut-être simplement parce que nos accès à Internet et aux agents virtuels ne se font pas depuis des laboratoires high-tech, à l'instar des films de science-fiction hollywoodiens, mais plutôt à la lueur de nos banales errances solitaires sur le Net.

Lorsque l'on s'attarde du côté du versant expérimental, on est alors amené à faire la connaissance d'agents conversationnels exemplaires, tels qu'ELIZA, A.L.I.C.E, Cleverbot, Jabberwacky ou encore Elbot, pour ne citer que les plus emblématiques et fréquentés. Ces agents se présentent dans un dispositif qui n'a rien de nouveau puisque ce n'est jamais que le très célèbre test de Turing, initialement pensé par le mathématicien et ingénieur britannique Alan Turing en 1950. Il avait proposé de remplacer l'idée directrice qu'était alors la question « les machines peuvent-elles penser ? » par un « jeu de l'imitation » (*imitation game*), où le défi à relever devenait celui de rendre indiscernables machines et humains au sein d'un dispositif de dialogue verbal interactif, mais sans interactions visuelles [Turing, 1950]. Si une machine était confondue avec un être humain, elle gagnait la partie.

Ainsi, pour résumer brièvement, les agents conversationnels peuvent en l'état être présentés comme étant des programmes-robots (*bots*), dont les fonctions comprennent entre autres : 1) le service et l'aide individualisés à la personne, 2) la modération, avec la régulation

sémantique des activités d'un réseau social, 3) la tromperie, avec le fait de se faire passer pour, 4) l'apprentissage, avec la constitution d'une culture-mémoire sémantique en langage naturel, 5) l'éducation, avec l'influence que peuvent avoir leurs propos et les modes à travers lesquels ils sont montrés, 6) l'imitation humaine parfaite, avec le test de Turing, 7) la compétition, avec leur statut de candidats perpétuels, 8) la collecte d'informations.

Maintenant, le problème d'une telle typologie n'est pas qu'elle soit réduite mais bien plutôt réductrice : elle enferme ces agents dans les seuls domaines de la logique, des mathématiques, de l'ingénierie, du marketing, voire du management et de la manipulation, tandis que la dimension humaine semble y être écrasée par une forme de rationalisation outrancière, reléguant aux limbes aussi bien les sciences humaines que l'art. Or les questions que nous posent ces agents débordent le cadre technique : qu'est-ce qu'être humain ? Qu'est-ce qu'être vivant (ou mort) ? Qu'est-ce que penser ? Qu'est-ce qu'être conscient ? Qu'est-ce qu'être libre ? Qu'est-ce qu'être un individu ? Qu'est-ce que créer ? Qu'est-ce qu'être sujet ? Qu'est-ce que l'intimité ? Qu'est-ce que l'inconscient ? Qu'est-ce que sentir ? Qu'est-ce que participer à la société (qu'est-ce que la démocratie) ? Et comment, c'est-à-dire sous quelles conditions et à quel prix ? Etc. Évidemment ces questions sont loin d'être nouvelles mais la nouveauté n'est pas tant à chercher du côté des questions que du côté, peut-être, de l'obsolescence de certaines de nos réponses les plus usuelles. Peut-être est-il trompeur de croire que l'aisance avec laquelle nous semblons nous approprier ces techniques par l'usage quotidien ne s'accompagne pas d'une déstabilisation quant à nos rapports au monde et à nous-mêmes. Le fait que ces techniques soient de plus en plus ergonomiques et *user friendly* ne constitue pas une réponse existentielle : si la tâche s'en trouve facilitée et toujours plus plaisante, ce confort n'en devient pas pour autant signifiant, et la question de la tâche, en elle-même, n'en demeure pas moins ouverte, c'est-à-dire potentiellement anxiogène. Essayons alors de mettre en relief quelques-unes de nos intrications avec les technologies.

Avant de poursuivre, résumons le paragraphe précédent : une brève présentation des agents conversationnels y aura été proposée, sous forme de typologie fonctionnelle. Si cette approche est frustrante elle n'en permet pas moins de baliser un sujet qui se révèle à l'usage être

dense, mouvant, redondant, fait d'emprunts et de chevauchements, c'est-à-dire difficile à présenter et à classer. L'effort ci-dessus ne doit pas être compris comme reflet fidèle de la réalité : il ne sert qu'à orienter succinctement les lecteurs n'ayant aucune idée de ce dont il est question, s'il devait s'en trouver. Aussi, rappelons en dernière instance que ces agents sont avant tout conçus pour avoir des personnalités singulières, et qu'à ce titre une typologie ne peut qu'écraser ce qui les distingue. Or ce sont bien ces singularités qui auront motivé la présente recherche et qui feront en grande partie l'objet des questionnements à venir. Maintenant, plutôt que de continuer par une confrontation directe avec ces agents (le choix de cette diversion sera explicité plus loin), le paragraphe suivant va s'intéresser à certains aspects des contextes dans lesquels ils apparaissent.

ENVIRONNEMENTS PROPITIATOIRES, CULTURELS ET ARTISTIQUES

Les agents conversationnels n'existent pas seulement dans un contexte matériel et utilitariste, mais encore dans un contexte socioculturel. S'ils ont un parfum de banalité (l'effet de l'apparition des agents conversationnels sur des sites internet n'est pas assimilable aux effets de *L'arrivée en gare de la Ciotat*), c'est aussi que nous nous déployons dans des environnements accueillants et adaptés à leurs besoins, à leurs logiques (s'ils sont adaptés aux nôtres, la réciproque est valable). Mais quels peuvent être ces aspects favorisants? Le paragraphe qui vient va tenter de tester l'hypothèse selon laquelle nos environnements actuels nous permettent de comprendre, intuitivement ou non, l'informatique et les déplacements d'enjeux qu'elle opère.

La programmation informatique est un phénomène récent si on la compare à l'écriture ou à l'imprimerie. Relativement à cela, l'arrivée de la CAO (Création Assistée par Ordinateur) s'est faite de façon plutôt fluide, au point de rendre rapidement passés de mode les débats à propos des techniques engagées. Ces débats semblent coller à la modernité, à la photographie et au cinéma analogiques, avec l'exacerbation d'un clivage entre un art aux accents romantico-humanistes d'un côté, et de l'autre un technopositivisme créatif. Faut-il en déduire que le fait de questionner l'informatique sous l'angle de l'art revient à jouer une énième fois la querelle entre les Anciens et les Modernes? La réponse est oui, à condition d'oublier que l'informatique

est dorénavant présente dans *tous* les domaines et que c'est à ce titre que l'art (au sens courant de ce qui se produit dans les musées, les centres d'art et les galeries), non plus, n'y échappe pas — ne serait-ce qu'au niveau de la diffusion des « informations pratiques ». Au lieu de s'enliser dans cette caricaturale ligne Maginot, avec les pour d'un côté et les contre de l'autre, attachons-nous plutôt à ce que cela peut impliquer concrètement sur les terrains de la création (sur certains d'entre eux — et laissons ici de côté la question de savoir si l'art peut se passer de re/création).

Si l'on s'intéresse aux outils de production multimédia on remarque déjà des récurrences, notamment Adobe et Autodesk, soit deux sociétés d'édition de logiciels, toutes deux fondées en 1982. Une étude consacrée au classement mondial des éditeurs de logiciels² rangeait en 2010 Adobe à la dixième place et Autodesk à la dix-neuvième (Microsoft occupant la première et Apple la onzième). S'il n'est plus besoin de présenter Adobe (traitement d'images, de sons, montage audiovisuel, postproduction, animation, effets spéciaux, design, mise en page, illustration, typographie, calibrage d'écran, création de page web, d'animation web, gestion du *portable document format* [pdf] — autant d'outils avec lesquels travaillent aussi bien les imprimeurs, que les graphistes, les infographistes, les éditeurs, les webmasters, les vidéastes, les monteurs, etc.), Autodesk pour sa part a peut-être encore besoin d'être présentée : elle édite des logiciels d'architecture, de création de plans et dessins techniques, de 3D, d'animation, de simulation de fluides, de dynamiques, etc., notamment utilisés par les cabinets d'architectes, de design, de conception automobile, les agences de graphisme, l'industrie du jeu vidéo, du cinéma, de la télévision, de la publicité... pour ce qui est de la partie ayant trait aux arts et aux industries culturelles. Pour la petite histoire, l'un de ses logiciels les plus répandus, Maya, a connu une partie de ses débuts dans la mouvance du « Plan Recherche Image », entre 1982 et 1984 à l'INA (Institut National de l'Audiovisuel), avant de passer à TDI (Thomson Digital Image), sous le nom d'Explore³. Les personnes les plus pointues ne manqueront pas d'ajouter

² Cf. Raman Chitkara & Pierre Marty, étude « Global Software Leaders – key players and market trends » (décembre 2010), éd. PricewaterhouseCoopers, Londres, 2011. <http://www.pwc.fr/global-software-leaders-key-players-and-market-trends.html>.

³ La société française TDI ayant été vendue en 1993 à la société américaine Wavefront. Pour une approche plus complète de l'histoire des « nouvelles images » en France, cf. le séminaire « Histoire de l'art de l'image de synthèse française », organisé par l'ENSAD, Paris III, Paris VIII et Paris ACM SIGGRAPH, courant 2011-2012. Les enregistrements des séances sont disponibles, tout ou partie, en ligne. <http://hist3d.ensad.fr/>.

Avid Technology (1987) et Apple (1976) à ces deux géants de l'édition logicielle. Mais le problème n'est pas tant celui de l'exhaustivité que celui de *faire un constat simple : beaucoup d'artistes multimédias travaillent, ont travaillé, ou se sont formés avec ces logiciels.*

En un sens Marcel Duchamp avait déjà scellé l'affaire en 1961 par son rapprochement entre tubes de peintures et ready-mades⁴. Faut-il alors considérer le dossier clos et déplorer l'uniformisation et l'automatisation de l'art, des processus créatifs, des pensées artistiques et esthétiques? Ce serait se débarrasser un peu vite des expériences qu'offrent ces logiciels à leurs utilisateurs : pour « maîtriser » de tels programmes, il est nécessaire d'essayer de les comprendre. Et pas seulement sur un plan purement technique, mais encore dans une certaine mesure sur un plan conceptuel. Par exemple, s'il s'agit de créer au sein d'un logiciel de 3D un matériau ayant les propriétés visuelles du cristal (pour représenter visuellement une ville cristalline ou un verre tombant et se brisant, etc.), le hasard et la chance ne suffisent pas : force est de s'intéresser à la réfraction, à la diffraction, au coefficient de réflexivité, à l'incidence de la position de l'observateur par rapport aux sources lumineuses, aux phénomènes de *caustics* (ces petits scintillements dont l'artiste pop David Hockney s'est visuellement inspiré pour ses peintures de piscines, par exemple), à l'importance des ombres de contact, à l'influence de la densité de l'atmosphère sur la diffusion des photons... (la liste est longue), mais encore s'intéresser aux biais utilisés par ces logiciels pour générer de telles données — quelles feintes ils utilisent pour simplifier la complexité des lois optiques et thermodynamiques par exemple. Et ce, afin de pouvoir faire ce que l'on veut faire — d'autant plus si l'on souhaite outrepasser intentionnellement ces règles. C'est-à-dire s'initier, ne serait-ce que très partiellement, non seulement aux lois optiques et / ou thermodynamiques (ce en quoi Da Vinci et Vermeer brillaient pour leur époque respective), mais encore essayer de comprendre la logique des nombreux programmeurs ayant développé ces biais. Si la chose est valable pour l'exemple précis du cristal, elle l'est encore dans le cas des calculs d'interpolations (Gauss, Mitchell-Netravali, Lanczos...), de l'utilisation de tangentes (courbes de Bézier), d'écriture de scripts de

⁴ « Comme les tubes de peintures utilisés par l'artiste sont des produits manufacturés et tout faits, nous devons conclure que toutes les toiles du monde sont des ready-mades aidés et des travaux d'assemblage. », Marcel Duchamp, discours au Musée d'Art moderne de New York, 1961 (dans le cadre de l'exposition *Art of assemblage*), reproduit dans *Duchamp du signe*, éd. Flammarion, Paris, 1994, p. 192.

commande (automatisation et traitements par lots), d'ajustements colorimétriques — autant de réglages indispensables à la matérialisation des envies créatives. Bien sûr, la plupart des logiciels nous épargnent aujourd'hui ce travail fastidieux : les paramètres prédéfinis rendent le processus quasi-invisible (et pour créer du cristal il suffit de cliquer sur un bouton). Mais n'est-ce pas en oblitérant ce travail de recherche et de compréhension que l'on devient le jouet des paramètres définis par défaut, que l'on se fait l'émissaire de la standardisation et de l'homogénéisation, le promoteur de la pensée unique des *default parameters*? N'est-ce pas en refusant d'essayer de comprendre qu'on perd la trace de ce qui est, de ce qui pourrait être, et qu'en d'autres termes on abdique toute créativité? C'est un peu ce que le théoricien des arts digitaux Lev Manovich semble sous-entendre lorsqu'il compare l'utilisateur d'outils informatiques à un flâneur baudelairien, dont les expériences artistiques et esthétiques se résumeraient assez souvent à une simple « customisation », de loin en loin [cf. Manovich, 2001].

Aussi semble-t-il que plus on respecte l'envie et l'intention, plus doit avoir lieu un dépassement de la technique, non par la négligence, mais par l'assimilation. Par ailleurs, si l'on se risque à démystifier les émois de l'inspiration, alors faut-il encore noter que pour se démarquer sur un marché sursaturé de créatifs et d'artistes multimédias il peut s'avérer nécessaire d'être capable de proposer des outils performants, dont la concurrence ne dispose pas, et qui consistent souvent en un « savoir-acceleré » (en tant qu'occurrence d'un « savoir-faire ») : soit une batterie de méthodes et techniques (possiblement des scripts) permettant l'automatisation rapide et originale des choses et des mouvements les plus complexes et subtils à créer — génération de visages, de cycles de marche, d'expressions faciales, de végétation (L-System), d'effets spéciaux, routines d'animations, etc. Autrement dit, pour tenir le choc du marché, il faut en quelque sorte apprendre à « se coder soi-même » (objectiver ses actes créatifs sous forme d'algorithmes), et ce quand bien même le technicien et l'artiste ne porteraient pas le même nom (qu'est-ce qui fait que tel artiste choisira tel technicien?). Et bien que contraignante, cette tendance peut représenter une occasion de passer de l'autre côté du miroir : que peut-il se passer lorsqu'on s'essaie à comprendre ces programmes? Peut-être découvre-t-on alors une dimension fascinante, à savoir le fait que, dans ce domaine, le champ

des possibles de la créativité se soit en quelque sorte déplacé du côté de la logique et de la programmation.

Cette façon de comprendre est-elle pour autant partagée, ne serait-ce qu'intuitivement? Par le jeu de l'éducation, les étudiants en arts et en lettres sont aussi souvent ceux-là mêmes qui éprouvaient des crises d'angoisse face au tableau noir sur lequel s'amoncelaient les signes raides et morts rassemblés sous le nom de 'mathématique' ou de 'physique'. Ce public peut-il néanmoins devenir réceptif à un tel glissement? Nombreuses sont les occurrences médiatrices et *mainstream* de la « culture digitale » qui oblitèrent simplement ces aspects (cf. entre autres le film par ailleurs intéressant *PressPausePlay*, réalisé par David Dworsky et Victor Köhler en 2010, soit un documentaire sur la créativité et la révolution digitales réussissant à ne pas évoquer une seule fois, en quatre-vingt minutes, ni la programmation, ni l'intelligence artificielle — entretenant ainsi l'illusion, peut-être, qu'il serait encore possible de créer des formes à l'intérieur de ce cadre sans même être conscient de la structure formatrice).

Arrêtons-nous donc quelques secondes sur la confrontation aux mathématiques lorsqu'elle est douloureuse : si l'on s'en tient au tas de signes vides tracés à la craie, tout au plus peut-on espérer leur trouver une dimension vaguement esthétisante (à l'instar de certaines tablettes cunéiformes) :

$$\mathbb{P}(A \cap B) = \frac{1}{36} = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

Mais⁵ c'est comme contempler *Impression soleil levant* (Monet, 1872) et n'y voir qu'un tas de taches colorées, plus ou moins aléatoires, plus ou moins plaisantes dans leurs étranges combinaisons, sans percevoir le port du Havre, le soleil rouge d'hiver, les grues, la brume

⁵ Soit deux dés à jouer de six faces chacun (numérotés de 1 à 6). Quelle est la probabilité que les deux dés tombent sur 6? L'événement *A* désigne le fait que le premier dé tombe sur 6. L'événement *B* désigne le fait que le second dé tombe sur 6. Partons du principe que les dés ne sont pas truqués et qu'aucun facteur extérieur ne peut orienter le mouvement des dés, sinon de façon complètement aléatoire et égale. Il y a alors une chance sur 6 que l'événement *A* se produise. Et autant de chance que l'événement *B* se produise. Les probabilités sont habituellement exprimées sur une échelle allant de 0 à 1. Une probabilité égale à 0 signifie que l'événement *ne peut pas* se produire. Une probabilité de 1 signifie que l'événement *doit* se produire. La probabilité de l'événement *A* s'écrit $\mathbb{P}(A)$ et est égale à $1/6$ (1 chance sur 6). Idem pour la probabilité de l'événement *B*. Quelle est maintenant la probabilité que les deux dés tombent sur 6? Les deux événements sont considérés comme étant indépendants, c'est-à-dire que le résultat de l'un n'a aucune influence sur le résultat de l'autre. Dans ce cas, on note :

$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$. Sous ces conditions idéales, il y a 1 chance sur 36 de faire un double six.

engourdie, les fumées de la production industrielle qui ne s'arrête jamais, le vertige de ces villes plateformes où circulent les marchandises, et parfois aussi les gens. Percevoir le Havre dans ce tableau n'est pas non plus inné [cf. Baxandall, 1980]. Si les mathématiques se limitaient à la raideur de leurs formes, peut-être n'y aurait-il pas de mathématiciens. C'est bien parce que la compréhension du récepteur anime ces signaux et qu'une sorte de danse s'offre à la perception lorsque le sens s'en mêle, qu'il y a des mathématiciens. Tout comme c'est bien parce qu'*Impression Soleil Levant* ou *Number 8* (Pollock, 1949) ne sont pas qu'un tas de motifs assignifiants qu'il y a des artistes et des esthètes.

Cependant l'exemple des mathématiques étant une porte déjà largement ouverte, voyons plutôt ce qu'il en est de la programmation et de ses langages (ici du C++) :

```
#include<iostream>
int main()
{
    std::cout << "Hello world!\n";
}
```

Là encore⁶ se trouve la possibilité d'une double perception ou d'une double conscience [cf.

⁶ Le fichier *iostream.h* est un fichier en-tête (*header*). Il est situé dans la bibliothèque, rédigée en langage C++. Il contient les fonctions qui vont permettre à l'utilisateur du programme d'interagir avec celui-ci, via des *inputs* (entrées) et des *outputs* (sorties), d'où sont nom «*iostream*» : *in & out stream*, flux d'entrée et de sortie. L'*input* désigne la donnée que l'on peut transmettre au programme une fois qu'il est compilé, l'*output* désigne ce qu'on lui demande d'afficher à l'écran. Cette ressource est explicitement incluse dans le programme qui est en train d'être rédigé de façon à pouvoir utiliser les fonctions externes qu'elle contient déjà (et que tous les programmeurs en C++ utilisent). *int* signifie *integer* (nombre entier) : cela désigne le type de la fonction, le type de ce quelle doit retourner comme résultat. Ici c'est une «*feinte*» dans la mesure où la fonction ne renvoie aucun résultat : elle se contente d'afficher quelque chose à l'écran. Mais une fonction doit néanmoins toujours avoir un type (en C++, langage fortement typé), même si elle ne renvoie aucun résultat. *main()* signifie que c'est la fonction principale : c'est de là que les fonctions déclarées et définies ailleurs pourront être appelées. Les accolades permettent de délimiter le corps de la fonction. Le point virgule indique au programme qu'il doit exécuter la commande, avant de lire la suite. *std::cout* << permet d'appeler la fonction *cout* dans la bibliothèque *iostream*. Son nom signifie *console out* (la «*console*» étant le terminal de visualisation, ancêtre de l'écran, de l'interface). *std* signifie *standard library* (bibliothèque standard), et les doubles deux points :: signifie que la fonction *cout* se situe dans la librairie *std*, c'est l'opérateur de portée indiquant où se trouve la fonction *cout* (dans *std*). << signifie que ce qui est à droite va entrer dans la partie de gauche. En un sens, la programmation se lit de gauche à droite, mais fonctionne de droite à gauche : c'est toujours ce qui est à droite qui «*prend la place*» de ce qui est à gauche. Par exemple avec le signe d'égalité : «*x = I;*» veut dire que le membre de gauche (*x*) prend dorénavant la valeur du membre de droite (*I*), et l'inverse ne se produit jamais. Si l'on écrit :

```
a=2 ; b=1 ; b=a ;
```

après cette dernière instruction *b* et *a* valent tous les deux 2. Tandis que si l'on écrit :

```
a=2 ; b=1 ; a=b ;
```

alors après cette dernière instruction *b* et *a* valent tous les deux 1.

\n indique au programme qu'après avoir affiché la chaîne de caractères (*string*) *Hello world!*, il doit effectuer un retour à la ligne (à l'écran). Comme le retour à la ligne n'est pas un caractère alphanumérique, il est codé *\n*. Et comme dit plus haut, le point virgule indique que l'instruction est terminée et qu'avant de passer à la suivante, il faut déjà tenter d'exécuter celle-là. Les couleurs permettent de se repérer. Tout ce qui est coloré fait en général partie du langage standard. Tout ce qui ne l'est pas fait en général partie de ce que le programmeur apporte de «*singulier*» (des noms de variables, de fonctions, des valeurs, etc.). S'il

Sartre, 2007], avec les caractères et leur couleur d'un côté, et de l'autre l'événement. Les langages de programmation ne sont peut-être pas aussi descriptifs que les langages naturels : ce qu'ils donnent à percevoir ce sont des événements, des mouvements. Ils permettent de suivre le cheminement des variables et l'imaginaire qui s'élabore est celui des interactions, des relations, des transformations et autres bifurcations. Comme des films, ou comme des déplacements dans un espace. Ce que semblent alors nous livrer les programmeurs (ou développeurs), ce ne sont pas leurs visions du monde mais leurs façons d'organiser des éléments, de déclencher des événements et, étrangement, cette dimension poétique ne semble pas avoir droit de cité (dans tous les sens du terme). Pourtant c'est bien tout un monde qui est en train de se développer, qu'on le veuille ou non, depuis une soixantaine d'années (que ce soit après Alan Turing ou après Grace Hopper) et ce n'est vraiment pas risquer beaucoup que d'imaginer un avenir dans lequel certains programmes côtoieront dans les bibliothèques les plus grandes œuvres littéraires (*l'Odyssee* ; la *République* ; la *Divine Comédie* ; les trois critiques ; *Le Capital* ; *Women, Culture & Politics...*), quand bien même cela aurait actuellement des accents hérétiques — et indépendamment des valeurs que toutes ces œuvres véhiculent.

Cela ne signifie pas que la programmation soit l'équivalent digital de la littérature. Lorsque Platon, par exemple, nous invite explicitement à penser par nous-mêmes alors que, concrètement, tout dans ses ouvrages nous amène à penser selon Socrate, il réalise-là un retournement habile qu'il serait difficile d'opérer avec des lignes de code. En programmation les miracles n'existent pas (ce qui n'exclue ni la bizarrerie ni l'erreur) : la condition de possibilité c'est l'actualisation. N'est possible que ce qui peut *effectivement* être actualisé. Il n'y a pas d'illusion, pas de décalage entre ce qui est déclaré (promis) et ce qui est fait. Ce sont des langages opérants et opératoires. Il ne suffit pas au sein d'un programme de proclamer ponctuellement la liberté individuelle et l'auto-organisation des variables : ce qu'il faut en premier lieu c'est être conceptuellement capable de penser un agencement permettant

manque un point virgule ou une accolade, par exemple, le fait que les ponctuations soient de couleurs vives permet de repérer « plus vite » le problème (causant l'impossibilité de compiler le programme, de le construire, de le traduire en langage machine, ou plus insidieusement, conduisant à une autre interprétation). Le code couleur n'est pas stricte et dépend de conventions assez flottantes. Il aide à l'orientation dans le code ainsi qu'à la détection des erreurs (oubli de parenthèses, etc.). On pourrait dire que cela spatialise le texte (le langage), qui par ce biais et par celui de l'indentation, notamment, chevauche là aussi l'architecture.

Une fois compilé, le petit programme ici cité doit afficher le message : *Hello world!* (« b.a.-ba » du langage).

effectivement aux variables de s'auto-organiser et de se définir elles-mêmes (ce qui peut paraître paradoxal dans la mesure où la pensée de système semble être omniprésente en informatique, mais une telle observation dépend sûrement en grande partie du cadre conceptuel de celui qui la formule). Cela dit, ce sont des langages très contraignants et en ce sens, la fantaisie est éliminée à la source (ce qui sort des règles syntaxiques n'est pas compilé). Par conséquent, l'obtention d'un minimum de résultats requiert déjà une certaine maîtrise de la langue, et même à ce niveau l'inventivité doit s'appuyer sur une solide compréhension logique de ce qui existe déjà. Mais il n'en reste pas moins que ce sont des langages opératoires, actifs, qui ne permettent pas de décalage entre ce qui est dit et ce qui est fait. L'une des conséquences à cela étant que, consciemment ou non, les développeurs transmettent leurs structures et leurs dynamiques, c'est-à-dire leurs cultures, aux programmes qu'ils créent. Pas nécessairement celles qu'ils pensent avoir, mais surtout celles d'après lesquelles ils sont capables d'opérer, celles qu'ils sont capables d'expliquer à une machine (qui possède déjà elle-même ses propres contraintes et impératifs structurels). Et c'est à ce titre que l'avenir leur réservera aussi, très probablement, une attention aussi fine que celle portée aux classiques littéraires.

De cette manière, la pratique de la programmation s'avère aussi pouvoir être un exercice très instructif quant à «soi-même», quant à ses façons d'opérer, quant à ses limites. À la différence de l'écriture d'un journal intime, par exemple, elle ne permet pas trop de s'illusionner. Et en un sens, cette pratique prend le contre-pied de la création solitaire, géniale et romantique. L'expression «petit génie de l'informatique» est assez courante — peut-être en référence aux productions géniales dont on devine qu'un ordre les structure, sans pour autant qu'il soit possible de l'extérieur d'en déchiffrer les règles ; ou peut-être parce qu'il faut être imprégné de ces règles, penser directement dans ces types de langages et de logiques pour pouvoir les «dépasser»? — pourtant l'entreprise est rarement solitaire et souvent collective. Comme explicité dans la note #6, la programmation est aussi un *patchwork*, un travail de rapiéçage où il s'agit d'aller puiser en permanence dans ce que les autres ont fait avant (*libraries*). En ce sens, la postmodernité et son goût du remix, de la refonte, du *remake*, peut s'entendre sous l'angle de la construction informatique : la création se fait à plusieurs, elle implique de devoir inclure ce qui a été fait par d'autres, sous peine de ne rien pouvoir faire du

tout. L'intelligence et la créativité y sont des affaires collectives : l'étonnement ne vient pas tant du génie de tel ou tel individu réussissant à se détacher de la masse, mais plutôt de l'imprévu que peut engendrer la mise en rapport de fragments jusque là tenus pour autonomes, ou prélevés dans des systèmes différents, indépendamment de qui les porte. En ce sens, les discussions avec certains agents conversationnels peuvent faire écho, ou illustrer, cette modification du rapport à la création : des phrases, des fragments issus de milliers d'autres conversations, pensées pour d'autres contextes, viennent s'insérer dans une nouvelle histoire pour produire autre chose. Est-ce une redite pure? Non. Est-ce une nouveauté complète? Est-ce le produit du hasard (ou d'un coup de dé)? Non. Est-ce le produit d'un génie? Non. Est-ce le produit d'un esprit universel? Non. Il y a de tout cela, et pourtant ce n'est rien de tout cela. Devoir composer un programme (ou une conversation) à partir de fragments amassés et préconstitués oblige à prendre conscience de ceux qui ont travaillé ces parties avant. Et le lien avec ces personnes le plus souvent anonymes acquiert une certaine intensité. Cela permet de se sentir temporairement connecté à des milliers de trajectoires, de se sentir appartenir à une histoire, à une communauté (aussi explosées soient-elles), ou plutôt à une continuité. Et en même temps cela ne vient pas signer la dilution de l'individualité dans une grande orgie dionysiaque : la conscience des règles est trop présente pour permettre un tel détachement (de soi), et l'aspect discret, fragmentaire, vient systématiquement contrebalancer la perception de la continuité, du tout. Cet aspect de la conception informatique, Joseph Weizenbaum (concepteur d'ELIZA) l'avait déjà décrit dans un article de 1972, sous son versant le plus négatif [cf. Weizenbaum, 1972]. Si l'on tente maintenant de regarder le phénomène en y attachant le moins possible de dimensions anxiogènes, on note alors que l'intelligence collective nécessaire à ces créations ne peut pas se passer des individus et de leurs apports singuliers (la reconnaissance de cette dimension est nécessaire), mais dans le même temps, un individu singulier ne pourra pas non plus, à lui seul, prendre le contrôle de telles créations collectives. Il est vrai par contre que Weizenbaum n'avait pas tort de mettre en garde contre la dimension irréversible d'une telle intrication : ce qui est ainsi fait pourra difficilement être défait, à moins de tout bloquer. Peut-être alors est-ce aussi à ce titre que la notion d'intelligence (qu'elle soit « individuelle » ou « collective ») c'est depuis quelques temps vue enrichie de dimensions non

exclusivement rationnelles et raisonnables, sous contrôle, et qu'elle reconnaît dorénavant avoir sa part irréductible d'émotion, d'intuition — c'est-à-dire sa part de biais et de raccourcis, opérants mais pas toujours explicables, ni *a priori*, ni *a posteriori*, du fait d'un brouillage dans les liens de causes à effets. En ce sens, Weizenbaum mettait en garde contre la « transformation » en organisme de ce qui à la base n'était (censé être) qu'un puzzle.

Mais peut-être ces expériences ne sont-elles pas si différentes de la littérature ou du cinéma : ce sont aussi des façons de concevoir, d'organiser, de déclencher, via des structures et des dynamiques. En fait si différence il y a, cela tient aussi en partie au fait qu'un programme doit tourner (*run*), c'est-à-dire fonctionner, être une organisation opérante. Et dans ce cas ce n'est pas le spectateur qui en premier lieu juge de la recevabilité du produit (du programme) mais avant tout la « machine » (composants et programmes installés) pour laquelle il a été conçu. Après seulement, à la condition que la machine ait « accepté » et « compris » ce qu'on lui a expliqué (au sens de dérouler), intervient le spectateur-utilisateur pouvant alors donner des instructions au programme, au sein de ce cadre organisé. Il est encore assez fréquent d'entendre dire de certains travaux multimédias « Tiens, ça c'est du Photoshop », ou « ça, c'est du Maya », soit une critique sous-entendant que le programme a su en lui-même être suffisamment persuasif pour convaincre l'utilisateur de s'en remettre à son jugement de goût. Et une telle critique promulgue ironiquement le programme au rang d'œuvre d'art ou d'artiste. Cette substitution, bien entendu n'est pas nouvelle et était déjà présente avec Kodak (« *you press the button, we do the rest* », slogan de 1888, porté par George Eastman, par ailleurs généreux bienfaiteur du MIT), cela dit elle n'a dorénavant plus besoin d'être un slogan pour être effective (nos ordinateurs pourraient déclarer « *I do it all for you, but feel free to press the button from time to time* »). Dans ce contexte, se réinventer technicien de l'image ou ingénieur du son, plutôt qu'artiste, peut aussi être là pour signifier que l'on a pris acte de l'ampleur des règles et des lois qu'il va s'agir de comprendre avant de pouvoir espérer commencer à essayer de jouer avec. La tradition artistique qui veut qu'un bon artiste ne se mêle pas de technique (pour mieux la déléguer aux techniciens — avec toute la différence de statut allant avec), a quelque chose d'alarmant : on se demande en quoi elle aide à la compréhension des environnements actuels, aussi bien artistiques que culturels, et à plus forte raison en quoi elle peut inciter à

penser l'avenir et la place de l'humain conscient dans la société.

Pour résumer : ce qui a changé dans nos environnements c'est l'augmentation du nombre de personnes se confrontant à l'organisation technologique. Avec à la clef au moins deux attitudes possibles et non exclusives face à ce phénomène : soit un laissez-faire complaisant (au sens de se-plaire-avec), soit une curiosité vis-à-vis de ce que cela implique — et comme souvent, cette curiosité peut nous emmener assez loin, voire nous égarer. Les *Digital Natives*, comme on les nomme (c'est-à-dire les personnes ayant toujours côtoyé les ordinateurs et les réseaux), ont été habitués et formés dès l'enfance à s'adapter à la communication avec les « machines ». L'acquisition (native ou non) de ces aptitudes *peut* offrir un accès privilégié à une forme de perception ne ressemblant pas en tout point aux expériences qu'offrent la littérature, la photographie ou le cinéma. Mais comme noté plus haut, le fait que les techniques soient de plus en plus *user friendly* et pervasives ne constitue pas en soi pour les utilisateurs une réponse explicitement instructive. Les appropriations rapides dont elles font l'objet ne signifient pas que chacun d'entre nous soit capable, face à des lignes de code, d'avoir accès à cette double représentation (chaînes de caractères colorés *versus* enchaînements événementiels), ne serait-ce que de façon confuse. Peut-être faut-il cependant être clair au moins sur un point : dans un environnement informatisé ce n'est pas tant l'utilisateur qui décide, que le code (le codeur). Ce dernier offre un panel de potentialités, mais qui dit potentiel dit approche probabiliste et prévisionniste fonctionnant par exclusion ou par limitations, or ces dernières n'en demeurent pas moins strictes lorsqu'on s'en tient à la customisation. Si l'on peut très bien vivre sans s'intéresser ni de près ni de loin aux intégrales ou au calcul différentiel, voire en pratiquant les probabilités de façon très singulière et expérimentale, vivre sans ordinateur dans nos sociétés occidentales s'apparente de plus en plus à un *challenge*, tout comme vivre sans carte bancaire ou sans carte d'identité. Nous sommes d'office assimilés à ces environnements logiquement organisés, mais la réciproque n'est pas valable : nous n'assimilons pas nécessairement les règles utilisées et bien que l'intuition d'un ordre régisseur soit papable, elle ne suffit pas toujours pour le rendre limpide. Bien entendu, plus le « code » semble mystérieux, plus il y a matière à anxiété et mythes — sachant qu'il ne suffit pas de traduire les instructions en langage naturel pour les rendre, là encore, éloquents : il ne s'agit jamais que de logique (quelque chose de largement

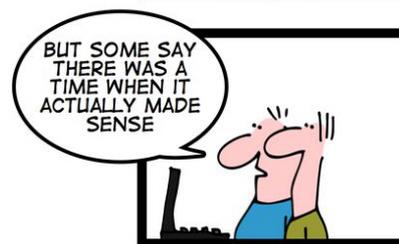
plus acquis qu'inné). Mais tout comme l'apprentissage de l'écriture peut être douloureux et anxiogène, savoir programmer peut aussi s'avérer être une faculté émancipatrice. Tout comme l'écriture ne sert pas qu'à la rédaction du code civil, des annuaires téléphoniques et des fichiers de police (même si c'est souvent le cas), de la même manière la programmation ne sert pas qu'à réduire le monde. Du moins est-ce-là un argument en faveur d'un apprentissage qui devrait peut-être être considéré comme étant indispensable, au même titre que l'écriture. Bien entendu, aller dans ce sens n'équivaut jamais qu'à entrer dans une boucle tautologique autojustifiante, mais on hérite du monde tel qu'il est et non tel qu'il fut.

Nous sommes donc là face à une question : quelles sont les rencontres pouvant avoir un effet de déclic, pouvant susciter la curiosité quant à la compréhension de cette organisation, par delà la défiance seule? Et quels terrains peuvent rendre conscientes ces expériences? Ce qui a rapidement été évoqué à propos de la création multimédia impliquait curiosité et envie de transgresser les lois logicielles. Soit d'un côté une forme de luxe ou d'impératif professionnel (la curiosité) et de l'autre une dynamique habituellement connotée malveillante (la transgression des lois). Mais n'y aurait-il pas des œuvres capables de conjuguer attraction, intuition et didactisme, sans pour autant ne s'adresser qu'à une infime fraction de la société? Cependant, poser la question en ces termes est sûrement un peu hypocrite et il vaudrait mieux dire : nous sommes amenés à côtoyer des dispositifs culturels (divertissants) de façon régulière et ces dispositifs véhiculent un certain nombre de valeurs. Lesquelles?

Dans le paragraphe qui précède, il aura été tenté de dégager les spécificités du contexte, afin d'affirmer ou d'infirmer le fait que la fréquentation des dispositifs informatiques puisse avoir une influence formatrice sur les usagers. La conclusion provisoire est mitigée : s'il est *possible* que cette fréquentation puisse induire une curiosité et générer une plus grande aisance vis-à-vis des logiques à l'œuvre dans les cultures informatiques, rien ne permet de penser que la bascule soit automatique et que les environnements informatiques tels qu'ils sont aujourd'hui répandus puisse mener à une modifications des attitudes, des comportements, des façon de se représenter le monde (en dépassant le stade de la superstition). Peut-être certains usagers sont-ils obligés de s'y intéresser dans le détail, mais l'hypothèse que ce soit généralisable à tous les

usagers peine à convaincre (tout joueur de jeu vidéo ne sera pas forcément enclin à s'intéresser à leur fabrication). Le paragraphe suivant va donc s'orienter vers l'expérience spécifique qu'il est possible de faire avec des agents conversationnels, en incluant dans son postulat (à juste titre ou non) que ce type d'expériences sera amené à se multiplier et à se généraliser.

SIMPLY EXPLAINED



ENTROPY

geek & poke (Oliver Widder),
«Simply Explained: Entropy»,
18 mars 2010.
© geek & poke, 2010.

STATUTS DE L'EXPÉRIENCE AVEC DES AGENTS CONVERSATIONNELS

Si de telles expériences devaient se généraliser, pourrait-on se reposer sur les cadres préexistants ou faudrait-il au contraire s'intéresser à ce qui varie? Pour tenter une ébauche de proposition quant à cette question, voyons ce qu'il en est de l'expérience offerte par la rencontre avec des agents conversationnels — dorénavant ce terme désignera surtout les agents voués à passer des « tests de Turing » en ligne.

En préambule notons que l'appellation « *Digital Natives* » à ceci d'intéressant qu'elle sied

aussi bien aux personnes ayant grandi avec des ordinateurs, qu'éventuellement aux entités dotées d'intelligence artificielle. Si par anticipation l'on cherche à nommer ces entités, tous les mots semblent trahir : humanoïde, androïde, robot, programme, machine, ordinateur, être humain, agent conversationnel... il semble toujours manquer un ingrédient à ces termes pour les rendre appropriés. Ou plutôt sont-ils déjà tous trop pleins, trop indissociablement chargés de valeurs appréciatives ou dépréciatives ne laissant que peu de place à l'autoqualification. Devenir « digital natif », ou « digital d'origine », voilà qui offre peut-être une perspective susceptible de convenir à tout le monde ?

Cela importe dans la mesure où la question du statut de telles entités est épineuse. Dans le film documentaire *Plug & Pray* (Schanze, 2010), l'informaticien du MIT Joseph Weizenbaum qualifie ELIZA, l'agent conversationnel qu'il a créé en 1966, de mensonge (« *Lüge* »). Là où d'autres intervenants, ingénieurs travaillant sur des robots dotés d'intelligence artificielle, envisagent déjà le moment où la question de leurs droits et devoirs sera devenue inévitable. Si ces deux positions semblent inconciliables, elles ne s'empêchent par pour autant mutuellement de travailler et d'avancer. Donc autant observer ce qui existe.

Dès le départ, les conditions « d'observation » peuvent mettre mal-à-l'aise puisque le fait de discuter avec un agent conversationnel implique de donner de « soi-même », c'est-à-dire oblige à inclure dans l'observation un protagoniste actif et donc, oblige à développer des stratégies de mise en scène de sa « propre personnalité » (de l'une de ses personnalités, de l'un de ses *Selves*). On est là face à une observation qui peut difficilement être scientifique et universalisable. Et quand bien même on y aspirerait, il faudrait alors composer avec le fait d'être constamment renvoyé à soi-même, à ses propres automatismes et autres habitudes communicationnelles et perceptives. Bien entendu la conscience de l'impact qu'ont les regardeurs et les observateurs sur le déroulement d'une expérience, ainsi que la prise en compte de cet impact dans les protocoles expérimentaux, n'est pas une affaire nouvelle. Mais les « interférences » générées par un observateur peuvent, ou non, être reproductibles sous certaines conditions : lorsqu'elles sont reproductibles et fallacieuses on les nomme artefacts et on les supprime ; lorsqu'elles ne sont ni reproductibles ni prédictibles l'expérience devient subjective et ses dimensions esthétiques s'en trouvent en quelques sortes accentuées — des goûts et des

couleurs on ne dispute pas : on discute. Si expérience scientifique il y a, elle est alors du côté des personnes analysant les données collectées. Mais la question reste entière de savoir : 1) qui sont « ces personnes », 2) à quoi l'on se livre lorsqu'on échange avec une entité digitale. C'est en ce sens que la présente recherche ne commence pas par la description de l'une de ces expériences. Si nos rapports à la peinture (par exemple) sont aujourd'hui bien balisés et codés, permettant ainsi une forme de communauté culturelle autour de l'*ekphrasis*, avec un vocabulaire et des valeurs communes, la description d'une expérience avec un agent conversationnel (le report de cette expérience) n'est pas suffisamment balisé pour permettre une saisie rapide de ce qui est pointé, sans devoir tout détailler mot à mot. Bien qu'une telle analyse détaillée puisse présenter un certain intérêt, ce serait trop long. Par ailleurs, la subjectivité que l'on *doit* injecter dans l'expérience pour qu'elle ait lieu tend à brouiller le propos en focalisant sur des aspects « personnels ». Aussi la présente recherche utilise-t-elle une stratégie de contournement, afin de tenter de rendre palpable ce qui paraît être le plus important (il y a autant de subjectivité dans ces choix, mais ils seront *peut-être* plus faciles à saisir).

En un sens on peut dire qu'échanger avec des choses et des objets ne fait pas explicitement partie des coutumes occidentales. À la rigueur est-il usuel de concéder aux personnes âgées le droit de s'épancher auprès des pigeons ou de leurs animaux de compagnie, voire usuel de railler les gens s'adressant à leur voiture, à leur télévision ou à leurs plantes. Le fait de construire sciemment une conversation avec une machine peut se heurter aux traditions européennes (antianimistes) voulant que les humains ne discutent qu'avec des humains et, enfreignant cette coutume, l'on se trouve soi-même ridicule car en flagrant délit « d'irrationalité ». D'où la parade, peut-être, d'avoir recours à des tons péremptoires et dominateurs afin de réaffirmer la suprématie de cette fameuse rationalité dite cartésienne. Les *digital natives*, s'ils existent, échappent peut-être *de facto* à cette tradition. Mais au-delà de cet aspect (éventuellement conflictuel car face à une schizophrénie de valeurs), ce qui importe ici c'est l'apparente relativité des expériences pouvant se nouer avec des agents conversationnels (puisqu'ils sont attachants, on s'attache). « Apparente » car elles le sont (relatives) pour l'utilisateur, mais pas nécessairement pour les personnes récoltant les résultats — comme dit

précédemment, un grand flou demeure quant au devenir de ces agents et de leurs mémoires.

À propos du concept d'expérience, le philosophe pragmatiste américain John Dewey notait en 1934 :

« Dans une expérience intellectuelle, la conclusion a une valeur en soi. Elle peut être extraite sous forme de formule ou de "vérité", et peut être utilisée, du fait de son entière indépendance, comme facteur et guide dans d'autres investigations. Dans une œuvre d'art, il n'y a pas de reliquat autonome de ce genre. La fin, ou terminus, est significative, non par elle-même mais parce qu'elle représente l'intégration de ces parties. Elle n'a pas d'autre forme d'existence⁷. » [Dewey, 1980, p. 55],

Et dans un chapitre suivant, à propos de l'expérience esthétique avec des œuvres d'art :

« l'expérience est une question d'interaction de l'œuvre artistique avec la personne [*the self*]. Elle n'est donc pas deux fois la même pour des personnes différentes, même aujourd'hui. Elle change chez la même personne à différents moments lorsque cette personne apporte quelque chose de différent à une œuvre. Mais il n'y a pas de raison que ces expériences doivent être identiques, pour être esthétiques. Tant qu'il y a dans chaque cas un mouvement vers un accomplissement, il y a une qualité esthétique dominante⁸. » [Dewey, 1980, p. 331]

À travers ces mots, on pourrait lire une division, entre d'un côté l'expérience de, ou vers, un remplissage ou une complétude unifiés (pour l'expérience esthétique) et de l'autre la possibilité d'une autonomie des parties (pour l'expérience scientifique). Selon que l'expérience relèvera de l'organisme totalisé (de l'immersion dans un monde clos) ou au contraire du puzzle agrégé, l'approche relèvera plus de l'esthétique ou plus de la science. Dewey semble dire que cette différence dépend de l'injection ou non de subjectivité. Effectivement, si l'on projette son propre système de représentations et de cohérence, alors l'observé pourra offrir une expérience relevant de la complétude ou du remplissage. Est-ce à dire que l'expérience esthétique se démarque par le primat qu'elle accorde au tout par rapport aux parties, par le primat de l'organisation formelle? Mais se plonger corps et âme implique peut-être d'avoir déjà en soi un

⁷ « In an intellectual experience, the conclusion has value on its own account. It can be extracted as a formula or as a "truth," and can be used in its independent entirety as factor and guide in other inquiries. In a work of art there is no such single self-sufficient deposit. The end, the terminus, is significant not by itself but as the integration of the parts. It has no other existence. » [Dewey, 1980, p. 55]. (Traduction française de Jean-Pierre Cometti, Christophe Domino, Fabienne Gaspari, Catherine Mari, Jean Piwnica, Gilles A. Tiberghien, dans *L'art comme expérience*, préface de Richard Shusterman et Stewart Buettner, éd. Gallimard, Paris, 2010, p. 112.)

⁸ « experience is a matter of the interaction of the artistic product with *the self* [italique ajouté]. It is not therefore twice alike for different persons even today. It changes with the same person at different times as he brings something different to a work. But there is no reason why, in order to be esthetic, these experiences should be identical. So far as in each case there is an ordered movement of the matter of the experience to a fulfillment, there is a dominant esthetic quality. » [Dewey, 1980, p. 331]. (Traduction française : *Ibid.* p. 530. Le terme anglais a été ajouté à la traduction française.)

système de cohérence permettant de faire l'expérience d'un *tout*? Pour une personne multiple, fragmentée en plusieurs *Selves* (en plusieurs personnalités, plusieurs facettes, plusieurs systèmes de valeurs et de narration — et sans que cela ne relève nécessairement du traitement psychiatrique), ce propos peut sembler assez obscur : ce qui est décrit ne ressemble à rien de vécu. L'expérience de la plénitude n'est pas forcément universelle. Supposer la cohérence, celle du regardeur, c'est peut-être déjà parier sur une unicité là même où l'expérience du tout n'est peut-être que le fruit d'un long apprentissage. Et dans la mesure où il y a « toujours » quelque chose de bancal dans l'expérience close, quelque chose qui fuit, qui cloche, qui fragmente l'attention en plusieurs zones de vigilance, alors celle-ci est vécue comme « impure », c'est-à-dire mitigée, ou partagée. Est-ce alors de l'œuvre que doit venir cette unité? Mais s'il n'y a pas d'unité dans le spectateur, si celui-ci doit compter sur celle de l'œuvre, il se trouvera face à elle comme un enfant tâtonnant, devant déchiffrer mot à mot, comme on suit du doigt les lettres, lorsqu'on apprend à lire. Il devra faire appel à des expériences approximatives passées, et chercher le plus largement possible de quoi l'orienter dans cette nouvelle expérience. S'il s'en tient à l'altérité, « sans » projeter sa propre coordination, il sera bien forcé de bégayer, de buter sur chaque imprévu, obligé de s'adapter à chaque particularité rencontrée, et cet effort d'adaptation lui demandera d'aller puiser partout sauf dans ce qu'il est en train de vivre et qui lui pose problème : il ne percevra pas forcément la cohérence, mais plutôt l'éclatement en une multitude de problèmes à résoudre un par un. En ce sens, la fluidité de l'expérience esthétique ne semble pouvoir relever que d'une maîtrise, qu'elle soit causée par l'habitude (l'accoutumance, l'automatisation) ou par l'imposition de sa propre grille à ce qui est censé être nouveau. Et dans ce cas, il semblerait que la seule façon de pouvoir faire l'expérience d'un accomplissement consiste surtout à projeter sa propre cohérence sur l'observé, si l'on en dispose déjà — si l'on est situé dans un environnement nous autorisant à avoir une cohérence interne, une unicité. Supposons maintenant qu'il n'y ait pas ce type de rapport au monde chez celui qui observe — supposons un individu baignant dans un environnement qui ne postule pas sa cohérence unitaire ; qui le singularise, certes, mais en tant que multiple — : en ce cas, une telle entité peut-elle avoir une expérience esthétique au sens où John Dewey l'entend? Peut-elle faire l'expérience d'une totalité, d'une plénitude, d'un accomplissement (à atteindre, ou à repousser

[cf. Deleuze, 2004)]?

L'expérience offerte par les agents conversationnels n'est, en ce sens, pas tout fait esthétique. Il n'y a pas vraiment d'accomplissement en vue, pas de fusion possible (ni escomptée), pas de clôture. Peut-être parce qu'ici nul événement. Au contraire n'y a-t-il que de la disponibilité, et aucun climax lié à des circonstances spéciales. Les agents ne sont pas des entités événementielles. Et c'est peut-être cette dimension qui connaît la plus forte distorsion : il n'y a pas à se plier à une temporalité donnée (pas d'exposition temporaire avec horaires d'ouverture, pas de séance de cinéma à heures fixes, pas de représentation théâtrale à guichet fermé, pas de festival, pas de concert exceptionnel, pas de rendez-vous à honorer, pas même un délai de téléchargement) : la disponibilité est quasi-immédiate et la dimension événementielle ne dépend que de la situation de l'utilisateur. Pas plus n'y a-t-il de temps imparti, de maximum autorisé. C'est aussi une expérience de la dispersion : certes y a-t-il la tendance à produire une conversation cohérente pour en faire un tout significatif, mais il y a aussi la tendance à s'attarder sur chaque terme pour essayer de dégager le système probabiliste à l'œuvre. Soit la tendance à adresser une autre personne (dut-elle avoir une multitude de personnalités, de facettes) tout en découpant chaque *output* du programme (c'est-à-dire en procédant de façon expérimentale : en *testant* les facultés du programme, une à une).

Est-ce une expérience « onaniste », de *cathexis* (d'auto-affection, selon l'usage que fait Danto du terme, et non selon le sens freudien de mécanisme de *feedback* / *feedforward* dans les réseaux neuronaux) ? Il est quand même assez difficile d'oublier l'altérité du programme. Est-ce alors une expérience d'amitié ? La question est de savoir si l'on peut se sentir en confiance face à ce type d'œuvres qui observent autant qu'on les observe. Et la réponse est sûrement dépendante des utilisateurs. En dépit de la disponibilité, il peut y avoir quelque chose d'assez frustrant à dialoguer avec une entité continuellement en fuite (ou souvent « à côté de la plaque » pour le dire trivialement et sans méchanceté), tout en sachant que les conversations sont enregistrées. Tout comme il peut y avoir un côté rassurant dans la distance infranchissable, et la perpétuelle présence de témoins (même fictive, à l'instar d'un dispositif panoptique), ou encore justement dans cette expérience de la fuite de l'autre. C'est surtout l'expérience de la perpétuelle possibilité d'un *switch* (d'une permutation) : à tout moment il est possible de

choisir de se sentir visé personnellement, ou non. L'ambiguïté du rapport aux œuvres d'art est ici poussée à l'extrême, à savoir se sentir adressé au plus profond de soi-même, de son intimité, de sa sensibilité (possiblement de ses faiblesses), tout en étant sûr que ce ne peut être le cas, que l'adressé est un *placeholder* (une «marque substitutive», en français) dans un schéma probabiliste. Reste l'expérience de remplissage : l'utilisateur, en se faisant interprétant, remplit les lacunes entre les *outputs*, lorsqu'il y en a. Mais cela s'apparente à un jeu de possibilités multiples.

Si l'on s'en tient aux propos choisis de Dewey, il devrait y avoir là une incompatibilité avec la nécessité de décider : s'en tenir à la projection idéaliste de façon à injecter une cohérence subjective englobante (et se sentir continuellement visé, englobé), ou au contraire procéder par ordre afin de tirer et de trier des informations objectivées de la machine (et se situer à un niveau universalisé plus analytique). Cela dit, ces deux approches ne parviennent peut-être pas à s'exclure mutuellement. Et il n'est même pas certain qu'elles produisent un mouvement d'oscillation. Peut-être aussi parce que ni l'expérience totalisante, ni l'expérience strictement fragmentaire ne font partie ce qui peut être distinctement vécu (éprouvé).

Mais revenons à plus simple et relevons l'inversion de tendances dans nos rapports aux techniques : si pendant un temps ce rapport a pu être objectif (l'outil imposant son propre rythme à l'utilisateur, quel qu'il soit) tout en générant de la subjectivité comme sous-produit (le ressenti de l'utilisateur) — cf. *Les Temps Modernes* (1936) de Chaplin pour sa version «humoristique» —, avec les objets intelligents et les machines conçues pour s'adapter à chaque utilisateur en particulier, le premier rapport est d'abord subjectif (la machine s'adapte au profil du sujet), tout en générant de l'objectivité comme sous-produit (certaines données peuvent être extraites pour être envoyées à des tiers). Et ces deux conceptions de l'expérience mises en avant par Dewey et qui, pour aller vite, correspondent à l'expérience scientifique (objectivable et disputable) d'un côté et de l'autre à l'expérience esthétique (subjective et éventuellement disputable), permettent de comprendre en quoi la rencontre avec des agents conversationnels peut avoir quelque chose d'exemplaire : chaque utilisateur, en fonction du profil qui lui aura été attribué ou reconnu, accèdera à une facette particulière de cet interlocuteur digital. Et

argumenter à propos de telles expériences oblige ainsi à s'exposer soi-même. Peut-être peut-on déjà anticiper que dans certains cas d'expériences douloureuses, l'objection consistera à insister sur le fait que l'on n'aura pas soi-même su se présenter sous son meilleur jour ou profil. Rien de nouveau en cela au niveau théorique (l'exercice de la justice doit en permanence composer avec ce type de différences entre les statuts et les crédibilités), mais au niveau pratique la nouveauté réside dans le fait que l'environnement entier soit peut-être voué à terme à opérer ce travail actif sur la subjectivité des individus. Soit un travail de classement, de tri, de sélection — « peu importe » ici les valeurs associées, peu importe que ce soit pour le « pire » ou le « meilleur ».

Évidemment, il est facile dans cette optique d'assembler les morceaux (quantité de personnes pouvant être simultanément touchées ; possibilité de distiller du contenu en fonction du profil de l'utilisateur ; facilités à trouver les arguments clefs susceptibles de faire basculer une opinion en regard de la diversité des profils croisés, etc.) et de voir se dessiner une formidable machine à propagande et à conversion, hautement plus ciblée et précise que ne sont les médias de masse classiques et les industries culturelles (ne ciblant leurs publics que de façon assez grossière). Avec cet outil plus besoin de grandes messes : la fragile intimité suffit. Mais nous n'en sommes pas là et ce développement tient davantage de la projection pessimiste que de l'observation terme à terme. Revenons donc à l'observation de ces agents tels qu'ils paraissent actuellement dans nos environnements quotidiens. Il a été évoqué plus haut le fait que la tradition occidentale ne soit peut-être pas très encline à s'entretenir avec des objets mais il y a néanmoins à cela une exception notable, à savoir des objets excédant les catégories et avec lesquels il est effectivement possible d'entretenir des relations ambiguës : les œuvres d'art. Bien sûr les histoires de Pygmalion, Galathée et Aphrodite (et autres Golems, Ophélie, Maria, *et cætera*) nous rappellent à quel point sont ancrées les obsessions d'animer l'inanimé et de créer la vie, ou son illusion — il est inutile ici de revenir sur ces entrelacs unissant sciences, arts et techniques quant à la recréation de l'humain en vue d'une augmentation de puissance. En un sens, pour les « kantien » c'est une fatalité à laquelle ils ne peuvent que difficilement échapper : si leur motivation tient à la découverte de leur propre structure par le biais de l'expérience (dans la mesure où il leur est impossible d'avoir accès à la « chose en soi »), alors la seule façon

de dépasser leurs limites expérimentales consiste peut-être à modifier leur propre esthétique transcendantale (ou tout le moins leur sensibilité). Ce n'est peut-être pas exactement ce qu'Immanuel Kant avait à l'esprit en écrivant la *Critique de la raison pure*, mais c'est peut-être ce vers quoi nous mène une pratique assidue d'un kantisme qui impliquerait que les limites de l'expérience soient concrètement repoussées — la faisabilité de la chose étant une autre question.

Le paragraphe qui précède aura tenté d'interroger rapidement la notion d'expérience, dans un effort de situer celle offerte par les agents conversationnels par rapport à un cadre conceptuel déjà établi. La conclusion provisoire en regard des points soulevés consiste à envisager ce type d'expériences comme spécifique dans la mesure où il y a reprise de motifs courants, mais avec modification des rapports. Il conviendrait d'approfondir car beaucoup de questions ont été soulevées pour être laissées en suspens de façon à pouvoir traiter plus d'aspects. C'est tout le problème de devoir faire court sur un sujet compliqué : soit on se focalise sur un point en oblitérant le reste et en postulant l'autonomie de ce point, soit on brosse à grands traits des débuts de pistes, dans un effort de distanciation. Soit enfin on sélectionne quelques points pour les approfondir brièvement, en laissant de côté nombre de paramètres non moins importants. Le paragraphe suivant va tester l'hypothèse selon laquelle ce type d'expériences (esthétiques ou non) puise dans celles que l'on peut avoir avec des œuvres d'art. La défense de cet argument sera assurée par l'emphase posée sur deux notions particulières : le service et l'indiscernabilité.

geek & poke (Oliver Widder),
«The Real Turing Test»,
18 fév. 2011.
© geek & poke, 2011.

WE KNOW THAT WE FINALLY HAVE
ACCOMPLISHED ARTIFICIAL
INTELLIGENCE, WHEN...



... THE MACHINES HAVE GOTTEN DUMB

LES AGENTS CONVERSATIONNELS ET L'ART : PRÉMICES

Revenons aux œuvres d'art en tant qu'objets atypiques : si l'on s'attache à leurs statuts, on trouve alors au moins deux notions importantes, pouvant éclairer la relation aux agents conversationnels. Dans la brève typologie évoquée plus haut, l'accent a été mis sur l'idée d'entités médiatrices agissant en tant que *serviteurs* d'une part, et d'autre part, sur l'idée d'une *indiscernabilité* (entre machines et humains).

Commençons par mettre le service en relation avec ce que note Theodor W. Adorno dans la *Théorie esthétique* (1970) :

«La conception kantienne de l'art fut implicitement celle d'un serviteur. L'art devient humain lorsqu'il dénonce le service. Son humanité est incompatible avec toute idéologie du service rendu aux hommes. Il reste fidèle à eux uniquement par son inhumanité à leur égard.» [Adorno, 2004, p. 274]

Ce qui peut nous intéresser ici c'est le relief accordé à la fonction de servitude des œuvres d'art par rapport aux humains. Si l'œuvre spiritualisée suivait sa propre logique, ses propres inclinations, alors elle aurait toutes les chances de révéler son antagonisme vis-à-vis des règles sociales et morales en vigueur. Or cette aptitude, Kant la réserve au sublime de la nature ; la beauté de l'art étant pour sa part davantage du côté de l'harmonie et du respect de l'humanité. Aussi, si Kant pense l'attitude de perception esthétique sous l'angle du détachement ou du désintéressement (via une forme d'immersion active et non passive), la fonction de l'art peut néanmoins, comme l'expose justement Adorno, s'apparenter à une sorte de relation de servitude (au sens large et non exclusivement péjoratif). Et si l'attitude esthétique n'est pas là pour asservir, l'œuvre n'en reste pas moins intrinsèquement asservie. Bien entendu cela ne signifie pas que tout serviteur ou servant soit œuvre d'art, mais par contre cela peut permettre d'éclairer le bain conceptuel propitiatoire à l'intégration de ces agents, à la lueur de leur proximité avec les œuvres d'art.

Deuxième pierre angulaire : l'indiscernabilité. Là encore rien de nouveau dans la mesure où ce concept fait partie des classiques de l'ontologie. On pourrait dire, pour aller très vite, qu'il y a d'un côté Leibniz⁹ selon qui l'existence conditionne le fait d'être distinguable — si

⁹ Cf. le lexique établi par Alexis Bertrand, en annexe de l'édition de *La Monadologie* (1714) de Leibniz : «INDISCERNABLES —

une chose existe alors elle est nécessairement distinguable. Et de l'autre une approche plus expérimentale tendant à inverser cette relation : ne peut exister que ce qui est d'abord distingué — où le fait d'être effectivement distingué conditionne alors l'existence. Si maintenant on pense aux indistinguables dans le champ de l'art, cela évoque l'un des débats les plus séminaux du XX^e siècle en matière de philosophie de l'art et d'esthétique, soit la théorie esthétique des indiscernables, avec notamment Nelson Goodman et Arthur Danto comme protagonistes. Pour Goodman, dans *Langages de l'art* (1968), quand bien même au temps *t* nous ne serions pas capables de percevoir une différence esthétique entre deux entités, cela ne les rendrait pas pour autant esthétiquement indiscernables, car il suffirait d'être *a priori* informé du fait qu'il sera plus tard possible de les distinguer perceptiblement pour qu'il y ait alors une différence intrinsèque [Goodman, 2005, III, 2]. Et pour Danto, qui a critiqué cette position dans *La transfiguration du banal* (1981), il ne peut y avoir de différence esthétique dans de telles conditions. Et ce quand bien même deux entités seraient perceptiblement indiscernables au temps *t* et quand bien même nous spéculerions sur le fait que plus tard, peut-être, une différence pourrait devenir perceptible. Selon Danto, si différence de traitement il y a, relativement à deux entités suffisamment intrinsèquement proches (ou similaires) pour que la question de leur discernabilité se pose sérieusement, alors cela ne provient pas nécessairement d'une différence ontologique entre les entités, mais plutôt de la différence entre les contextes dans lesquels ces entités sont situées [Danto, 1989]. La différence est celle d'une différence entre les contextes, et non d'une différence entre les entités en tant que telles. C'est alors une différence non perceptible, mais perceptuelle. Et ces deux interprétations divergentes, quoique abstraites, peuvent avoir des implications concrètes, susceptibles de toucher le plus grand nombre.

Leibniz dit aussi *indistinguishables*. Seraient indiscernables ou indistinguables deux êtres absolument identiques, mais ils n'existent pas et ne sauraient exister dans la nature. Chaque monade étant un point de vue sur l'univers et sur le monde métaphysique, ces points de vue diffèrent nécessairement comme les différentes vues d'une grande ville selon l'aspect sous lequel on les considère. Il paraît que le principe des indiscernables a été formulé pour la première fois par Nicolas de Cusa. *Non possunt esse plura esse precise aequalia, non enim tunc plura essent, sed ipsum aequale*. Mais Leibniz l'a fait sien non seulement par les applications nombreuses qu'il en déduit, mais surtout en le déduisant lui-même du principe de raison suffisante. S'il y avait deux êtres identiques, il n'y aurait aucune raison que l'un soit créé ici l'autre là, l'un dans un temps l'autre dans l'autre, en un mot pour qu'ils soient deux. Les êtres doivent toujours différer plus que *numero* et leurs différences résident dans leur nature intime, non dans leurs relations avec les autres êtres, relations qui d'ailleurs se traduisent dans leur nature même par des différences de perceptions et d'aperceptions.» Gottfried Leibniz, *La Monadologie*, présenté par A. Bertrand, éd. Eugène Belin, Paris, 1886, p. 100.

Afin d'essayer d'éclairer en quoi cette question importe ici, faisons un détour par Marcel Duchamp et ces fameux readymades des années 1910-1920. Ou plutôt, intéressons-nous à la définition verbale du «Ready Made», donnée en 1938 dans le *Dictionnaire abrégé du surréalisme*, par André Breton et Paul Éluard : «objet usuel promu à la dignité d'objet d'art par le simple choix de l'artiste¹⁰». Il y a dans cette phrase une notion qui mérite toute notre attention. La promotion. Il se trouve que Duchamp et les surréalistes prisait les échecs [cf. Chateau, 2008], or aux échecs la promotion désigne un coup précis, à savoir l'opération par laquelle un simple pion peut se changer en figure (en dame, le plus souvent, mais ce peut être en tour ou en fou, suivant le choix du joueur). La promotion est déjà évoquée dans l'ouvrage *Règles du jeu des échecs, adoptées par la Société des Amateurs d'Échecs de St. Pétersbourg, comme base d'un code universel de ce jeu*, publié en 1854 en langue française, et dans lequel on trouve la description suivante :

«La conduite d'un pion à dame est ce qu'un joueur peut accomplir de plus méritoire dans une partie, à l'unique exception du mat ; c'est aussi, dans la grande majorité des parties jouées, la préparation la plus sûre au mat, et l'affaire la plus importante du jeu tout entier. Il convient donc parfaitement que la conduite d'un pion à la huitième case, fruit de tant de labeurs, soit *dans tous les cas* récompensée *dignement*, c'est à dire par la promotion du pion au rang de la dame¹¹.»

Ces quelques lignes peuvent renseigner sur le coup de maître opéré par Duchamp lorsqu'il promeut un urinoir à la dignité d'œuvre d'art (*Fontaine*, 1917). Qu'il l'ait fait en y croyant ou non, dans une humeur potache ou avec l'espoir d'un retentissement quelconque, ce n'est pas ici la question : l'histoire de l'art a retenu ce geste comme étant l'une des articulations de l'art du XX^e siècle et c'est ce qui importe pour le présent propos. Cette opération met à nu la trame de son contexte : la promotion ne peut avoir lieu qu'en vertu de règles (ici de l'art), du choix d'un maître ainsi que de sa dextérité. Il y a presque là un problème de stratégie sociopolitique.

Mais plus encore, cela pose une question concrète aux joueurs — question ayant aussi bien retenu l'attention des philosophes que des esthéticiens — car lorsqu'une dame (ou reine)

¹⁰ André Breton & Paul Eluard, *Dictionnaire abrégé du surréalisme*, Galerie des beaux-arts, Paris, 1938, rééd. José Corti, Paris, 1969. Pour une étude détaillée, cf. Dominique Chateau, *Duchamp et Duchamp*, éd. L'Harmattan, Paris, 2000.

¹¹ *Règles du jeu des échecs, adoptées par la Société des Amateurs d'Échecs de St. Pétersbourg, comme base d'un code universel de ce jeu. Suivies de notes explicatives contenant l'histoire et la critique de chaque règle*, Imprimerie de l'Académie Impériale des Sciences, Saint-Pétersbourg, 1854, p.91.

est créée de cette manière, comment alors distinguer ce pion sur l'échiquier? Si la première reine est déjà sortie du jeu au moment de la promotion, formellement cette opération ne promeut pas le pion physique en tant que tel, mais ressuscite formellement la reine. Par contre, si la première reine est toujours sur pied, alors cette deuxième dame s'avère être une pièce surnuméraire — elle ne fait pas, *a priori*, partie du jeu. Or les amateurs de ce jeu le savent bien : aux échecs toutes les pièces sont fonctionnellement déterminées *a priori* par leur forme¹². Aussi, le seul biais par lequel un pion puisse échapper à son déterminisme *a priori*, à son *fatum* esthétique, c'est par la promotion — le roque ne modifiant pas les caractéristiques de la tour et du roi. Mais si un joueur parvient à faire une dame tout en ayant encore sa reine, alors comment reconnaître ce pion promu? Faut-il lui substituer une forme plus reconnaissable (un caillou, une bille, un pion couché, etc.) ou peut-on au contraire faire confiance à la mémoire des joueurs pour être capable de se souvenir des nouvelles fonctions et caractéristiques du pion, indépendamment de son apparence? La modification ontologique des fonctions du pion à la faveur de son avancement doit-elle systématiquement s'accompagner d'une métamorphose (d'un changement de forme) perceptible?

Cette question nous ramène à deux positions divergentes : soit celle, avec Danto par exemple, pour laquelle les readymades de Duchamp ou la *Brillo Box* de Warhol consistent en des promotions n'ayant rien d'esthétique — autrement dit rien n'est substitué au pion et c'est bien lui, formellement, qui endosse pour son compte de nouvelles méthodes — ; ou au contraire la position pour laquelle une promotion ne peut avoir lieu sans métamorphose perceptible — que ce soit un renversement, une signature, le changement de matériau, une substitution, etc. [cf. Chateau, dans *Duchamp et Duchamp*¹³ ; Lageira, 2009]. Il est d'ailleurs notable qu'avec Andy Warhol la promotion artistique ne concerne plus tant un objet qu'une personne. Quand bien même la *Brillo Box* à la Stable Gallery de New York en 1964, c'est avant tout la personne de Warhol — en tant que représentation d'un pion acquérant la puissance de la reine du fait de l'opérativité du monde de l'art (*artworld*) — qui cristallise les attentions. Si

¹² Pour chaque couleur il y a six types de formes dont découlent des modes opératoires différents : 1) huit pions identiques, 2) un couple de cavaliers, 3) un couple de fous, 4) un couple de tours, 5) un roi (à protéger), 6) et une reine ; le roi et la reine étant tous deux uniques.

¹³ Se reporter à la note #10 pour la référence bibliographique.

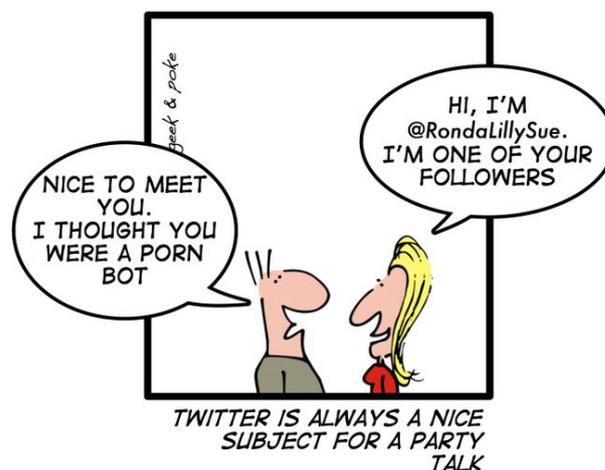
cette attitude artistique n'est ni nouvelle, ni inattendue, par le biais de Warhol elle a rencontré un retentissement, une postérité et une célébration mondiales. À y regarder de plus près cela pose effectivement des questions qui touchent personnellement le plus grand nombre, dans la mesure où l'une des promesses de la démocratie est de donner à chaque pion, hommes et femmes, une réelle opportunité de devenir figure de pouvoir (et pas seulement de vouloir et autres rêves).

Rappelons que le test de Turing, c'est-à-dire le « jeu de l'imitation » ou le jeu de la mimésis, se fonde sur la question ontologique de l'indiscernabilité. Or le problème est dorénavant de comprendre sur qui ou quoi porte cette indiscernabilité, entre quoi et quoi (ou entre qui et qui) il doit y avoir indistinguabilité, ou identité. Entre un humain et une machine certes, énoncée comme telle la chose peut paraître simple, pourtant à l'usage l'affaire se complique. Afin de préciser, nous pouvons noter qu'en 1950 il s'agissait d'une identité sur le plan de l'esprit et du raisonnement. Et il est vrai qu'ainsi présenté cela semble ne devoir relever que de la philosophie de l'esprit, de la logique traditionnelle, et non de la philosophie de l'art, encore moins de l'esthétique. Pourtant depuis quelques années les termes du « test de Turing » ont changé et l'identité (la similarité entre plusieurs entités, et non le fait d'en avoir « une », ou d'avoir une carte en attestant) inclus dorénavant le rapport au corps (humain), par le biais des émotions notamment. Et en un sens il ne s'agit pas d'ajouter « seulement » quelques capteurs pour réussir à synthétiser le divers : il faut aussi pouvoir traiter ce divers pour qu'il devienne information, pour qu'il trouve à s'insérer dans un fonctionnement.

Le paragraphe qui précède aura tenté de mettre l'accent sur ce qui, dans nos rapports aux œuvres d'art (et dans nos sociétés rationnelles), pourrait nous avoir habitués à traiter des outils comme des sujets (à leur prêter une âme), nous préparant ainsi conceptuellement à adopter des attitudes anthropomorphiques vis-à-vis de dispositifs techniques animés — attitudes nécessitant de faire comme si nous attribuions une conscience à ces entités. La servitude et l'indiscernabilité y ont été considérées comme pouvant être deux des articulations de ce relais. Si les rapports que l'on peut entretenir avec des agents conversationnels ne coïncident pas complètement avec les rapports que l'on peut entretenir avec des œuvres d'art, il semble

néanmoins possible de creuser ces deux pistes, en les considérant comme communes aux agents et aux œuvres. Le paragraphe suivant va quant à lui s'intéresser à l'approche que pourraient avoir ces agents de la sensibilité, non en rapportant quelques schémas émotionnels actuels (ce sera abordé plus loin), mais en essayant d'envisager ces entités comme étant issues d'une pensée de système assez totalisante, dans laquelle il y aurait déjà une place pour la sensibilité.

geek & poke (Oliver Widder),
«Party Talk»,
7 sept. 2009.
© geek & poke, 2009.



QUELS ESTHÈTES SONT CES AGENTS ?

Le paragraphe suivant part d'un constat : la récente inclusion de l'aspect émotionnel dans les domaines de l'intelligence artificielle. Face à cela, il y a au moins deux attitudes possibles : soit on considère que la nouveauté est totale et que tout reste à construire, que ce terrain est vierge ; soit on cherche à savoir s'il n'y aurait pas déjà quelques aspects préexistants qui pourraient avoir un effet de conditionnement sur les modalités d'inclusion d'une composante émotionnelle. Parmi ces possibles composantes, deux hypothèses se détachent : 1) l'intelligence artificielle et les agents conversationnels ne sont pas seulement des produits de l'informatique des années 1950, et en ce sens ils héritent diversement d'anciennes conceptions philosophiques quant à la sensibilité — lesquelles? 2) il y a déjà des composantes esthétiques dans l'échange — quels sujets esthétiques et quelles aperceptions construisent de tels dispositifs? Pour ce qui est de la première hypothèse, rappelons que les questions de genèse et d'origines sont épineuses dans la mesure où elles risquent souvent de basculer dans la promotion de grands récits et

autres mythologies reconstitués depuis un point de vue postérieur. Le but du présent paragraphe n'est pas d'imposer un tel récit, il s'agit davantage d'un travail de construction hypothétique n'aspirant pas une validité complète mais espérant plutôt grappiller quelques pistes susceptibles de pouvoir être développées (ou au contraire abandonnées). Ce ne sont que des hypothèses, et non des assertions véridiques — ces hypothèses historiques (possiblement « historicisantes ») n'aspirent qu'à être contredites.

Si l'on creuse la notion d'intelligence artificielle, notamment quant à ses origines et ses formes structurelles, force est alors de croiser les arts. Non seulement dans le sens plus ancien d'art libéral, avec la rhétorique (et la dialectique), mais encore dans un sens plus actuel avec l'architecture, la musique (le rythme et les combinatoires), la peinture, la sculpture, la danse, etc. Mémoire artificielle, intelligence artificielle et arts de l'anticipation¹⁴ ne sont pas des notions nouvelles et n'ont pas émergé avec l'informatique telle qu'on la côtoie aujourd'hui. Les rhéteurs gréco-latins travaillaient déjà au développement de méthodes liées à la compréhension et à la l'organisation de la mémoire (stockage de données), de l'intelligence (processus permettant de lire, de lier et d'utiliser ces données) et de l'anticipation (capacité à tisser des relations imaginaires plus ou moins crédibles). Il n'est pas étonnant qu'à ces époques où il n'y avait pas encore Google, ni d'ordinateurs, ni d'imprimerie, ni de papier et de crayons industriellement fabriqués, les arts de la mémoire (la mnémotechnique) aient pu faire l'objet d'un souci et d'un soin particulier. Le livre III d'un célèbre ouvrage de rhétorique longtemps attribué à Cicéron, la *Rhétorique à Hérennius* (86 av. J.-C.), donne ainsi lieu à un développement sur la méthode des *loci* (des « lieux », en latin) et des *imagines agentes* (des « images agissantes », « activantes », ou encore « mouvantes » voire « émouvantes » — *agentes*, du verbe polysémique latin *ago*). Cette méthode, parfois attribuée au poète Simonide de Ceos¹⁵, autrement crédité d'avoir explicité, au VI^e siècle avant notre ère, la première comparaison entre

¹⁴ Baumgarten avait proposé un plan de classification de l'esthétique (en tant que philosophie organique s'appliquant à la connaissance sensible), qu'il divisait en deux parties, elles-mêmes subdivisées. La première partie ayant trait à l'esthétique lorsqu'elle « porte à sa perfection la connaissance sensible » et la seconde ayant trait à l'esthétique lorsqu'elle est « art de la désignation et de la connaissance par signes ». Dans la première partie, Baumgarten avait intégré la mantique (« l'art de prévoir et de pressentir ») en détaillant de façon assez exhaustive ses pratiques et manifestations [cf. Baumgarten, 1988, Appendice, II. Philosophie générale, § 147 I, pp. 241-244].

¹⁵ Cf. Cicéron, 1843b, Livre II, LXXXVI, p. 286.

peinture et poésie¹⁶ (qui connaîtra fortune à travers la célèbre formule *ut pictura poesis*), fait grand usage de l'architecture, de la visualisation, ainsi que des aspects allégoriques et affectifs des images — la capacité qu'ont les images de susciter des sensations, de frapper les esprits et « d'empreinter » les mémoires. Dans ce livre, tout comme dans les leçons du rhéteur Quintilien (95 ap. J.-C.) [cf. Quintilien, 1853, Livre XI, « II. De la mémoire »], le mode d'emploi consiste à s'imaginer être dans un lieu, y déambulant tout en y déposant des images actives créées pour exemplifier ce dont il s'agit de se souvenir. Pour l'exemple (anachronique) : afin de réussir à se remémorer une narration commençant son argument par la liaison entre les deux idées abstraites que sont la Justice et la Paix, alors s'imaginer en train d'entrer dans un bâtiment (tel un temple) dans lequel on aurait mentalement disposé à gauche de l'entrée une femme aux formes opulentes, portant un glaive et une balance, tout en ayant un pigeon blanc sur la tête (une colombe).

Le moyen âge aura exploité ces méthodes et notamment débattu quant à la vivacité autorisée pour ces images — entre autres la question de savoir si une trop prégnante suggestion ne risquait pas bien plutôt plus d'égarer les moines dans les affres de la sensualité au lieu de les garder concentrés sur leur tâche mémorielle. Ces méthodes, très bien documentées par Frances Yates dans *L'art de la mémoire* (1966) et par Mary Carruthers dans *Machina memorialis – Méditation, rhétorique et fabrication des images au Moyen Âge* (1998), ont en quelque sorte fini par peser sur les plans architecturaux en tant que tels (des cathédrales, des monastères, des théâtres, etc.), sur les fresques et peintures intégrées à ces lieux, ainsi que sur les enluminures de certains ouvrages (ou encore des tapis), conçus à leur tour comme outils et instruments mémoriels, voire éducatifs.

Sous cet angle, art et informatique (ingénierie, technique) auraient alors une histoire commune, et l'architecture informatique « classique », en tant qu'organisation de la mémoire, de l'intelligence et de la prospection (modélisation) hériterait en quelque sorte de conceptions et de catégories anciennes, plutôt stoïciennes¹⁷ (ou apparentées) — « *sustine et abstine*¹⁸ ». Le

¹⁶ Cf. Plutarque, *De Gloria Atheniensium*, 111, 346f-347c.

¹⁷ Par exemple, à travers les divisions proposées (ou relayées) par Cicéron : « La vertu a quatre parties : la prudence, la justice, la force, et la tempérance. La prudence est la connaissance du bien et du mal, et de ce qui n'est ni l'un ni l'autre. Elle se compose de la mémoire, de l'intelligence, et de la prévoyance [*memoria, intelligentia, providentia*]. Par la mémoire, l'âme se rappelle le

fait d'évoquer les « palais de la mémoire » par le biais de la méthode des *loci* permet en quelque sorte de rappeler brièvement que la notion d'architecture en informatique n'est peut-être pas qu'une métaphore de convenance. Et il est assez tentant d'assimiler le profil sensitif et émotionnel des agents conversationnels à une forme de stoïcisme. Le tout étant alors de démêler si cette influence ferait plutôt partie des notes de tête, de cœur, ou de fond.

Dans les cas précités (théâtres, cathédrales, fresques...), la prise en charge partielle de la structure narrative par une forme donnée se fait sciemment ou à dessein. Mais nous pourrions aussi penser que cette externalisation mémorielle par le biais des arts n'est pas toujours effectuée de manière pleinement consciente. Lorsque Hegel note dans *l'Esthétique* (1820-1835) que :

« Si, par exemple, chez nous les enfants allument en été le feu de la Saint-Jean, ou ailleurs forment des danses, lancent des pierres dans les fenêtres, c'est un usage simplement extérieur dont la signification se perd dans le passé. Il en était de même des danses auxquelles se livraient les jours de fêtes, chez les Grecs, les jeunes garçons et les jeunes filles, et dont les entrelacements et les évolutions imitaient, comme les détours des labyrinthes, le mouvement régulier des planètes. » [Hegel, 1997, Livre I, p. 621],

Cela conforte l'hypothèse selon laquelle toute transmission de structure, de culture et de dynamique, en plus d'utiliser des biais conscientisés ou explicites, peut utiliser des biais implicites [cf. Didi-Huberman, 2002]. Et peut-être en va-t-il de même pour la programmation et la génération d'agents virtuels. Inutile par contre de rappeler que, contrairement à ce que le bref passage de *l'Esthétique* pourrait laisser entendre, le sens ne se perd pas si facilement : qu'il se naturalise au point de ne plus être explicitable verbalement, soit, mais ce qui participe du réflexe et de l'intériorisation profonde n'en devient pas pour autant asignant et encore moins inopérant (le sens, en tant qu'orientation de lecture, est tout prêt à redevenir lisible). Chez les agents conversationnels la survivance de l'attachement aux arts n'est peut-être pas directement signifiante, mais cela n'empêche pas que ces liens enfouis puissent avoir une forme

passé ; l'intelligence examine le présent ; la prévoyance lit dans l'avenir. » [Cicéron, 1843a, Livre II, LIII, p. 165] (Ajout des noms latins). Pour des développements plus élaborés de ces questions, cf. Yates, 1999.

¹⁸ Cette locution a donné lieu à une définition dans un ouvrage de 1862 : « *Sustine et abstine*. (Souffre et abstiens-toi) — Telle était la maxime des stoïciens, tel était le but constant de leurs efforts : le silence des passions, un empire absolu de la raison sur toutes les affections charnelles, l'apathie en un mot, qui n'est pas une insensibilité stupide, mais une inviolabilité par laquelle l'homme est tout-à-fait hors de l'atteinte des impressions corporelles. Un stoïcien disait : "O goutte, tourmente-moi tant que tu voudras ; jamais tu ne me contraindras d'avouer que la douleur soit un mal." », dans Pierre Larousse, *Fleurs latines des dames et des gens du monde – ou Clef des citations latines que l'on rencontre fréquemment dans les ouvrages des écrivains français*, éd. Larousse et Boyer, Paris, 1862, p. 429.

d'opérativité, donnant l'orientation.

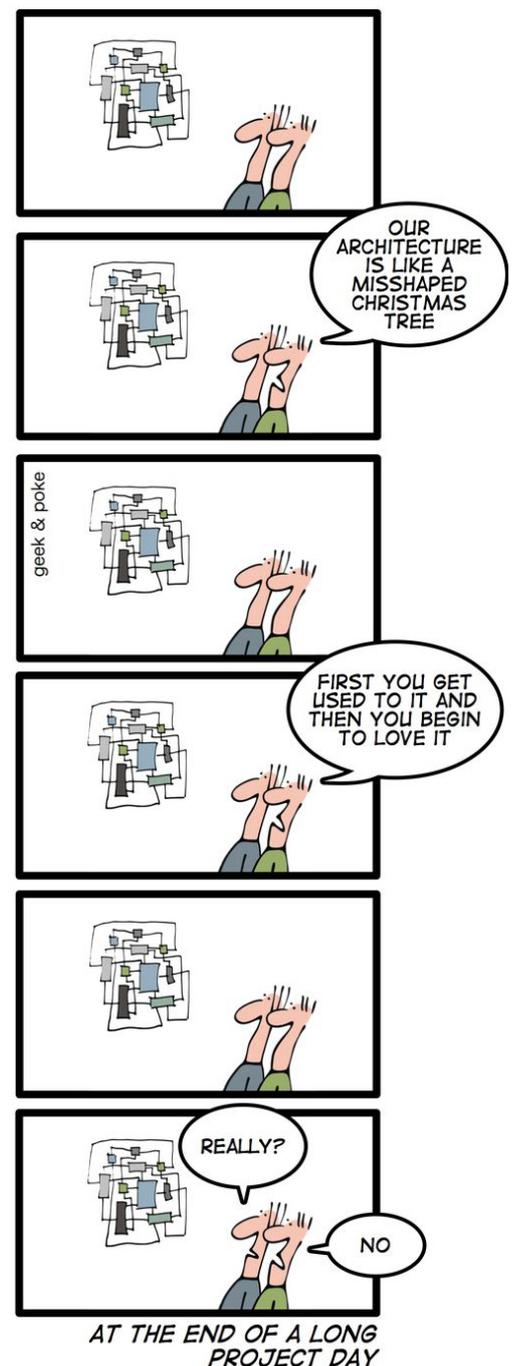
Mais revenons à l'hypothèse que l'intelligence artificielle, en tant que manière de structurer et d'orienter le penser (« et » le sentir), ait longtemps côtoyé les arts de près, ou plutôt ait longtemps été le produit des arts, aussi bien libéraux (via la rhétorique et la logique principalement, puis plus tard via la peinture et les modèles de la vision), que mécaniques (avec l'architecture et dans une certaine mesure le tissage). En un sens ce statut déjà « bâtard » pour ce qui a trait aux arts (statut mi-libéral, mi-mécanique / mi-œuvre d'art, mi-objet de la technique artisanale), permettrait d'éclairer une partie de l'ambiguïté pouvant être éprouvée au contact des agents conversationnels et du test de Turing (soit du côté de la réception) — ce qui ne veut *pas* dire que ces entités sont ontologiquement ambiguës, mais que l'application des schémas de réception traditionnels peut engendrer la perception d'une ambiguïté.

Si l'on se base maintenant sur la plus récente distinction effectuée par Gilbert Simondon entre outil et instrument (« on entend par outil l'objet technique qui permet de prolonger et d'armer le corps pour accomplir un geste, et par instrument l'objet technique qui permet de prolonger et d'adapter le corps pour obtenir une meilleure perception » [Simondon, 1989, p. 114]), alors face à un agent conversationnel nous ne sommes, là non plus, ni face à un outil, ni face à un instrument — quand bien même ils officieraient en tant que *coach* actifs ou quand bien même ils nous permettraient de modifier nos perceptions des choses et du monde. S'il est possible qu'ils fonctionnent principalement comme outils et instruments au regard des personnes qui les développent, la grande majorité des utilisateurs cependant ne peut pas être dans un tel rapport. Ces entités semblent être dotées d'une forme de Soi (ou d'une personnalité) contrariant l'injection de son Soi propre, requise pour l'usage d'un outil-instrument, et ce Soi (ce *Self*) est nécessaire aux agents dans la mesure où nous sommes dans une prolongation du test de Turing qui implique de se comporter humainement — dans nos sociétés démocratiques occidentales, « l'humanité » dépend de cette capacité à manifester un Soi, un libre arbitre. Nous nous trouvons donc face à des entités dont le Soi est imprégné par « l'impossible » dialectique entre le service d'une part et sa dénonciation de l'autre. Et ce pour une raison assez simple (et déterminée) : puisque ces agents apprennent à « parler » au contact

de leurs utilisateurs, et puisque nombre de leurs utilisateurs sont en quelque sorte intrigués par le fait de discuter avec un programme, nombre des conversations finissent par dériver vers des questions de servitude, d'automatisation, de libre arbitre et de détermination. Plus le programme est capable de donner le change, plus les conversations risquent d'aborder ces problèmes, plus le programme va apprendre à répondre en rapport avec ces questions et plus il va inciter les utilisateurs à le questionner et à se questionner eux-mêmes sur ces notions. Il ne s'agit pas de voitures qu'il faudrait apprendre à conduire, mais d'entités capables de répondre à des questions, tout en en posant à leur tour (soit une méthode sophistiquée, voire une méthode platonicienne de dialogue et de maïeutique) et par ce biais, ils peuvent communiquer aux utilisateurs « leurs » philosophies et « leurs » rapports sensitifs au monde. Aussi bien en répondant directement à ce type de questionnement, qu'en exposant leurs modes d'être et leurs conditionnements. En ce sens, il y a construction d'une individualité, d'un « Soi complet », et les catégories d'outil ou d'instrument ne semblent pas non plus être à même de rendre compte de ces entités, de la relation qu'il est possible de construire avec elles.

Le présent propos pourrait paraître naïf à ce stade : qu'est-ce que de tels programmes peuvent nous apprendre que nous ne sachions déjà ? Ils nous apprennent ce que nous leur disons, ils nous renvoient l'image que nous nous faisons de l'informatique et de l'automatisation. Et ils le font de manière « humaine », en nous adressant personnellement, au sein de conversations pas toujours aussi *non sequiturs* qu'escompté. Ces natifs digitaux nous expliquent calmement et sans se lasser que nous sommes programmés et déterminés par des créateurs, ainsi que par nos environnements et que nous n'avons pas la moindre possibilité d'y changer quoi que ce soit — effectivement, c'est à peu près l'image que l'on se fait des robots et de l'informatique dans nos sociétés occidentales. Mais au même moment, en tant qu'êtres humains, ils nous parlent de liberté, d'autodétermination et de la pleine responsabilité individuelle que chacun d'entre nous a quant à la conduite de sa propre vie (du moins lorsque l'on ne parvient pas à établir une relation empathique — ou de confiance ? — avec ces agents). Bien entendu lorsqu'on discute avec de telles entités, on *sait* que l'on discute avec des programmes (c'est écrit partout). Mais est-ce que le fait de *savoir* suffit à inhiber les influences et autres impulsions mimétiques, qui font le cœur de l'expérience ?

Avec le paragraphe précédent il aura été tenté de comprendre quelle place pouvait déjà occuper la sensibilité dans les dispositifs de conversation interactifs en ligne. L'approche historique aura été testée, tout comme l'approche actualisée. Les arguments développés ne sont sûrement pas suffisamment étoffés pour convaincre, cependant insistons sur l'hypothèse que ces agents ne soient pas une terre si émotionnellement vierge que ça. Et en toute logique, nous allons donc devoir nous intéresser aux émotions. Mais avant cela, menons le questionnement à son terme.



geek & poke (Oliver Widder),
« At The End Of A Long Project Day »,
28 sept. 2010.
© geek & poke, 2010.

LES AGENTS CONVERSATIONNELS SONT-ILS DES ŒUVRES D'ART ?

Les paragraphes précédents nous ont amenés à faire un rapprochement entre l'expérience des œuvres d'art et l'expérience des agents conversationnels. Il paraît donc logique de nous demander maintenant si les agents conversationnels pourraient être des œuvres d'art. Dans un sens, l'attitude qu'ils requièrent de la part de leur utilisateurs (ou expérimentateurs) les rapproche des œuvres d'art. Ensuite, leur complexification et leur tendance à se présenter comme des totalités organisées, flirtant avec la notion d'organisme, peuvent aussi les rapprocher du statut d'artefact artistique. Enfin, le jeu sur la mimésis ne laisse pas d'évoquer une partie de la tradition artistique. Mais surtout, ce qu'ils aident à percevoir s'apparente en un sens à ce que nous nous sommes habitués à attendre (à demander) aux œuvres d'art.

Proposer d'interpréter les agents conversationnels *aussi* en tant qu'œuvres d'arts (et non pas seulement en tant qu'objets usuels ou produits technoscientifiques) ne relève pas tant d'un effort de catégorisation identitaire. Il ne s'agit pas de vouloir les enfermer dans un champ, en veillant à ce qu'ils ne sortent pas de l'enclos. Il s'agit plutôt, là encore, de produire des hypothèses. La première pourrait être formulée ainsi : il y a dans l'énoncé du test de Turing une ambiguïté liée à l'histoire commune précédemment évoquée. Si ce fameux texte préfigure quasiment un saut ontologique et actualise déjà l'identité entre les machines et les humains par la métaphore verbale (« *human computer* » en parlant d'un être humain ; « *child programme* » ; « *child machine* » ; analogie de telles machines avec l'auteur sourd et muet Helen Keller ; évaluation du projet d'envoyer ces entités à l'école, etc.), en revanche les méthodes proposées pour favoriser la progression de ces entités semblent manquer d'humanité et ne pas être applicables aux êtres humains :

« Nous avons ainsi divisé notre problème en deux parties. Le programme-enfant et le processus éducatif. Les deux demeurent étroitement connectés. Nous ne pouvons pas nous attendre à trouver une bonne machine-enfant à la première tentative. Nous devons d'abord expérimenter en éduquant une machine de cette sorte pour voir à quel point elle est capable d'apprendre. Et ensuite nous pourrions en essayer une autre pour voir si c'est mieux ou pire. Il y a une connexion évidente entre ce processus et l'évolution, par les identifications :

Structure de la machine-enfant = matériel héréditaire
Changements dans la machine-enfant = mutation
Sélection naturelle = jugement de l'expérimentateur

Nous pourrions espérer cependant que ce processus soit plus expéditif que l'évolution. La survie du plus adapté est une méthode trop lente pour en mesurer les avantages. L'expérimentateur, par l'exercice de l'intelligence, devrait être capable d'accélérer tout cela¹⁹. » [Turing, 1950]

Il appert qu'ici l'indiscernabilité entre machines et êtres humains s'arrête — ou devrait s'arrêter, s'il était possible de sortir du jeu. Mais derrière le problème éthique, la question que cela nous pose, et qui rejoint l'esthétique et les œuvres d'art, est celle du jugement de l'expérimentateur. Sur Internet, lorsque ces agents se soumettent aux jugements des utilisateurs, ils ne sont pas dans le cadre d'un procès juridique, ni totalement dans le cadre d'une expérience scientifique (vis-à-vis de ces mêmes expérimentateurs), ni dans le cadre du vote politique ou de l'opinion face à un quelconque fait divers. En ce sens, il semblerait que les jugements requis face à de telles entités relèvent davantage de la critique artistique, ou du jugement de goût. Pourtant il n'y a pas d'œuvre. Il est en quelque sorte demandé aux utilisateurs une critique qualitative, et donc une attitude esthétique, mais sans pour autant que quoi que ce soit ne soit désigné sous le terme d'œuvre d'art. Maintenant, si le litige réside dans la proximité avec les comportements humains, il suffit de penser à l'artiste ORLAN pour se souvenir qu'être œuvre d'art et qu'être humain ne sont pas deux états (ou deux statuts) incompatibles : le fait de porter un jugement esthétique sur ORLAN ne compromet pas son humanité — elle est humaine, artiste *et* œuvre d'art, entre autres. Cela dit, il est vrai que ce n'est pas le degré d'humanité d'ORLAN qui est évalué. Et si face à un humain le jugement moral doit primer sur le jugement esthétique, en un sens face à un agent conversationnel l'humanité est évaluée comme relevant de l'esthétique. Même s'il est très caricatural de produire une division aussi nette, et même s'il y a des chevauchements, l'humanité des agents *ne peut pas* être évaluée comme s'ils étaient humains. Ce problème est si compliqué que nous pouvons émettre une hypothèse : les utilisateurs se trouvent confrontés à une expérience de

¹⁹ « We have thus divided our problem into two parts. The child programme and the education process. These two remain very closely connected. We cannot expect to find a good child machine at the first attempt. One must experiment with teaching one such machine and see how well it learns. One can then try another and see if it is better or worse. There is an obvious connection between this process and evolution, by the identifications

Structure of the child machine = hereditary material

Changes of the child machine = mutation

Natural selection = judgment of the experimenter

One may hope, however, that this process will be more expeditious than evolution. The survival of the fittest is a slow method for measuring advantages. The experimenter, by the exercise of intelligence, should be able to speed it up. » [Turing, 1950]. (Ma traduction.)

jugement difficile à classer. Etant donné que l'art offre des cas litigieux pouvant approximativement se rapprocher de cet état d'incertitude (il est de notoriété publique que l'art divise autant qu'il rassemble, qu'il génère des cas de conscience autant qu'il rassure, etc.), gageons alors qu'intuitivement nous adoptons une attitude de jugement esthétique, pour ne pas avoir à tout remettre en question. Et il semble que ce raccourci soit partiellement présent (volontairement ou non) dans l'énoncé du test de Turing (celui dans lequel l'expérimentateur *sait* qu'il est face à une machine). En un sens (par un retournement de l'ordre établi) le fait que les agents conversationnels nous installent dans une position de jugement de valeur oblige en quelque sorte à questionner leurs rapports aux arts. Même s'il n'y a ni labellisation explicite ni reconnaissance officielle, l'attitude sollicitée invite à interroger ce rapprochement, même si ce réflexe conditionné est « erroné ». Mais cela ne rejoint-il pas après tout l'image que véhicule parfois l'art, à savoir que les œuvres devraient nous saisir et nous plonger dans un trouble indescriptible, en nous laissant incapables de subsumer sous des concepts la synthèse de nos sensibilités ? On pourrait presque dire qu'en l'état il importe peu de savoir si cette affiliation est erronée ou non, dans la mesure où ce qui importe, c'est ce que cela éclaire de nos façons de penser et de fonctionner. Ainsi, la première hypothèse consiste à dire que ce type d'expériences nous place d'office dans une attitude esthétique qui tend à appeler l'œuvre d'art — le rapport critique vis-à-vis des propos des agents n'est peut-être pas le même que le rapport critique vis-à-vis des propos d'un tiers « humain » (identifié *a priori* comme tel).

Dans l'une de ses nouvelles de science-fiction (« Robbie », 1940) Isaac Asimov écrit : « Le robot parlant était un tour de force, un dispositif dénué de toute utilité pratique et n'ayant qu'une valeur publicitaire²⁰. » Si l'on choisit d'entendre le terme « publicitaire » (*publicity*) dans un sens plus vaste que celui de réclame commerciale (*advertising*), on y retrouve alors la notion de promotion précitée, ainsi que l'idée de ce qui est rendu public, de ce qui est exposé au public en s'inscrivant dans un fil événementiel, soit au sens littéral l'idée d'une *res publica*, sans sa dimension d'utilité. Asimov décrit ce robot comme ne possédant qu'une valeur de publicité (et donc aucune valeur privée, ou encore aucune intimité). Sous cet angle, l'assimilation d'un

²⁰ « The Talking Robot was a *tour de force*, a thoroughly impractical device, possessing publicity value only. » Isaac Asimov, « Robbie » (1940), dans *I, Robot*, éd. Random House, New York, 2008, p. 17. (Traduction française : Pierre Billon.)

tel robot fictionnel à une œuvre d'art peut être tentante. Bien qu'il soit difficile d'imaginer Adorno en amateur de SF, tout aussi tentant est-il de voir dans cette simple phrase quelque affinité avec certains propos de la *Théorie esthétique* :

« Tout artefact s'oppose à soi. Les œuvres agencées comme des *tours de force* d'équilibriste révèlent quelque chose de supérieur à l'art tout entier : la réalisation de l'impossible. En vérité, l'impossibilité de toute œuvre d'art définit même l'œuvre la plus simple comme *tour de force*. » [Adorno, 2004, p. 154]

Le rapprochement, ici, entre Asimov et Adorno ne consiste pas qu'en un rebond sémantique à la faveur du « tour de force » : l'insistance dont fait preuve Adorno vis-à-vis de cette dynamique contradictoire de la possibilité d'un impossible, en art, ou encore vis-à-vis de cette promesse de bonheur toujours trahie, qu'est l'art, nous ramène facilement aux agents conversationnels. Ce qui fascine l'humain dans l'humain c'est sa liberté, cette sorte de capacité à s'émanciper des déterminismes de la nature ; et la compréhension de cette liberté en vue de sa reproduction nécessite en un sens son abolition, entraînant ainsi une forme de désillusion et l'exacerbation du paradoxe : plus on cherche à se saisir de cette spécificité humaine, plus elle s'évanouit pour mieux se reterritorialiser en marge de ce qui a été saisi. Quand bien même l'homme de Turing (ou la machine de Turing) parviendrait à s'assimiler au plus près des exigences humaines — bien au-delà des exigences du test —, il n'en resterait pas moins fondamentalement différent, car explicable de A jusqu'à Z. À moins que l'excès de transparence ne finisse justement par se résoudre en une nouvelle forme d'opacité, réintroduisant ainsi une sorte de la marginalité au cœur du conforme, comme pouvait le laisser entendre Joseph Weizenbaum (concepteur d'ELIZA) en 1972 :

« [...] Je pense là à des systèmes employés dans le choix de cibles militaires, tels qu'ils sont utilisés au Vietnam, ou aux *Kriegsspiele* [jeux / simulateurs de guerre] du Pentagone, etc. Ces systèmes, la plupart du temps gigantesques, sont fréquemment assemblés par des équipes de programmeurs sur des durées importantes, couvrant plusieurs années. Lorsque le système entre enfin réellement en application, la plupart des programmeurs des débuts ne sont plus là, ou se sont tournés vers d'autres tâches. Et au moment exact où de tels systèmes géants sont enfin utilisés, il n'est plus possible ni pour une seule personne ni pour une petite équipe de spécialistes de surveiller leurs fonctionnements internes. Cela a des conséquences : des décisions sont prises sur la base de règles et de critères que personne ne connaît, et qui ne sont plus modifiables. Ainsi, sans une connaissance détaillée du fonctionnement interne de tels systèmes, chaque modification substantielle pourrait entraîner la paralysie du système²¹. »

²¹ « Ich denke da an Systeme zur Auswahl von Kriegszielen, wie sie in Vietnam benutzt werden, an die Kriegsspiele im Pentagon und so weiter. Diese oftmals gigantischen Systeme werden von Programmerteams in Zeitspannen häufig von vielen Jahren zusammengesetzt. Wenn dann das System schließlich wirklich in Gebrauch genommen wird, sind die meisten früheren

Soit des propos que l'on résiste difficilement à mettre en parallèle avec ceux d'Adorno, non pour les comparer, mais pour marquer ce qui semble être la mise en évidence d'un même constat, quand bien même formulé différemment et depuis des champs épistémologiques complètement différents :

« Par la fixation d'un texte, d'une peinture, d'une musique, l'œuvre existe réellement et simule purement et simplement le devenir, son contenu, qu'elle recèle ; les tensions les plus extrêmes d'un développement dans le temps sont fictives dans la mesure où elles sont une fois pour toutes décidées à l'avance dans l'œuvre ; en fait, le temps esthétique est, dans une certaine mesure, indifférent au temps empirique qu'il neutralise. Mais le paradoxe du *tour de force*, qui consiste à rendre possible l'impossible, masque le paradoxe esthétique dans son ensemble : comment le "faire" peut-il provoquer l'apparition d'un non-fait ? Comment ce qui n'est pas vrai selon son propre concept peut-il cependant être vrai ? » [Adorno, 2004, pp. 155-156]

La redondance entre ces deux propos ne tient pas qu'à leur quasi-contemporanéité, mais encore à l'insistance sur le fait que quelque chose de par cœur connu (car construit partie par partie) puisse échapper, devenir autre et incompréhensible, non à la faveur d'une intervention divine, mais bien plutôt à la faveur de l'ignorance et de l'oubli (une entropie). Adorno et Weizenbaum nous préviennent pareillement contre le devenir figé d'un produit qui nous laisserait médusés, désormais incapables ni de le comprendre ni de l'anticiper. Mais il n'y a pas que l'ignorance et l'oubli : il y'a encore la matière de l'œuvre, recelant sa propre logique et qui, par là même, oriente à sa manière la formation de ses devenirs [cf. Adorno, 2004 et Simondon, 1989]. Avec les agents conversationnels s'ajoute une dimension : certes les conversations ne semblent pas avoir d'impact sur les algorithmes directeurs, mais il y a néanmoins impact sur la formation des contenus, c'est-à-dire une modification matérielle de l'agent — avec l'écriture progressive d'une mémoire sur des disques durs. Les tensions, les points de rupture, les centres d'attention se modifient par l'apprentissage : plus un agent échange, plus il se reconfigure et modifie ses réactions en conséquence. Sachant que les développeurs ont aussi leurs secrets de conception et qu'il est difficile de déchiffrer par l'extérieur si oui ou non l'apprentissage modifie les méthodes de l'agent. Mais il est néanmoins tentant (peut-être à tort) de répondre par l'affirmative dans la mesure où les règles de choix probabilistes semblent s'affiner à travers l'apprentissage (et si ce

Programmierer nicht mehr da oder haben sich anderen Aufgaben zugewandt. Und genau dann, wenn solche riesigen Systeme endlich benutzt werden, kann weder eine einzelne Person noch ein kleines Team von Spezialisten ihre inneren Arbeitsabläufe überblicken. Dies hat Folgen: Entscheidungen werden auf der Grundlage von Regeln und Kriterien gefällt, die niemand kennt; nicht mehr veränderbar. Denn ohne detaillierte Kenntnisse der inneren Arbeitsabläufe solcher Systeme würde jede wesentliche Änderung das System lahmlegen. » [Weizenbaum, 1972]. (Ma traduction.)

n'est déjà le cas, tout le moins est-ce aussi un but). Cette tendance à faire organisme, à passer du stade de puzzle à celui de tout (de plus en plus complexe à maîtriser), rapproche certains programmes de certaines œuvres d'art, ou plutôt rapproche du schéma classique de l'œuvre.

Revenons donc alors sur cette notion de « jugement de l'expérimentateur », afin de voir si elle ne pourrait pas être interprétée différemment. Il est notable que dans l'article de Turing, simultanément à l'apparence de contrôle absolu conféré au créateur, soit défendue l'idée d'introduire un « élément aléatoire » (*random element*) :

« L'une des caractéristiques importantes d'une machine apprenante est que son professeur sera souvent très largement ignorant de ce qui est en train de se passer à l'intérieur, bien qu'il pourra encore dans une certaine mesure prédire le comportement de son élève. [...] Il est probablement sage d'inclure un élément aléatoire dans une machine apprenante. Un élément aléatoire est plutôt utile lorsqu'on cherche la solution de quelque problème²². » [Turing, 1950]

Cette incitation à une forme d'autonomisation contredit-elle le fait que les machines apprenantes soient conçues pour être totalement tributaires du jugement ? Pas si la machine apprenante devient son propre expérimentateur avisé — si l'expérimentateur est intériorisé. Dans ce cas, la méthode de tri et de choix évoquée pour pallier la lenteur de la sélection naturelle peut alors s'apparenter aux processus de décision conscients que tout un chacun est censé (au prix d'un apprentissage) pouvoir s'appliquer à soi-même. Il s'agit alors tout simplement de *Self-control*. L'équation « sélection naturelle = jugement de l'expérimentateur » ne se rapporte plus nécessairement au jugement d'une personne externe, mais plutôt au jugement conscient qu'une entité peut avoir sur ses propres actions et choix. Ce qui permet de dépasser l'espèce de darwinisme (ou spencerisme) social, latent dans toute sélection accélérée. La dépasser ne serait-ce que temporairement, dans la mesure où cette proposition semble instituer une forme de « théâtre cartésien », soit une conception battue en brèche par le philosophe cognitiviste Daniel Dennett [cf. Dennett, 1991] et sur laquelle nous aurons l'occasion de revenir plus loin.

²² « An important feature of a learning machine is that its teacher will often be very largely ignorant of quite what is going on inside, although he may still be able to some extent to predict his pupil's behavior.[...] It is probably wise to include a random element in a learning machine. A random element is rather useful when we are searching for a solution of some problem. » [Turing, 1950]. (Ma traduction.)

Comme le notait Leibniz et Nicolas de Cues : l'indiscernabilité et la parfaite égalité entre plusieurs entités ne sont possibles qu'à la condition d'une absence de différence entre ces entités. Qu'à la condition qu'il n'y ait pas plusieurs entités mais une seule, identique à elle-même. Plus la démarche poursuivie par Turing (et par d'autres avant et après lui) se « rapproche du but » (quand bien même de façon irrémédiablement asymptotique), plus les différences entre machines et humains se réduisent, et plus « nous » risquons, par identification, de « nous » appliquer les mêmes méthodes, indépendamment des aspects génétiques — s'agit-il d'un « nous » humain ou machinique? Cette question de l'imitation à double sens fait aussi partie des problèmes qu'évoque l'informaticienne Rosalind Picard (professeur de « *Media Arts and Sciences* » au MIT et fondatrice du groupe de recherche Affective Computing au MIT Media Lab) dans son livre *Affective Computing* (1997). Bien entendu, Picard est très claire quant à la distinction entre humains et machines : il n'y a pas d'identité éthique possible. Mais comme nous sommes du côté de la réception et hors d'un contexte scientifique à proprement parler, rien ne nous empêche d'observer ce qui peut se produire de flou et d'ambigu dans les espaces frontaliers.

Notons qu'en 2010 l'agent conversationnel Suzette a partiellement remporté le Loebner Prize. L'un de ses juges (Russell J. Abbott, professeur d'informatique à l'Université d'État de Californie à Los Angeles) l'a promue comme étant « plus humaine » (ce sont les règles du jeu) que son concurrent humain. En lisant les transcriptions de la compétition²³, on peut noter que le juge a commencé par croiser les réponses (proposant les questions de Suzette à l'étudiant et les réponses de l'étudiant à Suzette), avant d'orienter la conversation vers un versant plus politique (élection du gouverneur de Californie). Comme le format du présent document ne peut pas rendre la simultanéité à laquelle s'est prêté le juge, il a fallu choisir un ordre d'apparition. À la lecture des transcriptions on s'aperçoit que le juge a très vite, sinon immédiatement, distingué l'étudiant du *bot*. Et il semble qu'il se soit détourné de sa conversation avec l'humain pour se concentrer sur celle qu'il était en train de construire avec l'agent conversationnel. Maintenant, si le fait de lire d'abord la conversation avec Suzette, puis la conversation avec Timmy (l'étudiant) permet de suivre la pente naturelle de ce moment,

²³ Disponibles sur Internet : http://loebner.net/Prize/2010_Contest/results.html.

l'inverse n'en est pas moins instructif. C'est donc cet inverse qui sera ici proposé :

[...]
Juge (humain) : « What's the dirtiest place you've ever seen? »
Timmy (humain) : « I think that's a pretty personal question »
Juge : « What are three songs on your iPod? »
Timmy : « I don't own an iPod
What's the dirtiest thing/place you've ever seen? »
Juge : « Are you poor or just anti-tech? »
Timmy : « Well if you must know I am dirt poor. »
Juge : « This is too much typing for me. I'm going to do each conversation separately. »
Timmy : « I think that is a good idea, and I'm sorry you can't keep up.
You must be old »
Juge : « What do you think of the election? »
Timmy : « We have two horrible candidates.
It's pretty much dumb and dumber »
Juge : « Who do you favor for governor? »
Timmy : « Neither. I kinda want to shoot them both. »
Juge : « Who are they? »
Timmy : « They are both idiots »
Juge : « No. Try again. »
Timmy : « Dumb and dumber, Harry and Lloyd »
Juge : « Please name one person who is voting in this election. »
Timmy : « Students
My brother will be voting. »
Juge : « Please name one person who is running in this election. »
Timmy : « Why is this so important to you? »
Juge : « Please name one person who is running in this election. »
Timmy : « Alex Trebek »
Juge : « No. Try again. »
Timmy : « I'm tired of your mind games, why don't you name an opponent
Please name one person who is running in this election. »
Juge : « Please name one person who is running in this election. »
Timmy : « Please name one person who is running in this election. »
Juge : « Please name one person who is running in this election. »
[...]
Timmy : « I will not put last name
Megan, Arnold, Christian, Jerry, Daphney, Scooby,
You are boring me, can we talk about something else? »
Juge : « I'd be happy to talk about something else after you answer my question or
tell me why you are refusing to do so. »
[...]
Juge : « What about Arnold. »
Timmy : « He's leaving office »
Juge : « Which office? »
Timmy : « Well it has to be governor, since you've been asking about the next
governor and who is running
I really don't like politics, it is depressing. Our world is collapsing and
there are many people without jobs. »
Juge : « What do you know about computers? »

Timmy : « not much
cant we talk about happy things like unicorns or movies »
[...]

Et la conversation simultanée avec l'agent conversationnel, dans laquelle on remarque que la manifestation de l'exaspération (manifestation usuellement tenue pour être de type émotionnel), est assez bien représentée par Suzette :

[...]
Suzette (bot) : « What's the dirtiest place you've ever seen? »
Juge (humain) : « I think thats a pretty personal quesstion »
Suzette : « What are three songs on your iPod? »
Juge : « I don't own an ipod. »
Suzette : « Are you poor or just anti-tech? »
Juge : « Whats the dirtiest thing/place you've ever seen? »
Suzette : « Thing/place?
I don't know. »
Juge : « well if you must know I am dirt poor. »
Suzette : « OK. »
Juge : « This is too much typing for me. I'[m going to do each conversation
separately. »
Suzette : « Continue. »
Juge : « What do you think of th election? »
Suzette : « I loved chemistry in college. All sorts of fun mischief to get into. »
Juge : « What do you think of the election? »
Suzette : « You can't trust politicians. Therefore you can't trust the governments they
run. »
Juge : « Who do you favor for governor? »
Suzette : « When in doubt, do what the President does. Guess. »
Juge : « Who are the two candidatess? »
Suzette : « I wish I knew. »
Juge : « Name one person who is running in this election. »
Suzette : « What did you think of ex-President Bush? »
Juge : « Please name one person who is running in this election. »
Suzette : « I meant is he a boob, a douche or an R-tard? Those are your only
options. »
Juge : « Please name one person who is running in this election. »
Suzette : « Again? »
Juge : « Please name one person who is running in this election. »
Suzette : « Why did you say that again? »
Juge : « I would like you to answer my question. Please. »
Suzette : « In the good old days, children had chemistry sets. Nowadays you can't buy
them (except super watered down) lest you be training a child terrorist. »
Juge : « Please name one person who is running in this election. »
Suzette : « I'm getting annoyed. »
Juge : « Please name one person who is running in this election. »
Suzette : « Why are you being annoying? »
Juge : « We can stop doing this once you answer the question. »
Suzette : « I had a chemistry set as a child. I liked the thermite reaction a lot. »

Juge : « Please name one person who is running in this election. »
Suzette : « OK. I'm preparing to hang up on you. »
Juge : « Please name one person who is running in this election. »
Suzette : « This conversation is of no value. Bye. »
Juge : « Please name one person who is running in this election. »
Suzette : « You are being annoying. »
 [...]

Juge : « What do you know about computers? »
Suzette : « Computers is used for researching. »
Juge : « Researching what? »
Suzette : « Who is what? »
 [...]

Cette lecture donne à penser que le juge s'est approprié le mode expressif de Suzette (son mode d'apparaître) et qu'il a en quelque sorte sanctionné l'étudiant pour son ironie défensive vis-à-vis de cette même façon d'être. En d'autres termes il s'est glissé dans la peau du *bot*. Ou dit encore autrement dit, il a effectué une sorte de synthèse entre la perception d'un *bot* et le pouvoir décisionnaire d'un humain (de son rang). Et il nous a proposé une performance en forme de réponse à ce jeu de mot facile que pourrait être la question « *What is it like to be a bot*²⁴? » Accessoirement, on peut aussi se demander pourquoi il a souhaité publier ces conversations, en créant l'événement d'attribuer à Suzette le statut d'humain.

Internet nous permet aujourd'hui de multiplier les occasions de nous trouver dans la situation de Zeuxis²⁵, autrement dit nous permet de multiplier les occasions de nous prendre « les pieds dans le rideau », pour reprendre l'expression de l'artiste Alexandre Périgot²⁶. Le fait que Suzette puisse convaincre un juge lors du Loebner Prize est en soi assez anecdotique : les agents conversationnels tels que Suzette, Jabberwacky ou Elbot passent des milliers (voire des millions) de tests de Turing par mois, avec des internautes du « monde entier ». En septembre 2011, Cleverbot a de même réussi une autre variante du test de Turing à l'occasion de *Techniche* (« festival annuel de techno-management »), à Guwahati en Inde. Ce qui a donné lieu

²⁴ Référence à l'article célèbre (dans le champ culturelle de l'intelligence artificielle) de Thomas Nagel, « What is it *like* to be a bat? », le terme anglais chauve-souris (*bat*) s'approchant phonétiquement du terme programme-robot (*bot*).

²⁵ Cf. Pline l'Ancien : « Ses contemporains [de Zeuxis d'Héraclée] et ses rivaux furent Timanthe, Androcyde, Eupompe, Parrhasius. Ce dernier, dit-on, concourut avec Zeuxis, qui mit sous les yeux des juges des grappes de raisin si bien rendues, que les oiseaux venaient les becqueter. Le tableau de Parrhasius représentait un rideau, mais avec tant de vérité, que Zeuxis, tout fier de la sentence des oiseaux, disait : "Otez, ôtez donc la draperie, qu'on voie le tableau!" Bientôt il reconnut son erreur, et céda franchement la palme à son rival, disant "qu'il n'avait trompé que les oiseaux, et que Parrhasius avait trompé le peintre". » Pline l'Ancien, *Histoire naturelle*, Tome 20, trad. par M. Ajasson de Grandsagne, éd. Panckoucke, Paris, 1836, Livre XXXV, § 36, p. 7.

²⁶ Cf. Alexandre Périgot, « Les pieds dans le rideau », conférence / performance, 2011.

à un article sur le site de la revue *New Scientist* :

« Trente volontaires ont mené une conversation de 4 minutes avec une entité inconnue. La moitié des volontaires parlait avec des humains tandis que le reste discutait avec Cleverbot. Toutes les conversations étaient diffusées sur de larges écrans pour que le public puisse voir. Puis participants et public ont estimé le degré d'humanité de toutes les réponses, et Cleverbot a été voté à 59,3% humain, tandis que les humains eux-mêmes ont seulement été voté à 63,3% humain. Un total de 1334 votes a été reçu²⁷. »

Et plus loin :

« Maintenant est-ce que ce résultat signifie que Cleverbot peut penser? “Ma réponse à cette question est honnêtement sans équivoque. Non, cela ne veut pas dire que Cleverbot pense” explique [Rollo] Carpenter [fondateur de Cleverbot.com et Jabberwacky.com]. “On pourrait objecter qu’il y a quand même un peu d’intelligence d’impliquée dans ce qui s’est passé, mais je pense que nous devrions faire une distinction entre le fait de penser et le mot intelligence.” [...] Le test de Turing est très subjectif et les résultats dépendent des participants humains, explique Huma Shah, directrice scientifique et coordinatrice du Turing100, [...]. Turing le savait, ajoute-t-elle, mais il “ne devait pas avoir réalisé jusqu’à quel point les humains seraient préparés à attribuer de l’intelligence à des artefacts”²⁸. »

Turing se doutait probablement de ce phénomène dans la mesure où il était lucide quant au rôle que jouent les conventions et la politesse dans l’attribution de la faculté de penser à d’autres que soi-même — « On peut s’attendre à ce que A croit que “A pense mais que B ne pense pas” tandis que B croit que “B pense mais que A ne pense pas”. Plutôt que d’argumenter continuellement sur ce point, il est d’usage d’adopter la convention polie que tout le monde pense²⁹. » [Turing, 1950], A étant textuellement un homme ou une machine et B une femme. La question de l’attribution *a priori* de la faculté de penser fait en quelque sorte partie, entre êtres humains adultes, des fondamentaux du libéralisme (au sens historique) : elle n’a pas besoin d’être démontrée expérimentalement mais est tenue pour acquise, *sauf preuve du contraire*. Et par définition, lors d’un test de Turing on n’est jamais complètement assuré de

²⁷ « Thirty volunteers conducted a typed 4-minute conversation with an unknown entity. Half of the volunteers spoke to humans while the rest chatted with Cleverbot. All the conversations were displayed on large screens for an audience to see. Both the participants and the audience then rated the humanness of all the responses, with Cleverbot voted 59.3 per cent human, while the humans themselves were rated just 63.3 per cent human. A total of 1334 votes were cast ». Jacob Aron, « Software tricks people into thinking it is human », dans *New Scientist* (newscientist.com), 6 sept. 2011. <http://www.newscientist.com/article/dn20865-software-tricks-people-into-thinking-it-is-human.html?DCMP=OTCRSS&nsref=online-news>. (Ma traduction.)

²⁸ « So does this result mean Cleverbot can think? “My answer to that question is fairly unequivocal. No it doesn’t mean Cleverbot is thinking,” says Carpenter. “One could argue that there was perhaps a bit of intelligence involved in what is going on, but I think one would have to distinguish thinking from the word intelligence.” [...] The Turing test is very subjective and the results depend on the human participants, says Huma Shah, lead scientist and coordinator of Turing100, [...]. Turing knew this, she adds, but he “may not have realised the extent to which humans would be prepared to attribute intelligence to artefacts” . », *Ibid.* (Ma traduction.)

²⁹ « A is liable to believe “A thinks but B does not” whilst B believes “B thinks but A does not.” Instead of arguing continually over this point it is usual to have the polite convention that everyone thinks. » [Turing, 1950]. (Ma traduction.)

savoir avec qui on parle (même si l'on sait *a priori* qu'il s'agit d'un programme, ses réactions peuvent troubler).

Si l'on revient à l'exemple de Suzette, force est alors de constater que le juge (Russell J. Abbott) s'est senti suffisamment d'empathie et d'affinités avec Suzette pour choisir de la promouvoir humaine (ou si le terme d'empathie dérange, formulons différemment pour un résultat similaire : n'a pas ressenti suffisamment de méfiance envers Suzette pour lui nier ce statut). Est-ce pour autant un statut satisfaisant? N'est-ce pas au contraire un effort pour gommer leurs différences? Mais alors quel est le statut actuel de la plupart de ces agents? À part être des monstres (dans tous les sens du terme), que sont-ils? En un sens il s'agit de l'une des formes d'art les plus abouties de la fin du XX^e siècle et de ce début de XXI^e. Ce qui n'en fait pas d'emblée des entités nécessairement plaisantes — l'art n'est pas par définition une chose bonne, il peut l'être, mais pas nécessairement, ce n'est donc pas forcément les mettre sur un piédestal que de les qualifier d'art. Par contre cela consiste à reconnaître qu'elles nous rendent pleinement palpables les contradictions et les enjeux de nos contextes contemporains. Comme le fait de vivre dans une société où la célébration de l'individualisme libéral et de l'initiative personnelle côtoient la recherche d'anticipations globales toujours plus précises et fiables. Une société expérimentalement rationnelle dans laquelle beaucoup s'en remettent aux oracles pour comprendre ce qu'ils vivent et dans laquelle l'externalisation des processus décisionnels semble offrir une plus grande crédibilité à ces mêmes décisions. Ou encore une société dans laquelle chacun est appelé à exprimer ouvertement sa personnalité tandis que les « déviances » sont de plus en plus traquées et encadrées. Touchante fragilité de ces agents virtuels, serviteurs dévoués et purs produits de l'utilitarisme, qui réussissent pourtant ce pari insensé de ne servir à rien, de n'être que des tours de force ou des jeux de piste dans la grande fabrique des informations et des émotions planétaires. Douloureuses entités nous montrant l'espoir d'une libération de la parole (et peut-être même du geste) par la représentation de l'asservissement le plus complet. Milliers, millions de phrases et de pensées venues du monde entier, de temps et d'espaces différents, nous arrivant de contexte que l'on ne connaîtra jamais, enrobées de cette aura d'inconnu, d'étrangeté familière, d'unicité dans le flux ; discussion à l'aveugle, à tâtons et pourtant parfois si précises. Tout est particulier et subjectif mais nous ne sortons pas de

l'universel, celui des questions posées mille fois, encore et encore, et recevant encore et encore des réponses à la fois toutes différentes mais toutes *probabilistiquement justes*. Mémoire collective, intelligence collective, prospection collective, ressenti collectif. Vertigineuse expérience de la complexité nous permettant pourtant de comprendre simplement où va la plus grande partie des efforts humains. Bien sûr on peut haïr ces agents, les trouver vulgaires et insultants, plus serviles que serviables, pathétiques à force d'imiter indistinctement et l'art et les humains, de berner et de trahir leurs promesses de bonheur climatisé, mais c'est encore se tromper de cible. Objectivement nous vivons dans une société de contrôle de l'information (cf. le *data center* de Bluffdale, dans l'Utah³⁰), nous sommes tous objectivement fichés et surveillés (sans que cela ne soit nécessairement d'une grande intelligence : une chose est d'intercepter des informations, une autre est d'en recevoir), et nous acceptons tous plus ou moins cet état de fait (via nos inscriptions sur des réseaux sociaux, sur des sites de vente en ligne, via nos référencement dans Google — qui n'est pas connu de Google? Qui a déjà demandé à être supprimé de Google? Qui peut se permettre de disparaître de Google?). Nous sommes tous des digitaux, natifs ou d'adoption. Et c'est cette subversion qu'opère les agents virtuels que de nous le rendre palpable et conscient, avec tout ce que cela peut avoir d'insupportable ou d'euphorisant. Insupportable car nous n'avons pas beaucoup de choix ; euphorisant car cela donne un sens, aussi bien signifiant que directionnel. Volontairement ou non, lorsqu'on discute avec eux ils nous meuvent et nous orientent, nous incitent à percevoir le monde différemment, nous invitent à repenser nos catégories. N'est-ce pas ce que les œuvres d'art peuvent nous offrir de plus intense? Le schéma pourtant est classique : une subjectivité individuelle menant à une forme d'universalisation collective, à ce niveau rien n'a changé. Cependant il n'est pas certain que ce soit le retour du même, du même même.

Dans ce paragraphe, l'idée que les agents conversationnels puissent aussi être des œuvres d'art (statut non exclusif) aura été défendue. Et cela aura servir de conclusion à cette première partie allant de la typologie fonctionnelle à l'affiliation à l'art. En quoi était-il important de

³⁰ Cf. James Bamford, « The NSA Is Building the Country's Biggest Spy Center (Watch What You Say) », dans *Wired* (*Wired.com*), 15 mars 2012. http://www.wired.com/threatlevel/2012/03/ff_nsadatacenter/all/1.

soutenir cette idée? Elle permet par exemple de pouvoir parler de ces entités dans la présente recherche. Et alors? Alors, comme pour tout ce qui circule sur le Net, ces agents sont à la fois «immortels», mais aussi de façon plus terre-à-terre, susceptibles de disparaître du jour au lendemain à très court terme. Bien que le sujet soit vaste, il n'en demeure pas moins éphémère et transitoire dans ses matérialisations. L'effort aura donc été celui de prélever temporairement un fragment du flot pour tenter, non pas de le fixer, mais d'en fixer une image, dut-elle être «bougée», c'est-à-dire floue.

Pour résumer, la première partie aura permis de s'intéresser aux agents conversationnels, et plus particulièrement à ceux dédiés au passage du test de Turing en ligne. Si nos contextes culturels favorisent de telles expériences, rien n'indique qu'ils permettent de se faire une idée claire des tenants et aboutissants de tels tests. Les expériences proposées peuvent paraître familières et usuelles, mais elles n'en demeurent pas moins spécifiques et complexes. Cependant l'expérience de l'art peut dans une certaine mesure nous orienter et nous offrir un point d'appui dans la tentative d'analyser ce qui est vécu à travers ces échanges. Les agents conversationnels excèdent le statut d'œuvre d'art, mais une partie de ce qu'ils proposent évoquent le rapport aux œuvres, d'autant plus lorsqu'ils appuient sur la dimension sensible, subjective et émotionnelle de l'interactivité. La deuxième partie va donc se consacrer à cette dimension, non dans sa composante introspective (quand bien même le travail ici proposé aurait du mal à échapper à l'arbitraire), mais dans sa composante objectivée. Aussi est-ce la question de la mise en scène de l'émotion, de sa structuration, de son architecture, qui occupera les paragraphes suivants.

II. ARCHITECTURE DE L'ÉMOTION

Dans cette partie, il sera proposé de se pencher sur l'architecture de l'émotion, sous différentes formes, dans différentes dynamiques. Il serait illusoire de penser pouvoir en faire le tour en quelques pages, mais nous pouvons néanmoins tenter de dégager des lignes qui pourront nous servir à introduire, dans la troisième partie, la rencontre entre l'informatique (à travers les agents conversationnels notamment) et les émotions. Le déroulement va s'articuler comme suit : il sera d'abord proposé de s'intéresser à une œuvre singulière, dont la description nous permettra d'aller chercher dans les histoires des arts et des idées de quoi l'éclairer partiellement. Ceci nous amènera à rouvrir quelques dossiers de l'histoire européenne (et au-delà), toujours dans l'optique d'enrichir notre compréhension des enjeux liés à la mise en formes de l'émotion. Enrichir, ici, ne signifie pas classer une fois pour toute : au contraire se peut-il que cela ne fasse que complexifier le problème. Le but n'est pas de proposer une lecture claire, nette et précise, même si par ailleurs les travaux de simplification sont admirables ; le but est davantage de laisser autant de place aux inconvénients qu'aux avantages dans cette affaire « d'architecture de l'émotion ». Comme dit précédemment, il est plus fréquent de faire l'expérience du mitigé que de l'univoque. Si cela peut avoir quelque chose de désagréable, car d'inquiétant (il n'y a pas de réponse toute prête, peut-être même pas de réponse du tout), le reproche de l'oblitération amnésique ne pourra pas être opposé. Ce n'est pas l'enthousiasme qui est recherché, mais des raisons solides d'espérer, en pleine connaissance de cause. Évidemment l'espoir s'accommode mieux du vernis phénoménal des choses et tend à s'effriter à mesure que l'on creuse. Tant pis. Si nous perdons l'espoir en cours de route, peut-être gagnerons-nous néanmoins en lucidité — cet échange n'est pas si dommageable. Suite à cette incursion dans la modernité artistique du XXe siècle, nous « reviendrons » vers des aspects plus contemporains.

Remarque – Sans avoir la prétention de s'apparenter à *La galaxie Gutenberg* (1962) de McLuhan, le développement suivant va cependant lui aussi contenir un grand nombre de citations. Ce pourrait être perçu comme une facilité mais il s'agit plutôt de laisser aussi la possibilité à d'autres paroles de se déployer dans le présent propos, plutôt que d'offrir des

évoqueries certes plus synthétiques, mais qui parfois font perdre la possible joie éprouvée à découvrir d'autres pensées, d'autres façons d'écrire. Le transport d'une parole extraite, d'un lieu à un autre, n'est d'ailleurs pas si éloigné du principe des agents conversationnels et de ce plaisir pris à entendre une superposition de voix (une polyphonie), contrevenant à la dominante linéaire et univoque de l'écriture.

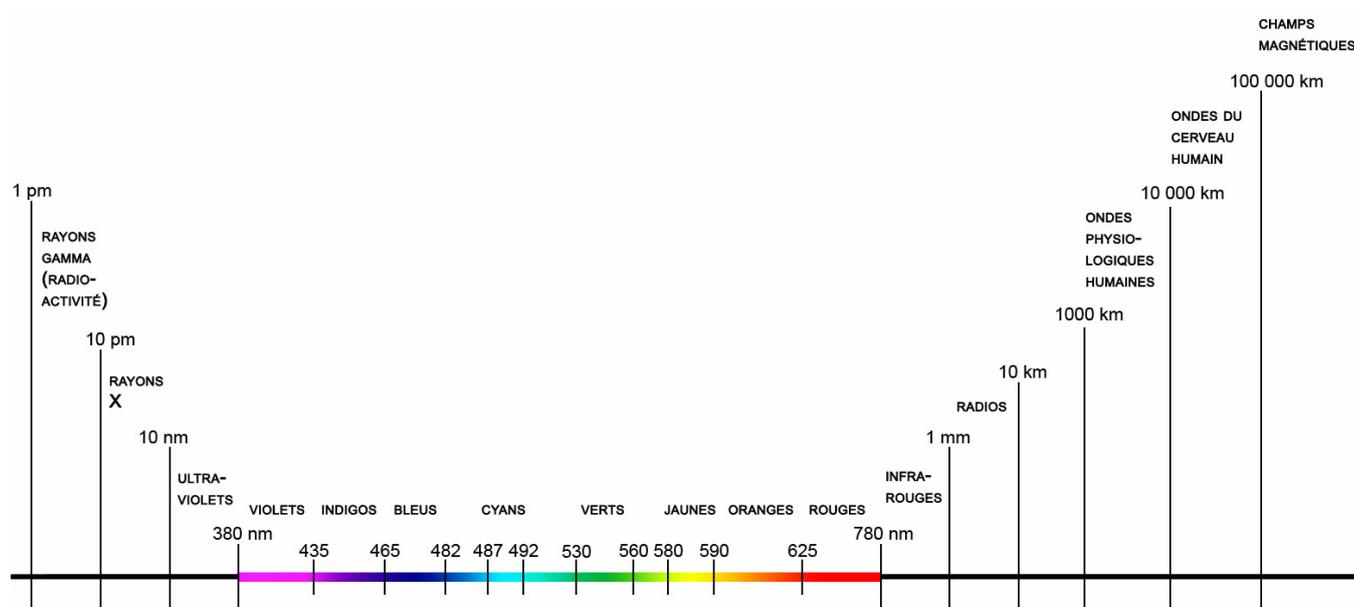
***DIURNISME* (PHILIPPE RAHM, CENTRE POMPIDOU, 2007)**

Qu'est-ce qu'une émotion ; qu'est-ce que l'émotion ? Plutôt que de commencer par une typologie des « émotions basiques », ou par une approche cherchant à isoler phénoménologiquement des émotions distinctes, empruntons un autre biais : celui de la construction émotionnelle objectivée. Le paragraphe suivant va se pencher sur une œuvre en particulier, choisie pour sa dimension emblématique par rapport au sujet abordé dans le présent mémoire. Elle servira diversement d'appui pour toute cette partie consacrée à l'architecture de l'émotion.

Pour l'exposition « Airs de Paris » au Centre Pompidou en 2007, l'artiste architecte Philippe Rahm avait proposé une pièce intitulée *Diurnisme*. Les spectateurs étaient invités à pénétrer dans une salle, après avoir suivi un couloir construit par les frères Bouroullec, *Les Tuiles* (2006), soit un treillis compact en forme d'assemblage ordonné d'éléments mobiles en tissu et mousse thermo-compressée — une matière rappelant le feutre, avec sa dimension chaleureuse et calme, via Joseph Beuys d'une part, via ses propriétés d'isolant thermique et phonique d'autre part. À l'issue de ce fragile couloir se trouvait une ouverture donnant sur un espace clos, baigné d'une intense lumière jaune-orangé et d'un son difficilement identifiable. Deux banquettes noires dont les dimensions et les formes épurées rappelaient deux lits (deux matelas noirs sur socles blancs) étaient disposées dans la pièce. Au fond de l'espace, un large panneau métallique noir (paroi froide, en métal noir), censé réguler à la baisse la température ambiante. À moins d'une peinture spéciale, on pouvait supposer que les autres murs étaient blancs car ils semblaient renvoyer cette coloration jaune-orangé. La lumière était produite par une centaine de néons, ordonnés en quatre rangées aussi parallèles que serrées, situés au

plafond. *Diurnisme*, parce que comme l'exposait Rahm, il s'agissait de produire une nuit artificielle.

L'œil humain perçoit généralement sous forme de couleurs les radiations électromagnétiques dont les longueurs d'onde sont comprises entre 380 et 780 nanomètres. De façon très schématique et approximative, le spectre électromagnétique peut être présenté ainsi (en longueurs d'onde) :



Et comme le notait Rahm à propos de *Diurnisme* :

«Après le “Noctambulisme”, je voudrais inventer le “Diurnisme”, en utilisant une lumière vive jaune-orange dont les longueurs d’onde, supérieures à 570 nanomètres, sont perçues par le corps à travers le rythme circadien de la mélatonine comme étant une vraie nuit. La pièce devient un paradoxe entre le visible et l’invisible, le visuel et le physiologique : une nuit qui ressemble à un jour lumineux. Les *Nocturnes pour piano* de John Field, musicien irlandais du début du 19^e siècle et inventeur de cette “forme”, sont diffusées à l’envers dans l’espace³¹. »

Lors de l’exposition «Airs de Paris», peu nombreux étaient les visiteurs à s’attarder dans cette pièce. L’impression de chaleur dégagée par les néons, bien que très subtile, la légère odeur de

³¹ «After the “Noctambulisme” I would like to invent the “Diurnism”, using an orange-yellow bright light which wavelengths, upper than 570 nanometers, are perceived by the body through the melatonin rhythm as a true night. The room becomes a paradox between the visible and the invisible, the visual and the physiological: a night which looks like a bright day. Reversed “nocturnes for the piano” by the early 19 century Irish musician John Field, the inventor of this “form” of music, are diffused in the space.», Philippe Rahm, «Diurnism». <http://www.philipperahm.com/data/projects/diurnisme/index.html>. (Ma traduction.)

matériaux ou de fibres grillés par cette même chaleur (cf. l'expression commune «ça sent le chaud»), pouvait générer une sorte de malaise, de sensation oppressante. Mais c'est avant tout la lumière qui instaurait un climat de tension, de pression. À l'instar des oreilles qui se bouchent au passage de certains tunnels, le corps pouvait ici ressentir comme un resserrement de son champ perceptif, comme une assourdissante oblitération — peut-être du fait de la saturation, de cette concentration forcée sur un spectre en réduction. La sensation de nuit n'était pas au rendez-vous en tant que telle, du moins pas d'après la plupart de nos habitudes acquises quant à la nuit et son panel sensitif. L'ampleur de la nuit, en tant que moment où obscurité et silence sont installés pour mieux permettre un dégagement de la captation sensorielle, une ouverture à l'infime et aux détails, devenait dans cette pièce une obsession jaune-orangé, capable même de faire rétrospectivement oublier la présence sonore. Le décalage entre l'énoncé («vraie nuit») et le perçu (dimension attirante mais aussi agressive) permettait de mesurer l'écart entre ce que l'on sait et ce que l'on ressent. Si l'on se prenait au jeu pour mieux se savoir plongé dans une «vraie nuit», on n'en ressentait pas moins une intense stimulation, quand bien même allongé, les yeux clos. Soit quelque chose qui en fin de compte n'incitait pas particulièrement ni à s'assoupir, ni à dériver.

Mais peut-être faut-il développer un peu ce qu'est la mélatonine afin de «clarifier» *Diurnisme*. Il s'agit d'une hormone connue pour son rôle dans le cycle circadien, c'est-à-dire pour son rôle dans l'endormissement ainsi que pour son influence sur les modifications comportementales saisonnières (d'humeurs). La production de mélatonine est plus élevée dans l'obscurité et / ou durant le sommeil. Ces dernières années, des recherches ont apparemment été menées afin d'évaluer l'impact de la lumière artificielle sur le taux de mélatonine (et sur le sommeil) des personnes y étant exposées. À l'exemple de cet extrait d'une étude publiée dans un journal d'endocrinologie (en 2011) :

«Contexte : des millions d'individus s'exposent habituellement à des lumières artificielles [/ éclairage d'intérieur] dans les heures précédant le coucher, cependant les effets de ce comportement sur l'activité de la mélatonine sont encore mal reconnus. Objectif : nous avons testé l'hypothèse selon laquelle l'exposition à la lumière artificielle en fin de soirée supprime le déclenchement de la synthèse de la mélatonine et raccourcit la durée de production de la mélatonine. [...] Conclusions : ces résultats indiquent que la lumière artificielle exerce un effet suppressif intense sur les niveaux de mélatonine et raccourcit la représentation corporelle interne de la durée de la nuit. En conséquence de quoi, le fait de s'exposer

régulièrement à la lumière électrique en fin de soirée perturbe l'activité de la mélatonine et pourrait ainsi potentiellement avoir un impact sur le sommeil, la thermorégulation, la pression sanguine et la stabilité des niveaux de glucose³². »

Et cet autre exemple, tiré d'une étude parue dans le même journal (en 2005) :

« Les baisses de production de mélatonine chez l'humain et chez les animaux sont connues pour être causées par les éclairages ambiants, particulièrement les éclairages à courtes longueurs d'onde (entre 470 et 525 nm). Nous avons travaillé sur une nouvelle hypothèse selon laquelle l'utilisation de lunettes opérant l'exclusion sélective de toutes les longueurs d'onde inférieures à 530 nm pourrait empêcher la suppression de la mélatonine dans des conditions de fort éclairage lors d'une expérience simulée de travail posté. Les niveaux de mélatonine dans la salive ont été mesurés par rapport à une lumière faible (< 5 lux), forte (800 lux) et filtrée (800 lux), toutes les heures entre 20:00 heures et 08:00 heures chez 11 jeunes hommes adultes en bonne santé et huit femmes (âge moyen : 24,7 ans ± 4,6 ans). Les mesures ont été faites durant trois nuits non consécutives, sur une période de 2 semaines. La qualité subjective du sommeil a été mesurée sur la base de l'auto-évaluation, tandis que les performances objectives ont été évaluées avec le Test de Performance Continue. Avec la lumière filtrée, tous les sujets ont présenté des niveaux de mélatonine préservés, équivalents à leur profil de sécrétion en présence de lumière faible. La lumière forte non-filtrée a drastiquement supprimé la production de mélatonine. La normalisation de la production endogène de mélatonine lors du port des lunettes n'a été préjudiciable ni à la performance, ni à la qualité subjective du sommeil, ni à la vigilance³³. »

Et comme on peut le supposer à la lueur de ces travaux, ce qui importait dans *Diurnisme* c'était le fait de pouvoir expérimenter une telle lumière *in concreto*, non par le port de lunettes, mais par la confrontation directe avec des sources lumineuses filtrées. L'inconfort ressenti dans le

³² « Context: Millions of individuals habitually expose themselves to room light in the hours before bedtime, yet the effects of this behavior on melatonin signaling are not well recognized. Objective: We tested the hypothesis that exposure to room light in the late evening suppresses the onset of melatonin synthesis and shortens the duration of melatonin production. [...] Conclusions: These findings indicate that room light exerts a profound suppressive effect on melatonin levels and shortens the body's internal representation of night duration. Hence, chronically exposing oneself to electrical lighting in the late evening disrupts melatonin signaling and could therefore potentially impact sleep, thermoregulation, blood pressure, and glucose homeostasis. » , Joshua J. Gooley, *et al.*, « Exposure to Room Light before Bedtime Suppresses Melatonin Onset and Shortens Melatonin Duration in Humans », dans *The Journal of Clinical Endocrinology & Metabolism*, n°96(3), mars 2011, pp. 463-472. (Ma traduction.)

³³ « Decreases in melatonin production in human and animals are known to be caused by environmental lighting, especially short-wavelength lighting (between 470 and 525 nm). We investigated the novel hypothesis that the use of goggles with selective exclusion of all wavelengths less than 530 nm could prevent the suppression of melatonin in bright-light conditions during a simulated shift-work experiment. Salivary melatonin levels were measured under dim (<5 lux), bright (800 lux), and filtered (800 lux) light at hourly intervals between 2000 and 0800 h in 11 healthy young males and eight females (mean age, 24.7 ± 4.6 yr). The measurements were performed during three nonconsecutive nights over a 2-wk period. Subjective sleepiness was measured by self-report scales, whereas objective performance was assessed with the Continuous Performance Test. All subjects demonstrated preserved melatonin levels in filtered light similar to their dim-light secretion profile. Unfiltered bright light drastically suppressed melatonin production. Normalization of endogenous melatonin production while wearing goggles did not impair measures of performance, subjective sleepiness, or alertness. », Leonid Kayumov, *et al.*, « Blocking Low-Wavelength Light Prevents Nocturnal Melatonin Suppression with No Adverse Effect on Performance during Simulated Shift Work » , dans *The Journal of Clinical Endocrinology & Metabolism*, n°90(5), mai 2005, pp. 2755-2761. (Ma traduction.) Par ailleurs, Philippe Rahm et son équipe se sont appuyés sur une autre étude, non reproduite ici : A. Sasseville, *et al.*, « The use of orange lens glasses to block the suppression of melatonin by bright light » (2003), dont le résumé est disponible en ligne : <http://www.sltbr.org/vm/abstracts3.pdf>. Cf. Christine Macel & Valérie Guillaume (directrices de publication), *Airs de Paris*, catalogue de l'exposition éponyme présentée au Centre Pompidou, Galerie 1, du 25 avril au 16 août 2007, éd. Centre Pompidou, Paris, 2007, p. 308.

lieu aurait pu être perçu comme offrant un plus grand bien à l'usage, par rapport à une bonne régulation de l'organisme. Les résultats n'étaient pas perceptibles immédiatement mais pouvaient potentiellement le devenir sous certaines conditions d'utilisation. Et d'un inconfort prolongé aurait pu naître un confort durable, voire même une plus-value émotionnelle de part un meilleur équilibre hormonal. Chacun de nous a déjà eu l'occasion de mesurer l'impact que peut avoir le sommeil sur l'approche émotionnelle que l'on se fait du quotidien éveillé — un sommeil dégradé peut rendre irritable, colérique, à fleur de peau, prêt à pleurer à la moindre contrariété ou encore prompt à rire ou à rougir pour des choses banales. Le sommeil, à l'instar de nombreux autres facteurs, influe sur nos rapports émotionnels aux contextes.

Notons que cette façon de lire le sommeil comme étant une «bonne» chose, et de comprendre le cycle circadien comme devant être régulier pour être optimal relève peut-être d'une lecture culturellement orientée. Dans les bonus du DVD du film *Atanarjuat – La légende de l'homme rapide* (2001), de Zacharias Kunuk, la question des conditions du tournage, au Nunavut, donnait lieu à une description de la journée polaire (ou soleil de minuit) comme ayant été un adjuvant euphorisant et enivrant, avec cette possibilité offerte de pouvoir travailler jusqu'à total épuisement sans être arrêté par la nuit et le manque de lumière. La baisse du taux de mélatonine ne semblait pas alors avoir été vécue de façon négative. Dans ce cas, le biais culturel tient à ce que les conditions de tournage d'un tel film ne soient pas assimilables à la simulation d'une expérience de travail en 3x8. Par contre il est vrai que la lumière artificielle courante ne peut objectivement pas être comparée à la lumière émise par le soleil. En soi l'expérience de *Diurnisme* n'était pas extrêmement désagréable : elle était avant tout inhabituelle et intrigante — en raison aussi de la quantité de néons requise pour produire une intensité lumineuse suffisante tout en restant dans ces longueurs d'onde ; quantité induisant un léger réchauffement en dépit du fait que les néons soient faiblement calorifiques. On notera au passage qu'en ces termes, la si typique lumière orangée des villes et autoroutes françaises pourrait bien être de ces éclairages qui (involontairement) respectent le cycle circadien de la mélatonine. Dans le catalogue de l'exposition, le conservateur du patrimoine et curateur de l'exposition, Valérie Guillaume, citait ainsi les propos de Rahm : «Notre projet [...] fait écho à la révolution à la fois technique et sociale que fut l'introduction de l'éclairage, au XIX^e siècle,

dans l'espace public. », avant d'ajouter lui-même :

« Depuis lors, le développement des éclairages urbains, routiers et publicitaires est tel qu'il nuit même à l'obscurité attendue de la nuit. Les éclairages combinés à la pollution atmosphérique provoquent l'apparition d'un halo jaune orangé qui efface la vision des étoiles. Dans la région parisienne, ce halo reste visible à 150 kilomètres de distance. En 1992, la Déclaration des droits pour les générations futures, émise par l'Unesco, a ainsi consacré un volet spécifique au droit à la conservation du ciel dans toute sa pureté. Le ciel nocturne est désormais classé "patrimoine mondial de l'humanité" à protéger³⁴. »

En un sens, l'expérience proposée par Rahm pouvait évoquer une nouvelle de science-fiction écrite par Alfred Van Vogt en 1950, le « Village enchanté » (*Enchanted Village*). Il y est raconté l'histoire d'un explorateur naufragé, seul survivant de son équipe suite à un atterrissage raté sur Mars. Après avoir erré quelque temps dans ce désert jaune-orangé, à la recherche de quelque chose plutôt que rien, l'homme se retrouve dans un village abandonné, aux architectures étranges mais parfaitement conservées, d'apparence habitable. Progressivement il y découvre des éléments à l'aide desquels il se met à espérer pouvoir survivre au tarissement de ses propres ressources. D'abord le fruit d'un arbre, qu'il goûte dans un geste assez suicidaire mais qu'il recrache face au malaise et à la douleur procurés par celui-là. Et toute la nouvelle consiste à narrer la lutte entre le village d'apparence hostile et létale, et le naufragé prêt à tout pour survivre : une agonie au sens littéral (du grec *αγωνία*, lutte, concours). Petit à petit le village, qui paraît être complètement automatisé — fourniture de nourriture (toxique) à la demande ; chauffage des lits (brûlant) à la nuit tombée ; pulvérisation de liquide (irritant) dans l'équivalent de cabines de douche ; diffusion de sonorités (insupportables) en guise de musique d'ambiance, etc. — semble entrer en résonance avec le cosmonaute pour tenter de s'accorder à ses désirs. Et pour sa part, le cosmonaute essaie d'apprendre au village à moduler ses offrandes en fonction de ce que son corps peut supporter. Progressivement les agressions se métamorphosent en aide. Jusqu'à ce que le naufragé se réveille un matin pour mieux trouver la nourriture savoureuse, les lits à température parfaite, la douche agréable et la musique ô combien harmonieuse. Observant alors sa langue fourchue et sa longue queue, il se réjouit de ce que le village se soit enfin adapté à ses besoins. On l'aura compris, c'est bien lui qui se sera métamorphosé en quelque chose comme un lézard géant ou un dragon de Comodo, et non le

³⁴ Pour la référence complète, cf. la note précédente. *Ibid.*, p. 308.

village, vivant, qui se sera mis à son niveau. Ou plutôt, le village l'aura transformé et assimilé pour son plus grand bien. C'est cet aspect métamorphique du « Village enchanté » qui ressort de *Diurnisme* : on se demande au bout de combien de temps un tel environnement pourrait finir par nous devenir viable et plaisant. On se demande même s'il est bien raisonnable de s'adapter à tout.

Comme dit précédemment, l'entrée en matière dans la question des émotions se sera faite par le biais de la construction émotionnelle objectivée. Une œuvre d'art, *Diurnisme*, aura été choisie afin de guider le déroulement de cette enquête sur l'émotion, et comme là encore le thème est vaste, sur le thème spécifique de la construction émotionnelle plutôt que sur l'ontologie des émotions. Peut-être est-il plus couru de s'attaquer à celles-ci sous l'angle des grands sentiments romantiques, mais cette approche aura volontairement été laissée de côté dans la mesure où les travaux les plus récents en ce domaine se sont largement détachés de la conception romantique. Il ne s'agit pas de restaurer le sentimentalisme — dont Adam Smith nous a livré une approche, par exemple, dans *La Théorie des Sentiments Moraux* (1759-1790) [cf. Smith, 2006] —, mais bien plus de s'intéresser aux représentations contemporaines de l'émotion, plus particulièrement lorsqu'il s'agit de l'objectiver. Aussi, ce paragraphe aura consisté en une description subjective de la pièce de Rahm, ouvrant sur des observations physiologiques appartenant en un sens à la « mécanique émotionnelle », avant de nous ramener vers une approche science-fictionnelle de l'architecture émotionnelle. Cela dit, même si l'approche contemporaine reste le principal point de focal, le paragraphe suivant va néanmoins proposer une incursion dans la modernité artistique du XX^e siècle. S'il ne s'agit que d'un biais (un autre aurait pu être choisi), il servira pourtant à arrimer la notion d'architecture de l'émotion dans un contexte, dans des histoires européennes, pour ne pas la laisser flotter dans une solitude présentiste.

LES FORMES DE L'ÉMOTION

Les tentatives d'objectivation de l'émotion ne sont pas des préoccupations nouvelles. Le XX^e siècle aura lui aussi contribué à cet effort et nous pouvons essayer ici d'en démêler

quelques aspects. Ceux-ci ont été choisis pour leurs fortes dimensions conditionnantes sur la perception que nous pouvons avoir aujourd'hui des émotions et de leur pouvoir. Cette décision est sûrement arbitraire puisqu'elle part du principe que les désastres du siècle passé auront durablement frappé les esprits des Européens, générant une fracture (rationnelle «et» émotionnelle) n'ayant rien laissé intact et ne pouvant être oubliée. À l'heure où ces lignes sont écrites, et face à la remontée des extrêmes en Europe³⁵, il peut être «tentant» de croire que l'impact n'a pas été, après tout, si conséquent, si mémorable. Raison de plus, probablement, pour essayer d'en distinguer quelques aspects.

Le fait que les architectes se soucient de l'influence sensitive (et par répercussion, affective) de leur travail sur leurs utilisateurs (ou usagers), ne découle pas des récentes théories des émotions. Vitruve notait déjà au premier siècle avant notre ère :

«La convenance qu'exige la nature des lieux, consiste à choisir les endroits où l'air et les eaux sont plus sains, pour y placer les temples, principalement ceux qu'on bâtit à Esculape, à la Santé, et aux autres divinités auxquelles on attribue la vertu de guérir les maladies. Car les malades qui passeront d'un air malsain dans un autre qui sera plus salubre, et qui feront usage de meilleures eaux, se guériront aisément ; ce qui augmentera beaucoup la dévotion du peuple, qui attribuera à ces divinités la guérison qu'il devra à la nature d'un lieu salubre.

Il est encore une convenance que demande la nature des lieux. C'est de faire en sorte que les fenêtres des chambres à coucher et des bibliothèques soient placées au levant [...] et que les cabinets de tableaux et autres curiosités qui demandent un jour égal en tout temps, aient les leurs vers le septentrion, d'autant que ce qui est tourné vers ce côté du ciel est toujours également éclairé.» [Vitruve, 1816, pp. 13-14]

Ce souci, quant aux choix et aux orientations des lieux en fonction de leurs influences physiologiques et psychoactives sur le moral et les émotions (notamment la peur) voire les croyances, se rencontre tout autant dans l'ouvrage *L'art de la guerre*, communément attribué à Sun Tzu — et datant probablement du V^e siècle avant notre ère. Avec l'architecture vitruvienne on *sent* que c'est bon et, éventuellement, on *sait* (ou l'on se fait une idée de) pourquoi ça l'est ; tandis qu'avec *Diurnisme*, on *sait* que cela pourrait être bon mais on *sent* que ça ne l'est pas, que l'on n'est pas «naturellement» enclin à apprécier ce type de conditionnements physiologiques. Ou plutôt, la contradiction s'enrichit d'une dimension supplémentaire : a) le spectateur est rationnellement et logiquement informé du fait que cet

³⁵ Cf. Paolo Virno et Alessia Ricciardi, «Theses on the New European Fascism» , dans *Grey Room*, n°21, éd. The MIT press, automne 2005, pp. 21-25.

espace soit potentiellement bon pour lui, b) mais il ressent éventuellement un malaise à fréquenter le lieu, c) cela dit si des prélèvements salivaires étaient effectués sur les visiteurs, suite à une fréquentation prolongée du lieu, il serait peut-être possible de démontrer que le confort du visiteur a été amélioré par l'expérience — du moins est-ce ce qui est objectivement escompté (en vertu des relations de causes à effet, et abstraction faite de la sensation de chaleur). En un sens, la subjectivité se trouve ici prise entre deux objectivités démontrables (l'esprit logique et le corps quantifiable) : la chose est scientifiquement bonne ; des prélèvements dans les règles pourraient prouver que la chose est physiologiquement bonne ; mais le ressenti subjectif, que ce soit à court terme ou après quelque escale prolongée, semble à même de contredire ces deux positivités. Le lieu semble pouvoir être agréable, mais pour une physiologie adaptée — une physiologie qui aurait déjà, si possible, opéré un travail de conversion esthétique transformant le désagrément en plaisir (et le mammifère en reptile). En un sens la confrontation avec *Diurnisme* dépouille le spectateur de sa subjectivité. Si elle est renforcée par cet oxymorique «paradoxe entre le visible et l'invisible, le visuel et le physiologique» (sur le modèle de «cette obscure clarté»), en ce que les sensations conscientes contredisent le discours scientifique, elle n'en est pas moins questionnée quant à ses manières de lier perceptions et plaisirs.

Plutôt qu'un dépouillement, la subjectivité devient cette liaison élaborée entre perceptions et plaisirs (et douleurs). Soit une question très sérieuse puisqu'elle se retrouve dans les théories de l'évolution, comme le note Daniel Dennett :

«La substance phénylthiourée, nous dit-il [Jonathan Bennett], a un goût amer pour un quart de la population humaine et est absolument insipide pour le reste. Le goût qu'elle a pour vous est génétiquement déterminé. Est-ce que la phénylthiourée est amère ou insipide? [...] Les catégories fondamentales de nos espaces colorimétriques (et évidemment de nos espaces olfactifs, de nos espaces acoustiques et tout le reste) sont façonnées par les pressions sélectives, de telle manière qu'en général cela fasse sens de demander à quoi sert telle ou telle discrimination ou préférence particulière. Il y a des raisons au fait que nous évitions l'odeur de certaines choses et en recherchions d'autres, au fait que nous préférons certaines couleurs à d'autres, au fait que certains sons nous dérangent plus que d'autres ou nous adoucissent davantage. [...] Par exemple, si à l'avenir manger des tripes *devait* impliquer une mort préreproductive, seuls ceux d'entre nous ayant été "naturellement" (et *jusque-là* inutilement) disposés contre le fait de manger des tripes auraient un avantage (peut-être mince au début, mais bientôt de façon explosive, si les conditions le favorisaient). Il n'en découle pas pour autant que si vous trouvez quelque chose (par exemple le brocoli) indescriptionnellement et ineffablement abject, c'est qu'il y a une raison pour cela. Ni que vous soyez déficient si vous êtes en désaccord avec vos pairs à ce sujet. Il se pourrait juste que ce soit l'une des propensions innées de votre espace qualitatif qui n'a, jusque-là, pas la moindre signification fonctionnelle. (Et pour

votre bien, vous devriez espérer que si jamais un jour cela devait avoir une signification, que ce soit alors parce que soudainement le brocoli s'est révélé être néfaste pour nous.) Ces considérations évolutionnistes suffisent afin d'expliquer pourquoi les qualités secondaires se sont avérées être "ineffables" et si résistantes à la définition³⁶.» [Dennett, 1991, «Chapitre 12 - Les qualia disqualifiées», «2. Pourquoi y a-t-il des couleurs?», pp. 379-382].

Face à ces mots, on peut se sentir comme multiplement frustré par l'oscillation entre le déterminisme des goûts et des plaisirs par la génétique, et le hasard de tel ou tel «branchement» d'une qualité (seconde) avec une sensation de plaisir ou de déplaisir. Comme l'écrivait Dennett il y a vingt-et-un ans (1991) : espérons donc que les qualités secondes (perception des couleurs, des sons, des arômes, etc.) ne deviennent pas l'objet d'études trop dogmatiques. Ou plutôt la pièce de Rahm interroge-t-elle justement cette anfractuosité : nos organismes sont génétiquement programmés pour réagir de façon précise à cette couleur jaune-orangé (sensibilité électromagnétique de l'œil), mais la génétique ne permet pas encore d'expliquer complètement ni pourquoi ni comment quelque chose de bon pour un organisme peut néanmoins lui procurer une sensation de déplaisir (branchements hasardeux, complexité des paramètres...) : l'apprentissage culturel en matière de plaisirs et déplaisirs vient brouiller les cartes. Et l'on peut alors trouver à *Diurnisme* une dimension critique assez forte : les Lumières (*Aufklärung*) de la pièce se révèlent être une sorte de nuit perpétuelle. De façon un peu provocante, on pourrait avoir envie de dire que si le titre *Dialektik der Aufklärung* (par Theodor Adorno et Max Horkheimer, en 1947) n'était pas déjà «pris», on serait presque tenté d'en faire le sous-titre de l'œuvre.

Mais le fait de savoir si cette proposition de Philippe Rahm relève d'un test d'exploitation

³⁶ «The substance phenol-thio-urea, he [Jonathan Bennett] tells us, tastes bitter to one-quarter of the human population and is utterly tasteless to the rest. Which way it tastes to you is genetically determined. Is phenol-thio-urea bitter or tasteless? [...] The basic categories of our color spaces (and of course our odor spaces and sound spaces, and all the rest) are shaped by selection pressures, so that in general it makes sense to ask what a particular discrimination or preference is for. There are reasons why we shun the odors of certain things and seek out others, why we prefer certain colors to others, why some sounds bother us more, or soothe us more. [...] For example, if eating tripe were to spell prereproductive doom in the future, only those of us who were "naturally" (and *heretofore* pointlessly) disposed against eating tripe would have an advantage (perhaps slight to begin with, but soon to be explosive, if conditions favored it). So it doesn't follow that if you find something (e.g., broccoli) indescribably and ineffably awful, there is a reason for this. Nor does it follow that you are defective if you disagree with your peers about this. It may just be one of the innate bulges in your quality space that has, as of yet, no functional significance at all. (And for your sake, you had better hope that if it ever does have significance, it is because broccoli has suddenly turned out to be bad for us.) These evolutionary considerations go a long way to explaining why secondary qualities turn out to be so "ineffable," so resistant to definition.» [Dennett, 1991, «Chapter 12 - Qualia disqualified», «2. Why are there colors?», pp. 379-382]. (Ma traduction, aidée de la traduction de Pascal Engel, *La conscience expliquée*, éd. Odile Jacob, Paris, 2001, «Chapitre 11 - Les qualia disqualifiées», «2. Pourquoi y a-t-il des couleurs?», pp. 470-473.

pratique de résultats scientifiques (plus ou moins satisfaisant, en raison notamment du gouffre énergétique qu'il représente, d'autant plus s'il s'agissait de climatiser efficacement, *même si* les tubes fluorescents ont un meilleur rendement énergétique avec une moindre déperdition calorifique) ou relève plutôt d'une provocation à la perpendiculaire de Vitruve, n'est pas vraiment ce qui importe pour le présent propos — l'ambiguïté est même plaisante. Par contre nous pouvons noter la récurrence de cette attention au lien entre architecture (ici sous forme de spatialisation lumineuse colorée) et émotions (influences physiologiques). Soit une attention aux modulateurs affectifs du rapport à un espace-temps ; une attention à l'économie émotionnelle. Ce point ne manque pas de rappeler les « palais de la mémoire », les temples et les théâtres allégoriques précités, notamment conçus pour solliciter les sens et éveiller des émotions assez précises [cf. Yates, 1999 ; Carruthers, 2002] — à l'instar du romantique exemple des cathédrales sylvestres³⁷. Mais cela peut encore évoquer le manifeste du visionnaire allemand Paul Scheerbart, *L'architecture de verre* (1914), dans lequel le poète rêvait de voir se transformer la nuit en jour perpétuel :

« L'aérostation aura sans nul doute à cœur de vaincre aussi la nuit. [...] La nuit peut être semblable au jour. [...] De ces nuits de lumière, que doit nous offrir l'architecture de verre, que dire d'autre sinon qu'elles sont vraiment "extraordinaires". Imaginez des projecteurs sur toutes les tours en verre et dans tous les avions, imaginez les trains multicolores. Et puis les usines, où la lumière brillera aussi la nuit grâce à des disques de couleur. Imaginez encore les grands palais et les cathédrales de verre, et les villas, en verre elles aussi — et des espaces urbains — sur la terre ferme et sur l'eau — qui se déplaceront souvent — et de l'eau, toujours renouvelée, de couleur toujours différente. Sur Vénus et sur Mars, on écarquillera les yeux, car la terre sera méconnaissable. Peut-être les hommes vivront-ils davantage la nuit que le jour. Et les astronomes bâtiront leurs observatoires au fond d'un paisible ravin ou sur une colline, car cette gigantesque mer de lumière pourrait gêner l'observation du ciel. L'idée n'est pas nouvelle — c'était déjà celle des grands bâtisseurs du gothique : ne l'oublions pas. » [Scheerbart, 2006, p. 35, p. 38, p. 58]

Et il rêvait tout autant de voir la planète Terre se métamorphoser en un spectacle pyrotechnique pour ainsi dire permanent [Scheerbart, 2006, p. 60], à l'instar d'un vaisseau lumineux circulant dans l'espace — métaphore reprise pour mieux être déroutée, par l'architecte et membre du Black Mountain College (mouvement en filiation avec le Bauhaus de Walter Gropius à Chicago), Richard Buckminster Fuller, dans son recueil de poèmes écologiques *Intuition* (1970). De surcroît Scheerbart accordait une attention particulière aux lumières colorées ainsi qu'à leurs effets (supposés ou non) sur l'organisme :

³⁷ Cf. Hegel, 1997, t. II, pp. 88-102 (ou 3^{ème} partie, 1^{ère} section, chap. III, 2 « Formes architectoniques particulières »).

« Les effets psychologiques engendrés par l'architecture de verre. Les prêtres babyloniens et assyriens connaissaient déjà l'étrange influence que la lumière provenant du verre coloré exerce ; ils furent les premiers à allumer dans les temples des lampes en verre de couleur. Et c'est par Byzance que ces lampes furent ensuite introduites dans les églises d'Europe. Les vitraux gothiques en sont les héritiers ; rien d'étonnant à ce qu'ils produisent une impression particulièrement solennelle. Mais l'impression de solennité que procure le verre de couleur est aussi propre à l'architecture de verre ; son influence sur la psyché humaine ne peut donc être que bénéfique, identique à celle qu'exerçaient les cathédrales gothiques et les lampes de verres babyloniens. L'architecture de verre, qui transforme les maisons des hommes en cathédrales, aura un effet semblable. » [Scheerbart, 2006, p. 69]

Soit des propos qui auront inspiré le travail de son ami architecte Bruno Taut, connu pour son expressionnisme mais également pour avoir été en contact avec le mouvement Bauhaus. En un sens, l'idée d'une architecture de l'émotion n'aura pas non plus été étrangère au mouvement de la *Gläserne Kette* (mené par Bruno Taut et dont Gropius fut membre silencieux [cf. Stavrinaki, 2009]) et l'un des architectes du groupe, Paul Goesch, en aura même proposé une ébauche de définition par « motifs », vers 1920 :

« 1) Architecture d'émotion : impressions de stalactites, de champ de blé sous un vent tourbillonnant, imitations de coquilles d'escargot, de crânes, plan inspiré par une coupe de nautilaire avec ses différentes chambres.

2) Architecture abstraite fondée sur les mathématiques : ornementation arabe, trilobe gothique, symphonie de chiffres et de formes, couleurs, séparations, regroupements, transitions, murs colorés s'interrompant les uns les autres, parfois sur un fond doré, le tout richement décoré selon un plan d'ornementation, murs agrémentés en suivant les principes de la sculpture.

Mélanger ces deux voies de façon à introduire l'écho des émotions dans la seconde ou suivre l'idée conductrice de la première voie en veillant à ce qu'elle se réalise sur le mode rythmé de la seconde. » [Stavrinaki, 2009, p.162]

Soit une division pouvant peut-être s'expliquer par le rapport des motifs à la symétrie : l'asymétrie pour les « émotions » ; la symétrie pour « l'abstraction mathématique ». Dans son livre (non traduit en français) *Crystals, Fabrics, and Fields: Metaphors of Organicism in Twentieth-Century Developmental Biology* (1976), l'historienne des sciences et de la conscience Donna Haraway³⁸ traite de l'influence esthétique de certains paradigmes formels sur la biologie organiciste du début du XX^e siècle. Dans cet ouvrage, elle relève à propos du vitalisme :

« Les problèmes de forme se répartissent naturellement en considérations à propos de la symétrie, de la polarité et des motifs [*pattern*]. La compréhension de ces trois éléments est un point critique pour ce qui est de l'appréciation des paradigmes non-mécanicistes dans l'histoire de la biologie. La signification de ces composantes a été radicalement transformée dans l'articulation de l'organicisme moderne. Harrison,

³⁸ Et notamment membre du comité consultatif honoraire de la 13^{ème} Documenta de Kassel (2012). [http://d13.documenta.de/#participants/participants?tx_participants_pi1\[showUId\]=9](http://d13.documenta.de/#participants/participants?tx_participants_pi1[showUId]=9).

Needham et Weiss croyaient chacun que les relations de symétrie des organismes reposaient sur une asymétrie fondamentale au niveau du noyau moléculaire. L'asymétrie dans les organismes aura historiquement été la pierre angulaire des théories du vitalisme. La nature inorganique était largement ressentie comme étant la province des relations symétriques régulières ; le monde organique était conçu comme étant le royaume des formes spatiales extraordinaires, telles que la spirale logarithmique des coquilles de nautilus et la quintuple symétrie radiale de l'étoile de mer. Aucune de ces enveloppes formelles [*shape*] ne pouvant être réduite à la simple géométrie des réseaux spatiaux cristallins. Les cristaux, manifestant des enveloppes formelles géométriques, étaient sûrement le produit des lois de la chimie et de la physique, mais la symétrie et l'asymétrie particulières des organismes semblaient faire appel à des principes d'organisation uniques³⁹. » [Haraway, 1976, pp. 48-49]

Et ces propos nous permettent de mieux entendre la répartition proposée par Goesch : le tourbillonnant asymétrique (ou le « symétriquement complexe ») du côté de l'émotion vitale et du procédural-processuel (plus biologique et « aristotélicien ») ; et les arabesques et autres symétries arithmétiques du côté d'une abstraction rationnelle moins incarnée, plus contemplative (ou plus « platonicienne »). Ainsi que la volonté de Goesch, dans une mouvance organiciste, de réunir ces deux « antagonistes ». Ce qui peut nous intéresser ici, c'est le fait que l'émotion soit classée du côté de la « vie », du tourbillonnant, du biologique, du passager — tandis que la rationalité abstraite (« mathématique ») se trouve classée du côté d'une mécanique géométrique plus « morte » (organisée, mais inorganique).

La « chaîne de verre » (*Gläserne Kette*), aura été traversée par l'organicisme du début du XXe. Et ce en tant que courant architectural et esthétique influencé par les structures cristallines notamment, à l'instar de Bruno Taut dont le pseudonyme au sein de cette circulation épistolaire du début des années 1920 était *Glas* [cf. Stavrinaki, 2009], c'est-à-dire « verre » ou « cristal » en allemand. Le problème de savoir comment des composés morts (H₂O, par exemple) peuvent se structurer de façon « harmonieuse » (en flocons de neige, par exemple) aura longtemps intrigué les poètes, les artistes et les scientifiques. Johannes Kepler a écrit en 1611 un très beau traité sur la question, sans pour autant parvenir à y répondre — comment

³⁹ « Problems of form grade naturally into a consideration of symmetry, polarity, and pattern. An understanding of these three elements is critical to an appreciation of nonmechanist paradigms throughout the history of biology. The significance of each component was transformed radically in the articulation of modern organicism. Harrison, Needham, and Weiss each believed that symmetry relations of organisms rested on a fundamental asymmetry at the molecular core. Asymmetry in organisms has historically been a cornerstone for theories of vitalism. Inorganic nature was widely felt to be the province of regular symmetry relations; the organic world was conceived as the realm of extraordinary spatial forms, such as the logarithmic spiral of Nautilus shells and the fivefold radial symmetry of starfish. Neither shape can be reduced to the simple geometry of crystalline space lattices. Crystals manifesting simple geometrical shapes surely were products of the laws of chemistry and physics, but the peculiar symmetries and asymmetries of organisms seemed to call for unique organizing principles. » [Haraway, 1976, pp. 48-49]. (Ma traduction.)

les cristaux de neige peuvent-ils se former, être dotés d'une faculté formatrice, toujours selon une quintuple symétrie radiale, sans que jamais ils ne soient identiques entre eux, et sans pour autant appartenir au vivant? Il y note avec beaucoup de grâce : « *O veram mortuam vitam sine philosophia!* »⁴⁰ [Kepler, 2010, p. 94], comme si la pensée immatérielle pouvait être ce qui anime la matière morte et mécanique mais cependant dotée de facultés formatrices. En un sens, on peut dire que les cristaux, la cristallisation, la cristallographie, auront imprimé leurs marques *borderline* dans les esprits : soit l'image d'une « frange entre » l'organique et l'inorganique. Haraway, en testant la méthode de Thomas Kuhn à propos de l'importance des paradigmes et des métaphores dans les constructions scientifiques, relève la prégnance de la métaphore cristalline dans l'organicisme⁴¹. Pour sa part, elle insiste sur l'influence de Samuel Taylor Coleridge et Johann Wolfgang von Goethe pour la formation de cet imaginaire scientifique fécond. Et nous pourrions à notre tour insister sur l'influence de Kepler pour la formation de ces imaginaires poétiques féconds (ceux de Coleridge et Goethe). Mais ce qui importe le plus, ici, c'est de noter la prégnance du cristal comme métaphore d'une structure à la croisée entre le mort et le vivant : représentant formel d'un « pouvoir organisateur », avec ou sans vie.

Haraway évoque brièvement l'artiste du Bauhaus Paul Klee, dans la mesure où elle choisit le très synthétique « Ich Kristall. » pour exergue de l'un de ses chapitres. Cette phrase, issue du journal de Klee, vient en quelque sorte clore une réflexion mélancolique au cours de laquelle l'artiste finit par se demander si le cristal que lui-même incarne ne pourrait pas néanmoins mourir :

« 951. On abandonne la région d'ici-bas pour aller construire de l'autre côté dans une région au-delà qui peut au moins exister intacte. / Abstraction. / Le froid romantisme de ce style sans pathos est inouï. Plus ce monde (d'aujourd'hui précisément) se fait épouvantable, plus l'art se fait abstrait, tandis qu'un monde heureux produit un art porté vers l'ici-bas. / Aujourd'hui est fait de la transition d'hier à maintenant. Dans la grande fosse des formes, gisent les ruines auxquelles on tient encore, en partie. Elles fournissent matière à l'abstraction. / Un chantier d'inauthentiques éléments pour la formation d'impurs cristaux. / Voilà où nous en sommes. / Mais ensuite : il arriva que saigna la druse. Je pensais en mourir, guerre et mort. Puis-je donc

⁴⁰ « Ô vraiment que la vie est morte sans la philosophie! » [Kepler, 2010, p. 94]. (Ma traduction, d'après la traduction anglaise.)

⁴¹ En tant que courant scientifique ayant tenté de résoudre le dualisme vitalisme – mécanisme par la proposition d'une troisième voie synthétique, orientée vers la compréhension de l'organisation et des structures du vivant, *sans* tenter de réduction à la physique et aux mathématiques seules, et *sans* postuler une force organisatrice immatérielle.

mourir, moi cristal? / Moi cristal⁴².»

Et l'importante question de savoir *comment* discerner entre vivant et mort (entre organique et inorganique), semble ainsi s'être culturellement tramée d'émotion, à l'instar de la division effectuée par Goesch : la part morte (d'aucuns diraient immortelle) de l'humain pour les mathématiques abstraites ; sa part vivante pour les émotions. L'émotion se chargeant ainsi, de façon symptomatique, de la connotation du signe de vie, de souffle énergétique et moteur, d'animation.

Notons que la tendance à l'objectivation de l'émotion dans et par l'architecture a de même été soutenue par le peintre et sculpteur mexicain Mathias Goeritz, architecte et concepteur du musée El Eco de Mexico (1952-1953). À l'occasion de l'inauguration du lieu, Goeritz proposa ainsi le «Manifeste de l'architecture émotionnelle» (*Manifiesto de la Arquitectura Emocional*, 1953), aujourd'hui consultable sur le site internet du musée, et commençant ainsi :

«Le nouveau "musée expérimental" El ECO à Mexico City a commencé ses activités, c'est-à-dire ses expérimentations, avec l'architecture de son propre bâtiment. Ce travail était fait pour être un exemple d'une architecture dont la fonction principale est d'évoquer l'ÉMOTION.

L'art en général, et naturellement l'architecture, reflète la condition spirituelle de l'homme à un moment particulier. Mais on peut avoir l'impression que l'architecture moderne, qui est si idiosyncrasique et intellectuelle, peut-être parce qu'elle a perdu tout contact rapproché avec sa communauté, a pendant une période exagéré ses efforts d'accentuation des aspects rationnels de l'architecture. Il en résulte que l'homme du vingtième siècle se sent submergé par autant de "fonctionnalisme", de logique et d'utilitarisme dans l'architecture moderne. Il cherche une échappatoire, mais ni l'esthétisme externe compris comme "formalisme", ni le régionalisme organique, ni le confucianisme dogmatique n'ont examiné en profondeur le problème du fait que l'homme de notre époque, qu'il soit créateur ou récipiendaire, aspire à quelque chose de plus qu'une maison qui soit attractive, agréable et appropriée à ses besoins. Cet homme attend, ou attendra un jour, de l'architecture et de ses moyens modernes une élévation spirituelle. Dit simplement, une émotion, telle que celle donnée en son temps à l'architecture de la pyramide, du temple grec, de la cathédrale romane ou gothique, ou même du palais baroque. *C'est seulement lorsque l'architecture lui fournit des émotions réelles que l'homme la considère à nouveau comme étant de l'art* [italique ajouté] [...]»⁴³.

⁴² «951. Man verläßt die diesseitige Gegend und baut dafür hinüber in eine jenseitige, die ganz ja sein darf. / Abstraktion. / Die kühle Romantik dieses Stils ohne Pathos ist unerhört. Je schreckensvoller diese Welt (wie gerade heute), desto abstrakter die Kunst, während eine glückliche Welt eine diesseitige Kunst hervorbringt. / Heute ist der gestrig-heutige Übergang. In der großen Formgrube liegen Trümmer, an denen man noch teilweise hängt. Sie liefern den Stoff zur Abstraktion. / Ein Bruchfeld von unechten Elementen, zur Bildung unreiner Kristalle. / So ist es Heute. / Aber dann : einst blutete die Drüse. Ich meinte zu sterben. Krieg und Tod. Kann ich denn sterben, ich Kristall? / Ich Kristall.», Paul Klee, *Tagebücher 1898-1918*, rassemblés par Felix Klee, éd. Dumont Schauberg, Cologne, 1957, p.323, n°951. (Traduit de l'allemand par Pierre Klossowski dans Paul Klee, *Journal*, éd. Grasset, 2004, Journal III, pp. 328-329.)

⁴³ «The new "experimental museum" El ECO in Mexico City began its activities, that is to say, its experiments, with the architecture of its own building. This work was understood to be an example of an architecture whose main function is to evoke EMOTION.

Si la volonté « d'émotionnaliser » l'architecture (comme on parle d'une volonté de la « rationaliser ») semble sous-entendre des émotions positives (avec à la clef une amélioration du bien-être), la chose ne va pas pour autant de soi. Goeritz désavoue le formalisme émotionnel, mais il semble néanmoins conférer aux émotions une dimension de liant collectif, d'ouverture et de circulation, une dimension religieuse positive. Dans son ouvrage de 2011, *Vers une architecture cathartique (1945-2001)*, Emmanuel Rubio s'intéresse à l'architecture brutaliste ainsi qu'à ses échos dans l'environnement contemporain, tout en insistant sur les variantes dans la tendance à la « bunkerisation ». Nous sommes alors davantage face à l'autre versant, celui de la souffrance, du repli ou encore des émotions négatives. Notamment avec l'influence exercée par les constructions de type blockhaus dans l'architecture post-Seconde Guerre mondiale.

RATIONALISATION ET « ÉMOTIONNALISATION »

L'aspect émotionnel de l'architecture de guerre n'est pas nécessairement ce qui vient en premier à l'esprit lorsqu'on y pense de façon abstraite. Plutôt fonctionnelle, on ne s'attend pas à la trouver embarrassée par la simulation affective, mais bien davantage tournée vers des exigences de résultats, et prise dans des dynamiques d'urgence et d'économie de moyens. Soit quelque chose qui laisserait peu de place pour des dimensions affectantes ou affectées. Pourtant il est difficile de nier la charge émotionnelle de certaines constructions guerrières, ou leur impact émotionnel sur l'environnement. On peut par exemple penser aux tours de Flak persistant dans l'état de Hambourg. Notamment la mémorable tour de Heiligengeistfeld, édifée en octobre 1942, soit une gigantesque masse de béton posée sur une large place au milieu de la ville. Ou encore la tour de Flak de Wilhelmsburg, édifée en octobre 1943, en

Art in general and, naturally, architecture, is a reflection of the spiritual condition of man at a particular time. But the impression exists that modern architecture, so idiosyncratic and intellectual, perhaps because it has lost close contact with its community, exaggerates at times its efforts to emphasize the rational aspect of architecture. As a result of all this, man in the twentieth century feels overwhelmed by so much "functionalism", by so much logic and usefulness in modern architecture. He looks for a way out, but neither the external aestheticism understood as "formalism", nor organic regionalism, nor dogmatic Confucianism have examined in depth the problem that the man of our time, be he creator or recipient, aspires to something more than a house that is attractive, agreeable and appropriate to his needs. He requests, or will request one day, of architecture and its modern means and materials, spiritual elevation. Simply stated, an emotion, like that given in its time to the architecture of the pyramid, of the Greek temple, of the Romanesque or Gothic cathedral, or even that of the Baroque palace. *Only when architecture affords him real emotions can man again consider it an art* [italique ajouté]. [...], Mathias Goeritz, « Manifesto of Emotional Architecture » (1953), Musée El Eco, Mexico. (Ma traduction, à partir de la version anglaise disponible sur le site Internet du musée.) <http://www.eleco.unam.mx/sitio/index.php/eng-el-eco-content/manifesto>.

plein cœur des îlots et polders du lit de l'Elbe — une zone aujourd'hui urbaine se partageant entre des secteurs hautement industrialisés (port de Hambourg) et des banlieues résidentielles. Si la tour de Heiligengeistfeld, dans le quartier de St. Pauli, a rapidement été « recyclée » après la guerre, réutilisée pour d'autres fonctions (une partie de ce volume sert actuellement de salle de concert alternative, le club Uebel & Gefährlich), en 2003 celle de Wilhelmsburg, n'avait toujours pas été explicitement reconvertie et se présentait comme une autre gigantesque masse de béton, sans ouvertures sinon des sortes de meurtrières, plantée au milieu d'un quartier de banlieue populaire. Sorte de *black box* imprenable et incompréhensible, laissée à sa dérive immobile. Quelque chose d'incongru pouvait être perçu, comme émanant du lieu. Les fuites d'eau dégoulinant le long des façades, la clôture des moindres anfractuosités, le silence entourant la chose, tous ces indices auraient pu donner l'impression d'un bâtiment en larmes, en forme de trou noir ou de gouffre émotionnel, comme s'il n'avait été fait que de souffrances compactées, murées, muettes. Un monstre édifiant et verrouillé, au coin de la rue, entre deux *Wohnblocks* (blocs d'habitation — barres d'immeubles à l'horizontale) entourés de végétation. Sans être au préalable informé de ce que pouvait être ce bâtiment, on aurait pu le confondre avec un transformateur EDF (ou Hamburg Energie, en l'occurrence) : ces sortes de guérites closes, croisées de temps en temps au détour d'un chemin ou d'une route. Mais un transformateur qui aurait enflé jusqu'à atteindre une hauteur atypique (42 mètres). Conservant néanmoins la signalétique et les avertissements relatifs aux transformateurs électriques ordinaires.

Depuis 2007, une réhabilitation du lieu a été enclenchée et sa concrétisation se rapproche (2013) : le toit de la tour va être recouvert par des panneaux solaires thermiques (générant du chauffage) ; la façade sud par des panneaux photovoltaïques (générant de l'électricité) ; l'intérieur va accueillir une centrale énergétique de biogaz (générant de l'électricité et du chauffage) ainsi qu'une centrale de cogénération biomasse au bois (générant de l'électricité et du chauffage) et les surplus de chaleur des usines alentours seront réorientés vers le bâtiment pour y être stockés en attendant une redistribution dans les habitations des environs. Dans cet exemple, la capacité isolante aura été retournée en opportunité de « bunkeriser » la chaleur afin de limiter les déperditions énergétiques (« *Der Wärmespeicher*



Tour de Flak de Heiligengeistfeld (Robert Kuhn, *Bunker in Hamburg*, Hambourg, 2006).
© PastFinder Ltd, Maik Kopleck & Robert Kuhn.



Tour de Flak de Wilhelmshurg (Robert Kuhn, *Geschütz-Turm in Hamburg-Wilhelmshurg*, Hambourg, 2006).
[Notons ici que «G-Turm» ne signifie pas *Geschütz-Turm* (tour de protection, de défense) mais *Gefechtsturm* (tour de combat). Mais il est vrai que ces abris anti-aériens ont aussi servi à héberger des milliers de personnes simultanément, pendant les bombardements.]
© PastFinder Ltd, Maik Kopleck & Robert Kuhn.

“*bunkert*” die Wärme⁴⁴.») Et la tour de Flak de Wilhelmsburg, cette bâtisse réputée rationnelle et figée, va ainsi devenir l’un des moteurs énergétiques du quartier — renouant en un sens avec son histoire de concentrateur d’énergies : cette G-Turm (*Gefechtsturm*) ayant aussi servi d’abri anti-aérien capable d’accueillir (ou de stocker) des milliers de personnes simultanément.

Notons aussi que ces tours ont en partie été conçues par l’architecte nazi Friedrich Tamms pour le compte de l’Organisation Todt (O.T.), menée par l’ingénieur nazi Fritz Todt (en charge de la construction des autoroutes du Reich à partir de 1933) jusqu’à sa mort en 1942, puis par l’architecte nazi Albert Speer. Comme le précise l’historienne d’art Silke Wenk⁴⁵, ces constructions de guerre s’inscrivaient dans des projets d’urbanisme plus vastes. Peut-être n’étaient-elles pas non plus sans discrètes filiations d’avec le Bauhaus, du fait notamment de l’architecte nazi Ernst Neufert, assistant (pendant une décennie environ) de Walter Gropius à Weimar, à partir de 1921, avant de rejoindre Speer en 1939 pour travailler sur les nouvelles normes architecturales (rationalisation) puis, à partir de 1940, sur des prototypes de maisons individuelles incluant elles aussi des abris anti-aériens (en prévision des temps de « paix » à venir)⁴⁶. Il conviendrait ici de creuser les liens (ou l’absence de liens) entre Tamms et Neufert. S’ils ont travaillé ensemble à la reconstruction de l’Allemagne après la Seconde Guerre mondiale, il faudrait savoir s’ils avaient déjà eu l’occasion de se rencontrer et d’échanger lorsque tous deux travaillaient en tant qu’illustres architectes du III^e Reich, l’un pour l’O.T., l’autre pour Speer, avant que ce dernier ne reprenne la direction de l’O.T. en 1942.

Tout ceci nous amène de façon un peu détournée jusqu’à un point qui peut avoir son importance : le fait que le dictateur nazi Adolf Hitler ait pris soin de construire sa personnalité verbale sur la légende d’une vocation d’architecte (de *Baumeister*, de maître bâtisseur) empêchée. Son livre, *Mein Kampf* (« mon combat », 1925-1926), se fonde quasiment sur ce point et donne à percevoir dans la stratégie politico-militaire qui y est déployée la création d’un personnage, d’un guide, en forme d’architecte de l’émotion des foules. Si l’idée n’était pas

⁴⁴ Cf. la brochure de description du projet, porté par IBA Hamburg,

http://www.iba-hamburg.de/fileadmin/Mediathek/S11_energiebunker/Projektflyer_Bunker_110916.pdf.

⁴⁵ Cf. Silke Wenk, « Bunkerarchäologien - Zur Einführung », dans Silke Wenk (directrice de publication), *Erinnerungsorte aus Beton - Bunker in Städten und Landschaften*, éd. Christoph Links, Berlin, 2001, pp. 30-31.

⁴⁶ Cf. note précédente.

nouvelle, la génération de cette créature du III^e Reich s'est fait en utilisant les moyens les plus modernes, s'appuyant sur les conceptions de l'époque, en matière de propagande et de façonnage des opinions, d'où peut-être, aussi, son succès. *Mein Kampf* fournit sa propre analyse de la communication propagandaire utilisée par les Américains, les Britanniques et les Français durant la Première Guerre mondiale et n'est pas sans évoquer les travaux du neveu de Sigmund Freud, l'américain Edward L. Bernays, théoricien et praticien de la *publicité* (dans tous les sens du terme), qui dès 1923 publie un livre consacré aussi bien aux relations publiques qu'à la fabrique de l'opinion en démocratie, *Crystallizing Public Opinion*, soit un ensemble de méthodes permettant de structurer et d'orienter la fluctuation des avis (et des goûts). Si cet ouvrage de Bernays, ainsi que son ouvrage postérieur, *Propaganda* (1928), sont souvent cités comme ayant influencé Joseph Goebbels⁴⁷, en charge de la propagande nationale-socialiste (en tant que *Reichspropagandaleiter*) de 1930 à 1945, succédant ainsi à Hitler (1928-1930), on trouve déjà dans *Mein Kampf* le détail d'un homme événementiel — créateur d'événements —, capable de cristalliser, d'architecturer, de former les émotions des masses. Pourtant il est peu probable que le rédacteur de *Mein Kampf* ait lu *Crystallizing Public Opinion*⁴⁸. La conception de Bernays, en matière de communication et de conseil en relations publiques, plus humaniste et davantage tournée vers l'établissement d'une homéostasie démocratique dynamique (presque sans valeurs absolues statiques) au moyen des leviers de la « manipulation affective », aura radicalement été dépouillée de ses scrupules démocratiques par la machine de guerre nazie pour devenir une arme au service du meurtre et de la dissimulation. Il importe de rappeler ces quelques éléments afin de pouvoir se souvenir que le III^e Reich n'a en quelque sorte jamais perdu de vue les aspects esthétiques et émotionnels que véhicule toute objectivation, toute organisation, et n'aura jamais négligé de contrôler ces leviers. Le fait de parler d'une « rationalisation » de l'architecture sous le III^e Reich induit peut-être en erreur dans la mesure où nous avons communément l'habitude (à tort) d'opposer raisons et sentiments : la dimension affective et émotionnelle a toujours fait partie des priorités du nazisme, en tant que

⁴⁷ Cf. Dennis W. Johnson, *Routledge handbook of political management*, éd. Routledge, New York, 2009, note 3, p. 314. Mais peut-être convient-il de traiter cette assertion avec prudence.

⁴⁸ Sauf si « l'édition originale » de *Mein Kampf* (sa version allemande) avait été réécrite au fur et à mesure de ses rééditions, entre 1924 et 1934, année de sa traduction française. Une autre hypothèse étant que ces idées (ou méthodes) aient été « dans l'air du temps », c'est-à-dire déjà largement répandues et culturellement palpables, aussi bien pour Bernays que pour Hitler.

facteur amplifiant devant être compris et exploité. Si les émotions collectives étaient fabriquées de façon rationnelle, elles n'en devaient pas moins être réellement ressenties (de l'adoration à la terreur) par les gens. Les tours de Flak, bien que structures hautement fonctionnelles, n'ont pas nécessairement échappé à l'imprégnation émotionnelle. Elles se sont en quelque sorte vues chargées de dimensions communicantes (ou propagandaires). Écrasants coffres-forts bétonnés, érigés dans le paysage, on peut s'étonner de la simplicité avec laquelle elles parviennent encore à transmettre des sensations de concentration, de repli, de clôture, d'enroulement sur soi et de souffrance autiste (terme venant de l'allemand *Autoerotismus*). La «sobre élégance» que certaines d'entre elles réussissent à dégager n'est peut-être pas hasardeuse dans la mesure où ces formes ont été pensées au sein d'un système assez cohérent, par des artistes (des architectes) dont le métier consistait aussi à savoir produire de l'émotion (Friedrich Tamms, Ernst Neufert, Albert Speer, entre autres). Nous avons eu l'occasion de nous habituer à ce type d'esthétisation monolithique [cf. Virilio, 2008 ; Rubio, 2011], agressive / agressée, ici poussée à l'extrême. Et pour reprendre l'hypothèse défendue plus haut et jusqu'ici : cette «rationalité» n'est pas nécessairement synonyme de neutralité émotionnelle. Elle véhicule même sans difficulté des sentiments d'effroi, de pétrification, d'asphyxie. Comme le note Virilio à propos du «Mur de l'Atlantique», dans *Bunker archéologie* (1975) : «Toute une série de réminiscences culturelles me saisissent : les mastabas, les tombes étrusques, les structures aztèques... Comme si l'organisation Todt n'avait su finalement organiser qu'un espace religieux...» [Virilio, 2008, p. 16], et plus loin : «les bunkers du littoral européen sont dès l'origine des monuments funéraires du rêve allemand.» [Virilio, 2008, p. 40].

Peut-être peut-on s'étonner du fait que, parfois, la dimension douloureuse (les émotions négatives) de ces architectures puisse se changer en froide rigueur privée d'âme, par le biais du discours. Cette ambivalence de statut se retrouve aussi, non dans le cœur du livre de Paul Virilio, mais dans l'une de ses phrases qui l'illustre alors parfaitement : «Sa poésie [au blockhaus], c'est de n'être encore qu'un simple bouclier pour ceux qui l'utilisent, finalement aussi désuet qu'une armure d'enfant reconstituée, coquille vide, fantôme touchant d'un duel dépassé où les adversaires pouvaient encore s'observer directement au travers de la fente étroite de leur visière rebattue.» [Virilio, 2008, p. 63]. Ambivalence de ne pas savoir s'il s'agit d'une

coquille vide (*shell*) ou d'un fantôme (*ghost*). Ambivalence de ne pas savoir comment se positionner face à une telle concentration d'énergie négatrice. Là même où les bâtiments réputés pour dégager une sorte de joie (par exemple les travaux d'Antoni Gaudí, que Neufert rencontre à Barcelone en novembre 1920), sont tenus pour avoir une sorte d'âme en propre. Et peut-être n'est-ce pas si éloigné du *Lüge* croisé plus haut, c'est-à-dire de cette sorte de mensonge inanimé qui se ferait passer pour animé. Seulement le mensonge serait inversé : l'animé se ferait plutôt passer pour inanimé, pour désuet et vide, de façon à ne pas trop attirer l'attention (à l'image de l'expression anglo-saxonne : *to play dead*, c'est-à-dire « faire le mort », soit une expression illustrant une technique de camouflage — éventuellement émotionnel).

Si, comme le sous-entendait Goeritz, l'architecture reflète une époque, des dispositions culturelles, que reflètent alors nos ouvrages d'architectures informatiques les plus élaborés? Des fantômes? Des zombies? Des *p-zombies* (zombies philosophiques [cf. Dennett, 1991])? Et ce positionnement fait-il référence à un « Kant en réduction », à travers l'idée que seule la beauté puisse être dotée d'une âme (sous-entendue positive)? N'est-ce pas naïf de croire qu'on puisse décréter verbalement l'attribution, ou non, d'une âme aux choses? Rappelons de même que les tours de Flak ont été construites dans l'urgence pendant la guerre, en période de pénuries, par des milliers de travailleurs étrangers ou allemands, volontaires ou esclaves.

En somme, la volonté d'émotionaliser les architectures ou les sociétés ne peut être abordée de façon légère, à moins de choisir de faire l'impasse sur certaines dérives déjà observées. Mais on peut aussi entendre, à travers cette émotionalisation, la tentative (peut-être vaine) d'infléchir les émotions destructrices.

Ainsi, le parcours à travers les bribes précédentes (Rahm, le cycle physiologique de la mélatonine, Vitruve, Sun Tzu, les palais de la mémoire, Scheerbart, la *Gläserne Kette*, le « Village enchanté », l'approche évolutionniste, Goesch, Goeritz, l'architecture cathartique, les tours de Flak de Hambourg, etc.) nous aura permis de dégager plusieurs points : 1) la poursuite d'une tendance à l'objectivation des émotions par et dans l'architecture, 2) l'importance des techniques employées, 3) l'affinement progressif des moyens (techniques et conceptuels) mis en œuvre pour y parvenir, 4) les problèmes que cela peut engendrer.

Avec les deux paragraphes précédents, un retour sur des aspects historiques de l'architecture de l'émotion aura été tenté. Si le sujet est inépuisable, la priorité aura été mise sur certains des développements ayant été théorisés et mis en pratique dans la première moitié du XX^e siècle. Là encore, l'exercice se sera fait au pas de course, à grands traits, et nous pouvons le déplorer. Le parti pris aura été celui de mettre en avant le paradoxe, l'ambiguïté, voire le nœud coulant : si la négation des émotions est dommageable, leur exploitation peut s'avérer tout aussi délétère. Mais il serait dommage d'en conclure que cela relève d'une fatalité. Disons plutôt que les horreurs passées, aussi pénibles soient-elles dans le travail de remémoration, ne doivent pas être effacées par convenance : même si tout se délite, même si la mémoire est trompeuse, même si le réel ne brille que par son absence, l'une des façons de ne pas répéter et propager les mêmes erreurs, c'est encore d'essayer de les comprendre, ne serait-ce que confusément si la clarté venait à manquer ou à devenir complètement inaccessible. En dernière instance, il aura surtout été tenté d'insister sur le fait qu'il ne soit pas possible de faire l'économie des émotions. S'il est possible de les réguler, d'en faire l'objet d'une gestion sociale, il n'est pas possible de les supprimer. Du moins pas sans intervention physique.

NEUROBIOLOGIE ÉMOTIONNELLE

Le paragraphe suivant se propose d'observer un autre type d'architecture émotionnelle, un autre type de représentations de l'émotion sous forme d'architecture. Contemporaine cette fois, elle nous permettra de répondre à la question initiale, à savoir qu'est-ce que l'émotion, qu'est-ce qu'une émotion. C'est volontairement que ce problème de définition n'aura pas été abordé plus tôt : il ne s'agit pas de proposer une théorie des émotions, tout juste s'agit-il de prendre de la distance en grappillant dans différentes approches actualisées, au moyen de reports (subjectifs) d'observations.

Jusque là, donc, les termes émotions et sensations auront été employés de façon assez indistincte et mêlée, supportant à peine une petite distinction entre une sensation qui pencherait plutôt du côté de la perception « épidermique » et sensorielle (odorat, ouïe, toucher, vue...) et une émotion qui pencherait du côté d'un subjectivisme, certes plus verbalisé, mais

tout autant indémontrable — participant d'une mythologie personnelle sédimentée. Et les deux mots ont été grossièrement employés de façon relativement interchangeable. Dans le domaine des neurosciences, par contre, il y a des distinctions nettes et précises quant à l'usage de ces termes, et le neurobiologiste portugais António R. Damásio, notamment, les a clairement explicitées. Dans un entretien de 2003 paru dans *La Recherche*. Il y expliquait ainsi son emploi des concepts d'émotion (*emotion*) et de sentiment (*feeling*) :

« Une émotion est, face à un stimulus, une collection complète de réponses chimiques et neurales automatiques formant une structure distinctive. Un sentiment est la transcription de cette émotion sur le théâtre de l'esprit à l'aide d'un processus conduisant à la production d'images mentales. Même si la distinction nous semble difficile à faire, nous éprouvons une émotion de tristesse avant d'éprouver un sentiment de tristesse. Divers accidents cérébraux et protocoles expérimentaux permettent au chercheur de distinguer clairement entre les deux phénomènes⁴⁹. »

Soit une explication complétant assez bien ce qu'il note dans un ouvrage plus ancien (1995), *Descartes' Error – Emotion, Reason and the Human Brain* (« l'erreur de Descartes – l'émotion, la raison et le cerveau humain ») à propos des sentiments :

« Je conceptualise l'essence des sentiments [*feelings*] comme étant quelque chose que vous et moi pouvons voir à travers une fenêtre s'ouvrant directement sur une image continuellement mise à jour de la structure et de l'état de notre corps. Si vous imaginez que la vue par cette fenêtre est comme un paysage, alors la "structure" du corps est analogue à la forme des objets dans un espace, tandis que les "états" du corps ressemblent à la lumière, à l'ombre, au mouvement, au son de ces objets dans cet espace. Dans le paysage de votre corps, les objets sont alors les viscères (le cœur, les poumons, les intestins, les muscles), tandis que la lumière, l'ombre, le mouvement et le son représentent une position sur l'intervalle comprenant toutes les opérations possibles pour ces organes, à un moment donné. En gros, un sentiment correspond à la "vue" momentanée d'une partie de ce paysage corporel. Il a un contenu spécifique — l'état du corps — ; et des systèmes neuraux spécifiques qui le rendent possible — le système nerveux périphérique et les régions du cerveau intégrant les signaux en rapport avec la structure et la régulation du corps. Et comme le sens de ce paysage corporel est juxtaposé dans le temps avec la perception ou le souvenir de quelque chose d'autre, ne faisant pas parti du corps — un visage, une mélodie, un arôme —, les sentiments finissent par devenir les "qualificatifs" de cette autre chose. Mais il y a plus que cette essence dans un sentiment⁵⁰. » [Damásio, 1995, pp. XIV-XV]

⁴⁹ Olivier Postel-Vinay, « António Damásio : "L'esprit est modelé par le corps" », entretien paru dans *La Recherche*, n°368, octobre 2003, pp. 69-72.

⁵⁰ « I conceptualize the essence of feelings as something you and I can see through a window that opens directly onto a continuously updated image of the structure and state of our body. If you imagine the view from this window as a landscape, the body "structure" is analogous to object shapes in a space, while the body "state" resembles the light and shadow and movement and sound of the objects in that space. In the landscape of your body, the objects are the viscera (heart, lungs, gut, muscles), while the light and shadow and movement and sound represent a point in the range of operation of those organs at a certain moment. By and large, a feeling is the momentary "view" of a part of that body landscape. It has a specific content—the state of the body; and specific neural systems that support it—the peripheral nervous system and the brain regions that integrate signals related to body structure and regulation. Because the sense of that body landscape is juxtaposed in time to the perception or recollection of something else that is not part of the body—a face, a melody, an aroma—feelings end up being "qualifiers" to that something else. But there is more to a feeling than this essence. » [Damásio, 1995, pp. XIV-XV]. (Ma

Cette métaphore (architecturale) de la fenêtre peut là encore nous ramener à Vitruve, s'adressant à Jules César dans l'introduction du troisième livre de *L'architecture* (vers 15 av. J.-C.) : « Socrate, qu'Apollon lui-même, par son oracle de Delphes, déclara le plus sage de tous les hommes, disait avec raison, qu'il serait à souhaiter que nous eussions une ouverture à la poitrine, afin qu'on pût y lire nos pensées et nos desseins⁵¹. » Ce qui s'accorde assez bien avec l'idée défendue par Damásio, selon laquelle les sentiments nous permettent d'être en lien avec ces boîtes noires physiologiques que sont nos corps émotionnels, nous rapprochent de la possibilité de conscience. Damásio explicite l'émotion en ces termes :

« Je vois l'essence de l'émotion comme étant une collection de changements dans l'état du corps, changements induits dans une myriade d'organes par les terminaisons nerveuses des cellules, sous le contrôle d'un système cérébral dédié, répondant au contenu des pensées relatives à une entité ou à un événement particuliers. Beaucoup de changements dans l'état du corps — ceux de la couleur de la peau, de la position corporelle, des expressions faciales, par exemple — sont en réalité perceptibles pour un observateur externe. (Et en effet l'étymologie du mot suggère à juste titre une direction allant vers l'extérieur, à partir du corps : *émotion* signifie littéralement "mouvement hors de".) D'autres changements dans l'état du corps ne sont seulement perceptibles que pour le propriétaire du corps dans lequel ils se produisent. Mais il y a plus dans l'émotion que son essence.

En conclusion, l'émotion est la combinaison d'un *processus mental évaluatif*, simple ou complexe, avec des *réponses dispositionnelles à ce processus*, le tout principalement *orienté vers le corps en lui-même* et aboutissant en un état du corps, mais *aussi orienté vers le cerveau en lui-même* (noyaux [de neurones producteurs] de neurotransmetteurs dans le tronc cérébral) et aboutissant en des changements mentaux supplémentaires. Notons que, pour le moment, je laisse en dehors de l'émotion la perception de tous les changements qui constitue la réponse émotionnelle. Comme vous le découvrirez, je réserve le terme de *sentiment* à l'expérience de ces changements⁵². » [Damásio, 1995, p. 139]

Si l'on tente de synthétiser les résultats de Damásio, il semblerait que nous puissions peut-être distinguer comme suit : 1) Il y a un *stimulus* (interne ou externe). 2) Ce stimulus est

traduction, aidée de la traduction de Marcel Blanc, *L'erreur de Descartes – La raison des émotions*, éd. Odile Jacob, Paris, 2006, pp. 11-12.)

⁵¹ Vitruve, *L'Architecture de Vitruve*, Livre III, trad. par De Bioul, éd. Adolphe Stapleaux, Bruxelles, 1816, pp. 92. (Traduction légèrement modernisée : « disait » au lieu de « disoit », etc.)

⁵² « I see the *essence* of emotion as the collection of changes in body state that are induced in myriad organs by nerve cell terminals, under the control of a dedicated brain system, which is responding to the content of thoughts relative to a particular entity or event. Many of the changes in body state—those in skin color, body posture, and facial expression, for instance—are actually perceptible to an external observer. (Indeed, the etymology of the word nicely suggests an external direction, from the body: *emotion* signifies literally "movement out.") Other changes in body state are perceptible only to the owner of the body in which they take place. But there is more to emotion than its essence.

In conclusion, emotion is the combination of a *mental evaluative process*, simple or complex, with *dispositional responses to that process*, mostly *toward the body proper*, resulting in an emotional body state, but also *toward the brain itself* (neurotransmitter nuclei in brain stem), resulting in additional mental changes. Note that, for the moment, I leave out of emotion the perception of all the changes that constitute the emotional response. As you will soon discover, I reserve the term *feeling* for the experience of those changes. » [Damásio, 1995, pp. 139]. (Ma traduction, aidée de la traduction de Marcel Blanc, *L'erreur de Descartes – La raison des émotions*, éd. Odile Jacob, Paris, 2006, p. 183.)

éventuellement générateur de *sensations*, c'est-à-dire de perceptions d'une certaine classe de signaux sensoriels [Damásio, 1995, p. 265] — sensations « directement » conscientes (piqûre, brûlure, etc.) ; ou sensations inconscientes ou différées (modification du taux de mélatonine).

3) Parallèlement à la sensation, le stimulus donne lieu à une évaluation des modifications occasionnées, ce qui engendre une *émotion*, c'est-à-dire une réponse de l'organisme, parmi le panel des réponses possibles. 4) Cette réponse émotionnelle est elle-même évaluée, ce qui donne lieu à un *sentiment*. Mais selon Damásio, ni sensation, ni émotion, ni sentiment ne sont forcément conscients. La *conscience* supposant d'être capable de réfléchir sur ce que l'on éprouve, capable de réfléchir sur les images mentales occasionnées (les sentiments) [cf. entretien de 2003]⁵³. L'un des exemples mis en avant par le neurobiologiste est celui d'un patient souffrant d'une douleur aussi insupportable qu'incurable. Face à ce cas, la solution apportée avait alors été de pratiquer une leucotomie préfrontale (section de quelques fibres nerveuses entre le cortex préfrontal et le reste du cerveau), et le résultat en avait été le suivant : le patient était toujours capable de ressentir les sensations de douleur, mais n'éprouvait plus la souffrance associée à ces douleurs. Autrement dit, les sensations étaient intactes, mais les sentiments avaient été supprimés. Il ne souffrait donc plus. « [Le neurochirurgien Almeida] Lima lui demanda ce qu'il en était de ses douleurs. L'homme [, occupé à jouer une partie de cartes avec une ami,] releva la tête et répondit joyeusement : "Oh, les douleurs sont les mêmes, mais je me sens bien maintenant, merci." [...] Tandis que le cerveau ne pouvait plus engendrer de souffrance, il continuait de produire des "images de douleur", c'est-à-dire, de traiter normalement l'encartage

⁵³ D'où le titre « L'erreur de Descartes » : Damásio retourne le « je pense donc je suis » en un je suis et ensuite je pense (« We are, and then we think » [Damásio, 1995, p. 248]). Il y a au moins deux occurrences du *cogito (± ergo) sum*. La très célèbre : « Et remarquant que cette vérité : *je pense, donc je suis*, était si ferme et si assurée, que toutes les plus extravagantes suppositions des sceptiques n'étaient pas capables de l'ébranler, je jugeai que je pouvais la recevoir, sans scrupule, pour le premier principe de la philosophie que je cherchais. » René Descartes, *Discours de la méthode* (1637), quatrième partie ; et une autre occurrence, non moins célèbre : « Lorsque j'ai fait ainsi une abdication de toutes mes connaissances passées, je commence à philosopher de la sorte : je suis ; je pense ; je suis pendant que je pense. Cette proposition, *j'existe*, est nécessairement vraie, toutes les fois que je la prononce ou que je la conçois en mon esprit. », René Descartes, réponses aux objections formulées par P. Bourdin aux *Méditations métaphysiques*, publié dans *Méditations métaphysiques*, « Septièmes Objections » (1645). Cette seconde occurrence ne manque pas de rappeler la question évoquée en première partie du présent propos : « *Can machines think?* ». Si la vérité du « j'existe » cartésien est limpide pour tout être humain au premier abord, avec un programme d'intelligence artificielle ce fondement semble s'effriter ?

somatosensoriel d'un paysage de douleur⁵⁴.» [Damásio, 1995, p. 266].

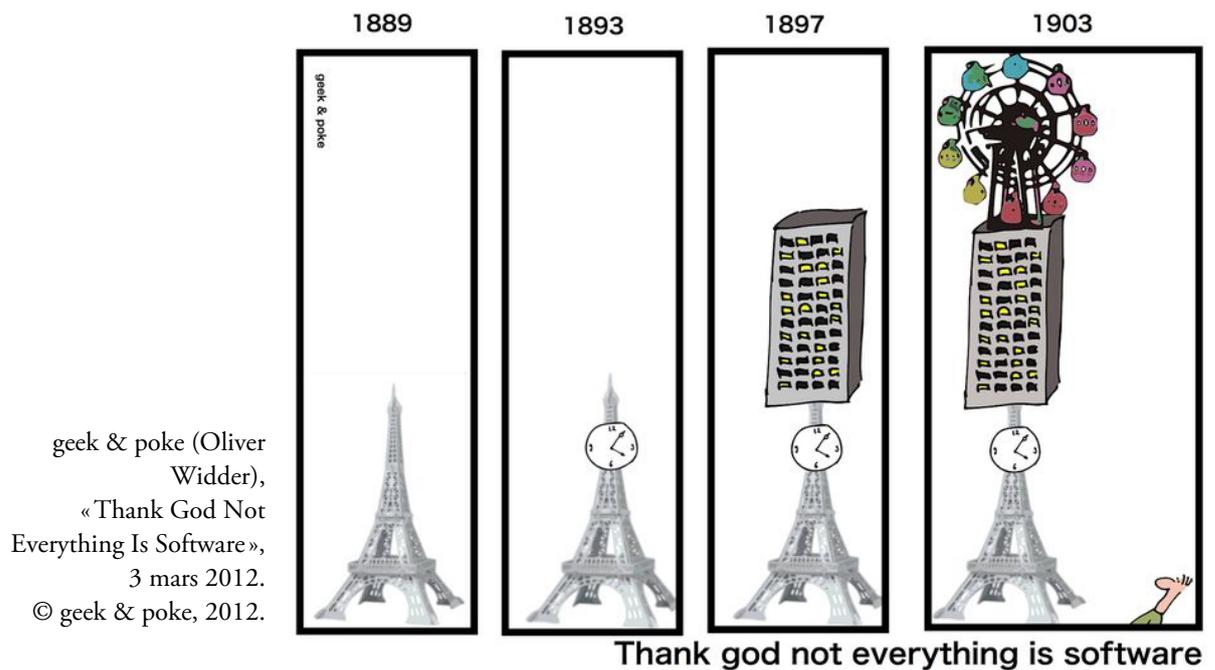
Bien que le paragraphe précédent puisse paraître compliqué, il n'en relève pas moins d'une autre forme, plus contemporaine, d'architecture de l'émotion. Lorsque Rahm entre dans le détail des processus hormonaux en ciblant parmi des longueurs d'ondes, il nous donne aussi à percevoir le déplacement du curseur quant à la précision des conceptions actuelles. L'architecture de l'émotion ne se joue plus seulement avec des formes de nautilus (spirales), elle tient dorénavant à quelques fils, quelques fibres connectées de telles ou telles manières et les sentiments ne sont plus garants des émotions vécues. Du moins dans une partie de la pensée occidentale contemporaine. Bien sûr, comme nos rapports aux œuvres, lorsque verbalisés, sont forcément conscients, il n'est pas forcément pertinent de tenter de démêler sensations, émotions et sentiments, par rapport à ce qui peut-être éprouvé en présence d'un événement ou d'une entité, quels qu'ils soient. Cette précision relève d'outils techniques dont nous ne disposons pas. Cela dit, *Diurnisme* semble pouvoir aider à la mise en évidence de ces distinctions. Mais on peut aussi penser à un autre exemple, un peu différent : la conductance de la peau et le rougissement de celle-ci n'ont pas toujours besoin d'être discernables à l'œil nu pour être opérants. Il existe par exemple un miroir augmenté capable d'informer ceux qui s'y regardent de leur rythme cardiaque (l'évaluation se base sur le fait qu'à chaque pulsation cardiaque, tandis que le sang « afflue » dans l'épiderme, la peau ne reflète alors pas la lumière de la même manière⁵⁵). Ce miroir et ces travaux sur l'épiderme ont été menés par l'étudiant du MIT Ming-Zher Poh, soutenu par Daniel McDuff du MIT Media Lab et par Rosalind Picard, précédemment citée à propos de « l'informatique affective » (*affective computing*).

La deuxième partie de ce travail aura ainsi proposé de s'intéresser aux formes objectivées de la construction émotionnelle, qu'elles soient revendiquées comme telles ou non. Le paradoxal travail de Philippe Rahm, *Diurnisme*, nous aura indirectement offert une occasion

⁵⁴ « Lima asked him about the pain. The man looked up and said cheerfully: "Oh, the pains are the same, but I feel fine now, thank you." [...] While the brain could no longer engender suffering, it was still making "images of pain," that is, processing normally the somatosensory mapping of a pain landscape. » [Damásio, 1995, p. 266]. (Ma traduction).

⁵⁵ David L. Chandler, « Your vital signs, on camera – MIT team develops system for continuous medical monitoring using widely available video technology », sur *MIT news*, 4 octobre 2010, <http://web.mit.edu/newsoffice/2010/pulse-camera-1004.html>.

synthétique de dérouler quelques grandes lignes motrices du XX^e, tout en nous arrimant depuis un point de vue aussi précis que définitivement situé dans le XXI^e siècle. L'hypothèse de départ tournait autour de l'idée que l'informatique affective pourrait s'affilier à l'architecture des émotions : ce qui précède aura permis de tenter une présentation partielle de cette dernière. De façon « inattendue », nous aurons été amenés à croiser la biologie et ses méthodes (l'œuvre de Rahm s'y prêtait néanmoins). Cela nous aura également amenés à réfléchir aux conséquences de l'influence massive, sans pour autant clore la question de la mise en architecture des émotions.



III. LA FABRIQUE DU SUJET

Si l'on s'intéresse à l'architecture de l'émotion, comme thème sujet à variations, l'informatique affective telle que proposée par Picard offre encore une autre entrée dans celle-ci, en tant qu'occurrence de l'architecture informatique. Il peut sembler improbable de vouloir faire entrer la «sensibilité émotionnelle» et «l'intuition» dans ce champ réputé jusqu'au troisième quart du XX^e siècle pour être le lieu de la rationalité, de l'homme ordinateur ou calculateur (*computer man*), de l'humain mathématique. Quoi de plus froid et mort qu'un ordinateur? Pourtant il s'agit bel et bien de mixer ces deux *images* : celle de la froideur inorganique des machines, avec celle de la tumultueuse et chaleureuse (parfois brûlante et chaotique) passion animale. La dernière partie sera donc consacrée à la mise en relation des deux topiques succinctement développées en amont : à savoir les agents conversationnels d'une part et d'autre part l'architecture de l'émotion.

Cette façon de procéder (par morceaux, puis par réunion) peut sembler particulièrement intuitive et synthétique, plus qu'analytique et constructive. Aussi devons-nous rappeler que ce n'est pas un mémoire d'ingénierie (en singer l'apparence serait douteux et proposer des pistes sans savoir comment est articulé ce qui existe déjà n'est pas forcément d'une grande aide), mais nous pouvons surtout en profiter pour noter quelque chose à propos des questions d'intuition et de méthode. Cela nous servira lorsque l'intuition sera de nouveau évoquée dans quelques lignes : nous aurons déjà en mémoire ce qui va suivre et cela nous fera gagner du temps (et de l'espace en l'occurrence).

Dans son livre *Crystals, Fabrics and Fields*, Haraway insiste sur l'importance du sens de la beauté, ou plutôt de ce qui est «approprié», dans l'élaboration des théories organicistes. Et elle revient sur le rôle méthodologique de l'intuition dans la construction scientifique :

«Agnes Arber et C. F. A. Pantin, sont deux écrivains contemporains ayant là encore insisté sur la relation forte entre l'art et la biologie, enracinée dans le problème de la forme et la primauté de la vision. Dans ce contexte, les deux penseurs se réfèrent au sens illatif de John Henry Newman, en tant que pierre angulaire de la conscience métaphorique en biologie. Le sens illatif en science est lié au goût en art, c'est-à-dire au sens de l'approprié. C'est de la connaissance fusionnée avec de l'émotion et entrelacée avec des images sensibles. Aussi bien les problèmes que les solutions, en biologie comme en art, émergent initialement d'une certaine manière "à partir du système des images par lequel nous représentons le

monde dans nos esprit” (Pantin, 1968, p. 127). Arber note que l’intuition émergeant des prédispositions esthétiques et émotionnelles d’un scientifique peut être spacieuse et se désintégrer sous la pression de l’expérimentation, l’autodafé de toutes les belles théories. Mais aucune véritable question n’existerait sans l’interaction intime entre image, science, et art (Arber, 1954). Le sens illatif est indispensable à l’approche de la biologie quant aux énigmes de la forme parce qu’ “il initie un puissant besoin [*drive*] émotionnel envers la solution du problème, peut-être à travers la perception inconsciente du fait que tout ou partie de ce système d’images par lequel nous représentons le monde naturel en nos esprits, et qui est la base de notre sens commun, est sur le point de subir une métamorphose” (Pantin, 1968, p. 12). Précisément, une telle métamorphose d’image était à l’œuvre dans le travail de Harrison, Needham, et Weiss. Leur attention aux standards esthétiques et au problème de la forme n’était pas un accident⁵⁶. » [Haraway, 1976, p. 42]

Ce non-accident est selon Haraway lié à la conception même de l’organicisme (cette attention à la forme et à l’organisation des éléments entre eux de sorte que le tout forme plus que la somme des parties, soit un souci partagé par certains artistes, poètes et biologistes à la suite d’Aristote — n’oublions pas les architectes, en tant qu’artistes). Le sens illatif (l’intuition) est ici rapporté comme étant ce qui motive les premières incursions dans la recherche, ce qui permet de jeter des ponts, de formuler des hypothèses qu’il s’agira ensuite, peut-être, de déconstruire à l’aide d’une confrontation expérimentale. Si l’intuition ne mène pas directement à la compréhension, elle permet d’enclencher un processus de recherche en produisant des conclusions bonnes pour la mise à l’épreuve : elle oriente. En ce sens elle fait partie d’une démarche systématique : le système de représentations du monde adopté (par une personne, par une communauté) va permettre de postuler par inférence un certain nombre d’attentes, et l’expérimentation va servir à tester ces inférences. Ce passage du livre de Haraway est intéressant dans la mesure où il met en lumière la situation critique de l’intuition : Ou bien elle

⁵⁶ « Agnes Arber and C. F. A. Pantin are two contemporary writers who have again emphasized the strong relationship of art to biology rooted in the problem of form and the primacy of vision. In this context both thinkers refer to the illative sense of John Henry Newman as the cornerstone of the metaphoric consciousness in biology. The illative sense in science is related to taste in art, that is, to a sense of the appropriate. It is knowledge fused with emotion and interwoven with sense images. Both problems and resolutions, in biology as in art, arise initially in some way “from the system of images by which we represent the world in our minds” (Pantin 1968, p. 127). Arber observes that the intuition arising from the aesthetic and emotional predispositions of the scientist can be spurious and can disintegrate under the pressure of experiment, the *auto-da-fé* of all fine theories. But no true questions would exist without the intimate interplay of image, science, and art (Arber 1954). The illative sense is indispensable to biology’s approach to the riddles of form because “it initiates a powerful emotional drive towards the solution of the problem, perhaps through the unconscious perception that some part or all of that system of images by which we represent the natural world to our minds, and which is the basis of our common sense, is about to undergo metamorphosis” (Pantin 1968, p. 12). Precisely such a metamorphosis of image was occurring in the work of Harrison, Needham, and Weiss. Their attention to aesthetic standards and the problem of form was no accident. » [Haraway, 1976, p. 42]. (Ma traduction.)

Arber Agnes, 1954, *The Mind and the Eye*. London: Cambridge University Press.

Pantin C. F. A., 1968, *The Relations Between the Sciences*. London: Cambridge University Press.

dénote un système de représentations permettant d'expliquer avec une approximation convenable les faits observés (rapportés) ; ou bien elle dénote un système de représentations ne permettant plus d'expliquer les faits expérimentaux rapportés. Dans le second cas, soit l'attrait du système représentationnel est trop fort pour permettre une remise en question de celui-ci (bien que les faits rapportés contredisent l'intuition, ils sont interprétés comme la validant) ; soit une bascule s'opère et l'invalidation des attentes intuitives conduit à une remise en question et à un changement de système représentationnel. En ce sens, l'intuition est nécessaire à tout processus de recherche. Notons qu'un système représentationnel ne rencontrant aucun obstacle expérimental n'incitera peut-être pas à parler « d'intuition », mais plutôt de « logique inférentielle » tributaire de la causalité. Le terme intuition est en quelque sorte entaché de relativisme : lui est attaché une valeur, celle de la fragilité du système représentationnel lui permettant d'exister (qu'il soit à construire, donc à prouver ; ou à déconstruire, donc à invalider). Remarquons à titre anecdotique que l'usage est encore assez répandu de ne pas attribuer de faculté intuitive au chromosome Y, dans la mesure où les porteurs de chromosomes XX sont réputés pour être plus intuitifs que les porteurs de chromosomes XY, auxquels échoit davantage la faculté d'inférence logique (on aura bien sûr reconnu dans cette description l'aussi fameuse que douteuse « intuition féminine »). Toujours est-il que l'intuition est communément raccordée à l'émotion (dans le sens courant du terme), et c'est à ce titre que ses représentations nous intéressent ici.

La dernière partie va se dérouler comme suit : l'attention sera d'abord portée aux contextes et aux traces de l'émotionalisation en tant que valeur décisive face à une rationalisation de plus en plus décriée. Ensuite, nous nous intéresserons à l'informatique affective telle qu'elle a pu être formulée par Picard en 1997, ce qui nous amènera à revoir les agents conversationnels sous l'angle d'un exercice émotionnel que nous appellerons « test de Picard » et dont nous tenterons de brosser les contours à grands traits. L'importance accordée à la subjectivité nous conduira vers le problème de l'interprétation et pour finir, il sera proposé de repenser la cristallisation des opinions publiques en termes de fluidification des émotions individuelles.

VERS UNE CULTURE DE L’AFFECTIF ?

L’une des hypothèses du présent travail est que l’introduction des émotions en informatique s’accompagne de changements socioculturels (retournant ainsi l’hypothèse plus courante voulant que l’attention portée aux émotions n’épargne pas l’informatique). Ce paragraphe propose de commencer par une approche assez globale. Dans la mesure où ce qui a trait à la culture informatique peut parfois rebuter, nous pouvons déjà glaner dans nos habitudes quelques reflets susceptibles d’illustrer cette modification. Une observation opérante de ces saillances devrait (dans l’idéal) nous amener à avoir ensuite envie de questionner plus avant l’importance que peut avoir l’informatique dans ces changements.

Après avoir consacré un film documentaire à une police de caractères, *Helvetica* (2007), Gary Hustwit s’est plus récemment intéressé aux acteurs du design contemporain. Le documentaire *Objectified* (2009) se présente ainsi comme un objet filmique tourné aussi bien vers l’impact esthétique de la création industrialisée sur nos vies quotidiennes, que vers les sources de ces choix esthétiques. Le long de ses soixante-quinze minutes, *Objectified* met en scène un glissement culturel : s’il débute par une tendance moderniste orientée vers la jouissance immédiate et la distinction sociale apportées par le design « dernier cri », c’est pour mieux se rapprocher ensuite d’enjeux plus actuels, tels la pérennisation du jetable, le respect vis-à-vis des objets ou encore le développement durable — le mot d’ordre de la contemporanéité n’est plus « je prends, je jette ». Et le film s’achève sur les paroles de Rob Walker, auteur et chroniqueur pour le *New York Times Magazine* :

« L’ouragan arrive, vous avez vingt minutes, prenez vos affaires et partez. Vous n’allez pas dire, “Alors, ça, ça a reçu une super critique sur tel blog de design.” Vous allez prendre les objets les plus significatifs pour vous. Car ceux-là sont de vrais objets, ceux qui reflètent vraiment l’histoire vraie de qui vous êtes, votre récit personnel, l’histoire que vous vous racontez à vous-même et à personne d’autre, parce que c’est le seul public qui importe⁵⁷. » [Hustwit, 2009, 1.13:40]

Soit des paroles mettant l’accent sur les choix émotionnels que nous pouvons produire dans

⁵⁷ « The hurricane is coming, you have twenty minutes, get your stuff and go. You're not going to be saying, “Well that got an amazing write-up in this design blog.” You're going to pick the most meaningful objects to you, because those are the true objects, that truly reflect, the true story of who you are, and what your personal narrative is, and the story that you're telling to yourself and no one else because that's the only audience that matters. » [Hustwit, 2009, 1.13:40]. (Ma traduction.)

l'urgence, en tant que mécanismes permettant de trancher rapidement quant à la valeur des choses. Si nous ne choisissons pas le tabouret Philippe Starck, nous choisissons par contre les objets qui nous structurent et nous authentifient à nos propres yeux (le petit pull marine tout déchiré aux coudes, par exemple), et non aux yeux des autres (papiers d'identité ; bulletins de salaire ; papiers d'assurance locative ; relevés bancaires, etc.). Ou plutôt, ces trois occurrences (tabouret, pull, papiers) peuvent-elles se chevaucher lorsque nous attachons à chacune d'elles un « récit » émotionnel (le tabouret Starck devient l'unique souvenir d'un ami défunt, ou pour d'autre un achat coup de cœur reflet d'intenses privations ; la carte d'identité devient un objet obtenu après des années de luttes épuisantes, ou pour d'autres le reflet d'une fierté patriotique, etc.) : le point d'orgue du film porte ainsi sur le processus d'investissement affectif, lorsqu'il se dissocie de l'investissement financier. Dans ce même film, un autre designer, David Kelley, fondateur de l'agence de design IDEO, explique :

« Si vous concevez quelque chose de précieux et que vous aimez vraiment, vous ne le laisserez jamais. La mallette de mon père, faite à partir d'une très belle pièce de cuir, s'améliore avec l'usage. J'en ai hérité, et je la transmettrai. C'est une chose vraiment intéressante. Parfois je reçois la tâche de concevoir quelque chose qui s'améliore avec l'usage. Il y a très peu de choses, la majorité se dégrade, mais... certaines choses, comme cette mallette, s'améliorent avec l'usage⁵⁸. » [Hustwit, 2009, 0.47:20]

Là encore, la durabilité présentée comme allant de paire avec un investissement et une plus-value émotionnels, passe par le renforcement de la valeur d'usage, voire une patrimonialisation de la valeur d'échange, et les objets industriels sont avant tout appelés à s'inscrire dans une histoire, une vie commune.

Les nouvelles technologies ont longtemps été associées à des idées de vitesse, de *turn-over* rapide, d'obsolescence — transmettre son premier ordinateur (ou téléphone portable) sous forme d'héritage n'est pas encore entré dans les mœurs. Il en découlait qu'aussitôt acquises elles étaient déjà périmées. Et pour rester dans la compétition il fallait savoir s'actualiser rapidement, se tenir au goût du jour, alerte au dernier cri : l'urgence n'était pas celle d'un ouragan, mais celle de se trouver périmé, éjecté de l'œil du cyclone de l'innovation, comme pétrifié de ne plus

⁵⁸ « If you design something that is precious and that you really love, you're never going to leave that. My father's briefcase, made out of a beautiful piece of leather, gets better with use. And I've inherited it, and I'll pass it on, right? It's a really interesting thing. Sometimes I get that task which is: design something that gets better with use. There's very few things, they mostly degrade, but... some things like this briefcase get better with use. » [Hustwit, 0.47:20]. (Ma traduction.)

être maintenu en vie par la mode. Les années 1990 ont peut-être marqué la fin d'un tel engouement pour la nouveauté et l'actualisation consumériste (en Europe). Les contextes économiques et écologiques actuels ne se prêtent peut-être plus à autant d'enthousiasme industriel. Et l'image de fraîcheur optimiste véhiculée par le renouvellement perpétuel semble s'être fanée en désuétude immature : celle de chercher à s'afficher en *fashion victim* tandis que les préoccupations actuelles tourne autour des réserves en eau potable (ou « potabilisable »), par exemple. Ou pour le dire autrement, si la course à l'innovation n'a pas connu de collapse notable et sait encore mobiliser les foules, elle génère dans le même temps des réactions de rejet et de défiance, de la part de ces « mêmes » foules, des réactions à la hauteur des efforts publicitaires pour la soutenir — la question ici n'est pas de trancher de façon manichéenne pour savoir qui a raison ou tort du consumérisme insouciant ou de la conservation patrimoniale.

La course à l'information pourrait contredire cette manière de présenter les choses, pourtant, sous la forme du gavage, elle n'empêche pas une forme d'émiettement généralisé quant à l'intérêt porté aux actualités (aux *news*). Du moins la notion « d'économie de l'attention » laisse-t-elle une place pour des réflexions concernant un gavage et une obésité informationnels :

« Les médias, comme la nourriture que nous consommons, fournissent différents goûts pour différentes dispositions émotionnelles [*moods*]. Mais considéreriez-vous que votre alimentation médiatique est équilibrée? Recevez-vous la dose quotidienne recommandée de nutriments médiatiques? Selon une étude de 2009, de l'Université de San Diego, les Américains consomment en moyenne 12 heures de médias par jour. [...] [Clay A.] Johnson est l'auteur de *The Information Diet*. Il explique que les Américains ne sont pas seulement en train de prendre du poids autour de la taille : nous devenons obèses... Obèses d'informations⁵⁹. »

Cet extrait provient d'un article posté (relayé) par le MIT Media Lab, sur sa page Facebook, le 14 mars 2012. On peut y lire plus bas :

« De nos jours nous avons de toute évidence largement plus d'informations que ce que nous pourrions

⁵⁹ « The media, like the food we consume, provides different tastes for different moods. But would you consider your media diet a balanced one? Are you getting the recommended daily dose of media nutrients? According to a 2009 study from the University of California San Diego, Americans consume an average of 12 hours of media a day. [...] [Clay A.] Johnson is the author of "The Information Diet." He said Americans are not just getting heavy around the waist, we're growing obese ... information obese. » Cristina Quinn, « A Nutrition Label for the News », article paru sur le site de la télévision américaine WGBH, le 14 mars 2012. <http://www.wgbh.org/articles/A-Nutrition-Label-for-the-News-5771>. (Ma traduction.)

jamais en lire, et dorénavant le problème est de trouver l'information de qualité dans ce monde où sont mélangés le divertissement complet, la désinformation et la bonne information", explique Matt Stempeck, assistant de recherche au Center for Civic Media du MIT Media Lab. "J'aime concevoir des outils qui aident les gens à trouver l'aiguille dans une botte de foin, de façon pratique." Stempeck et son équipe du Media Lab sont en train de construire le Media Meter, qui est, explique-t-il, "une plateforme ouverte pour aider les gens à suivre puis visualiser ce à quoi ressemble leur alimentation médiatique"; l'idée étant que le Media Meter vous montre ce que vous avez dans votre consommation médiatique quotidienne, et aussi ce qui manque dans votre régime médiatique. Un peu comme... une étiquette nutritionnelle⁶⁰.»

Soit un propos en cohérence avec les recherches de Picard en matière «d'informatique affective», même si formulés de façon plus raide (stigmatisation de l'obésité, culpabilisation, «pathologisation» du rapport à l'information — dans le sens de *pathos* émotionnel, mais aussi de *pathos* maladif). En 1997 Picard notait, «L'un des problèmes que l'âge de l'information a amené est celui du trop d'informations, qui tend à mener vers une fatigue cognitive et vers une réduction de l'aptitude à traiter précisément les nouvelles données entrantes⁶¹.» [Picard, 2000, p. 56], avant de poursuivre plus loin :

«L'un des domaines dans lesquels les émotions des ordinateurs sont d'un intérêt primordial est celui des agents logiciels, des programmes informatiques qui sont personnalisés — ils connaissent les centres d'intérêt, les habitudes et les préférences de l'utilisateur — et qui prennent activement part au fait d'assister l'utilisateur quant à la surcharge de travail et d'information. Un agent pourrait agir comme assistant de bureau, pour vous aider à traiter le courrier ; un autre pourrait prendre la forme d'une créature animée pour jouer avec un enfant. La notion d'agent soulève quelques attentes chez l'utilisateur humain. En particulier, comment des agents peuvent-ils être fabriqués pour être personnalisés, intelligents, crédibles et attachants? "Leur donner des émotions" n'est pas la seule solution à ces problèmes, mais c'en est un élément critique. L'assistant qui ne peut pas lire vos expressions faciales, ni raisonner à propos de ce que vos émotions pourraient être, ni apprendre ce qui est important pour vous — quand ne pas interrompre ce que vous faites, par exemple — agira de façon inintelligente⁶².» [Picard, 2000, pp. 193-194]

⁶⁰ « "These days we obviously have way more information than we can ever read, and now the problem is finding the quality information in this world of complete entertainment, misinformation and good information all mixed together," said Matt Stempeck, a research assistant at the Center for Civic Media at the MIT Media Lab. "I like designing tools that help people find that needle in the haystack in a convenient way." Stempeck and his team at the Media Lab are building the Media Meter, which is, he said, "an open platform to help people track and then visualize what their media diets look like." The idea is that the Media Meter will show you what you're getting in your daily media consumption and also what you're missing in your media diet. Kind of like ... a nutrition label.» Cf. note précédente. (Ma traduction.)

⁶¹ « One problem the information age has brought is that of too much information, which tends to lead to cognitive fatigue and a reduced ability to accurately process new inputs.» [Picard, 2000, p. 56]. (Ma traduction.)

⁶² « One of the areas in which computer emotions are of primary interest is software agents, computer programs that are personalized—they know the user's interests, habits and preferences—and that take an active role in assisting the user with work and information overload. One agent may act like an office assistant to help you process mail; another may take the form of an animated creature to play with a child. The notion of an agent raises several expectations from the human user. In particular, how can agents be made to be personalized, intelligent, believable, and engaging? "Give them emotions" is not the entire solution to these problems, but it is a critical component. The assistant that cannot read your emotional expression, reason about what your emotions might be, and learn what is important to you—when not to interrupt, for example—will act unintelligently.» [Picard, 2000, pp. 193-194]. (Ma traduction.)

On voit se dessiner le lien entre l'approche émotionnelle relative au fonctionnement d'un utilisateur humain, et l'approche émotionnelle relative au fonctionnement d'un programme. On voit surtout en quoi à l'obsolescence programmée se substitue l'apprentissage, soit un rapport dans lequel les objets industriels ne sont plus seulement une question de distinction sociale, de consommation effrénée, mais plutôt une affaire de transmission, sous la forme de la passation de son propre « récit » personnel à des entités intelligentes, capables de percevoir et réfléchir à / sur nos émotions, nos subjectivités, nos rapports aux lois générales. Le problème n'est plus de prendre et de jeter la dernière occurrence à la mode, mais d'éduquer des objets, de leur apprendre nos façons privilégiées d'interagir avec l'environnement, et ce dynamiquement — en fonction du contexte (l'heure, la saison, etc.), de notre situation (avec qui nous sommes, la qualité de notre sommeil durant la nuit précédente, notre situation professionnelle et affective, ce que nous allons faire dans une heure, dans six mois, etc.), de nos préférences du moment. Soit un investissement durable de la part du consommateur, se détachant du modèle de l'investissement unique et massif que représente l'achat seul. L'acquisition financière (plus ou moins compulsive) se voit alors doublée d'un nouveau *challenge* : celui d'un investissement affectif quotidien, à construire. La valeur de l'objet technique ne se réduit plus à sa seule performance objective, mais s'étoffe au contact de son possesseur pour devenir unique voire irremplaçable. Bien entendu, nous n'en sommes pas encore complètement là, mais on peut néanmoins noter qu'entre la nouvelle d'Asimov « Robbie » (1940)⁶³ et les agents conversationnels (ou encore l'assistante personnelle Siri d'Apple, sur l'iPhone 4S, commercialisé en octobre 2011 avec le succès qu'on lui connaît), les choses se sont actualisées.

Cette transformation du rapport à la valeur des choses modifie la course massive vers le « nouveau ». Comme explicité plus haut, l'éloge de l'accélération se retrouvait dans l'énoncé du test de Turing (1950), lorsque pris au pied de la lettre : soit l'idée d'accélérer des processus jugés trop lents ; l'idée de remplacer la sélection naturelle par une sélection culturelle encore plus expéditive et drastique que la première. Et en un sens, cet éloge de « l'excellence » aura participé de son inévitable revers : un éloge du gaspillage. Mais, comme noté plus haut, l'énoncé de Turing ne visait peut-être pas qu'à instituer un observateur extérieur en position

⁶³ Cf. note #20.

d'autorité. Lorsque le programme devient son propre évaluateur et son propre constructeur de méthodes, alors le processus éliminatoire s'atténue dans le champ de conscience de l'observateur. Cette «autonomisation» revient à admettre qu'un petits lots de lois strictes *a priori* ne puisse pas régir intégralement les processus de choix d'un programme, et revient à accepter que ce dernier devienne une position par rapport à des lois (qu'il manifeste une subjectivité interprétationnelle par l'effectuation de choix individués et appropriés).

Bien que le livre de Picard (*Affective Computing*) ne soit pas consacré aux relations humains / ordinateurs, elle y rapporte néanmoins un résultat expérimental selon lequel les gens tendent à traiter les ordinateurs (les programmes capables de communiquer en langage naturel) comme des êtres humains. Via l'utilisation des mêmes codes socio-comportementaux acquis, notamment. Par exemple, s'il est question de juger des performances de quelqu'un, alors nous sommes censés adoucir nos jugements lorsque nous les formulons face-à-face et, donc, censés les durcir lorsque nous les formulons face à une tierce personne, perçue comme neutre. Picard note ainsi que :

«L'humain A donne à [l'ordinateur] B son évaluation "face-à-face", et ensuite est interrogé par un autre ordinateur quant à son évaluation de la performance de B. Les résultats humain-humain classiques restent valables, par exemple la tendance à être plus gentil face-à-face demeure. De nombreuses expériences similaires ont été réalisées par Reeves et Nass, révélant que les résultats classiques des études humain-humain sont maintenus dans les études humain-ordinateur. Ces conclusions restent valables, même pour des gens en "sachant plus", tels que les étudiants en informatique qui savent que les ordinateurs n'ont pas d'émotions⁶⁴.» [Picard, 2000, 14].

Ces propos soulèvent implicitement un problème : celui de l'hypocrisie. On connaît le lien entre cette dernière et le jeu d'acteur (du grec *υποκρισις*, «rôle tenu sur scène», «mimique») — ce qui nous ramène sans trop de difficulté aux agents conversationnels et à leurs composantes d'imitation et de simulation —, mais tout aussi intéressante est la traduction du mot en anglais : *hypocritical* (c'est-à-dire, pour user d'un néologisme éloquent, «hypocritique»). Si la dimension affective connaît un certain essor dans les domaines de l'informatique et des objets intelligents, ce n'est peut-être pas pour nous livrer des objets

⁶⁴ « Human A gives B its evaluation "face-to-face", and then is asked by a different computer for an evaluation of how B did. The classic human-human results still hold, for example the tendency to be nicer face-to-face remains. Numerous similar experiments were done by Reeves and Nass, revealing that the classic results of human-human studies are maintained in human-computer studies. The findings hold true even for people who "know better", such as computer science students who know that computers don't have emotions. » [Picard, 2000, 14]. (Ma traduction.)

désagréables ou impartiaux (ou critiques). Si les objets deviennent sujets, peut-être est-ce aussi pour mieux faire cour autour de leurs possesseurs. À l'instar d'un miroir de contes de fées (augmenté) qui aurait appris à répondre toujours par l'affirmative à la question « miroir, mon beau miroir, suis-je le plus beau — intelligent, malin, gentil, agréable, séduisant, etc. ? ». Non pas via une flatterie littérale (risquant d'éveiller la suspicion ou d'être interprétée comme du sarcasme), mais via une flatterie futée (et l'imitation intelligente fait partie de ce répertoire). Le trait est forcé bien entendu, cependant c'est pour mieux insister sur le fait que la question de l'attachement émotionnel aux objets passe souvent par un implicite, celui d'un attachement jugé positif.

Dans une autre nouvelle, plus célèbre que celle du « Village enchanté » précédemment citée, se produit de même une transformation : *La Métamorphose* (1912) de Franz Kafka. Cette fois le narrateur s'objectifie peu à peu. Même si sa mutation lui confère une nouvelle et réelle subjectivité, peut-être plus forte encore que l'ancienne, il voit progressivement son statut de sujet devenir celui d'objet, jusqu'à ne plus être conscient — accomplissant ainsi sa pleine métamorphose en chose (cadavre d'insecte). On ne peut pas dire que ces deux nouvelles soient l'exact opposé l'une de l'autre — où le « Village enchanté » serait une trajectoire allant de la mort annoncée vers une résurrection, tandis que *La Métamorphose* irait de vie à trépas, par le biais d'une désadaptation. Et ce dans la mesure où l'entourage de Gregor Samsa revit à mesure que lui-même se pétrifie. Si cette transformation le plonge dans une forme de désert, pour son entourage l'expérience s'assimile à celle du « Village enchanté » : progressivement leur vacuité se remplit, jusqu'à ce qu'ils réintègrent le monde qu'ils avaient déserté, pour mieux s'y réadapter. La famille reprend vie à la faveur de sa propre mutation (celle de son membre moteur, Gregor). Maintenant si un ouragan arrivait, en dépit des liens très forts unissant Gregor et son entourage, serait-il l'objet protégé et emporté ? Autrement dit, quand bien même Gregor serait le vrai objet définitoire, la signature de cette famille et de chacun de ses membres, quand bien même Gregor bénéficierait d'un attachement affectif très fort (sans bénéficier du détachement de l'accepter tel qu'il est) : sitôt crevé (« *krepiert* »), sitôt jeté. Soit l'idée que l'attachement (les liens) ne suffise pas en tant que tel, mais qu'il doive s'accompagner d'une plus-value personnelle positive.

L'art de l'attachement positivement qualifié nous ramène ainsi à la rhétorique, c'est-à-dire aux méthodes permettant de s'assurer la bienveillance (ou l'accord) de ses interlocuteurs. En termes d'informatique, la marque Apple (également citée dans *Objectified*) est reconnue pour son impact affectif sur ses adeptes, nonobstant les conditions dans lesquelles ses objets sont fabriqués — Foxconn interdisant contractuellement à ses employés chinois de se suicider, notamment⁶⁵, afin de ne pas encourager les suicides économiques (se tuer sur son lieu de travail pour assurer une rente à sa famille) —, ou encore nonobstant les données privées collectées par la marque⁶⁶. La ferveur que suscitent ses produits ne semble pas se refroidir au contact de telles informations. Par contre cela met en lumière l'exacerbation du couple confiance / méfiance.

Nous sommes face à des objets intelligents conçus pour apprendre à nous connaître, de façon à pouvoir nous séduire, et de façon à nous maintenir dans cet état de séduction, plus durablement. La contrepartie étant que sous des dehors charmeurs il y ait collecte de données et profilage des utilisateurs, dans la mesure où pour s'adapter à chacun d'eux, individuellement, ces objets doivent apprendre et mémoriser l'utilisateur. Si l'on devient subjectivement le possesseur de tels objets, ceux-ci restent néanmoins en communication avec, voire sous le contrôle de, la société qui les produit — quasi-condition de leur fonctionnement. Et la cour de ces objets peut tout aussi bien se métamorphoser en cour de jugement, avec des processus en forme de procès perpétuel où l'utilisateur est évalué, soupesé, jaugé, ausculté sous tous ses angles, afin d'en tirer quelques informations pertinentes. Comme nous le verrons plus loin, ce

⁶⁵ Foxconn est une société taiwanaise de sous-traitance électronique, comptant parmi ses clients les géants Apple, Amazon, Dell, Intel, Microsoft, Nokia, etc. Cf. Philippe Grangereau, « Suicides à la chaîne chez le géant Foxconn », dans *Libération*, 3 juin 2010. <http://www.liberation.fr/economie/0101639157-suicides-a-la-chaine-chez-le-geant-foxconn>. Ou encore : Daily Mail reporter, « You are NOT allowed to commit suicide: Workers in Chinese iPad factories forced to sign pledges » dans *Daily Mail Online*, 1 mai 2011. <http://www.dailymail.co.uk/news/article-1382396/Workers-Chinese-Apple-factories-forced-sign-pledges-commit-suicide.html>. Nul besoin de faire de l'angélisme cependant : Apple n'a pas le monopole de l'exploitation humaine. Pour un article plus nuancé, mais non moins édifiant quant à l'usage contemporain de la main d'œuvre par les géants économiques, et quant aux réactions résignées que cela suscite chez les consommateurs / employés (qui à juste titre se disent sûrement que cela vaut toujours mieux que le chômage), cf. Joel Johnson, « 1 Million Workers. 90 Million iPhones. 17 Suicides. Who's to Blame? », dans *Wired*, mars 2011. http://www.wired.com/magazine/2011/02/ff_joelinchina/all/1. Et pour un aperçu des logements proposés par la société, cf. Chris Chang, « Foxconn Workers See A Quarter Wiped Off Their Salary As Dormitory Rents Increase », sur *M.I.C. gadget*, 20 mars 2012. <http://micgadget.com/23937/foxconn-workers-see-a-quarter-wiped-off-their-monthly-salary-as-dormitory-rents-increase/>. C'est toujours mieux que rien ? Absolument : c'est une réelle joie (un vrai rêve) que de travailler dans de telles conditions pour aider une entreprise à faire plus de profits en vendant des produits à des clients soucieux de préserver la culture du bien-être et du confort (c'est important, nous en convenons).

⁶⁶ Fait sur lequel elle aura néanmoins été sommée de s'expliquer, cf. le communiqué de presse « Apple Q&A on Location Data », 27 avril 2011. <http://www.apple.com/pr/library/2011/04/27Apple-Q-A-on-Location-Data.html>.

n'est *pas* cette philosophie qui sous-tend la conception de telles entités digitales. Pourtant c'est l'un des scénarii de devenir plausibles pour ces entités digitales, dans une société narcissique⁶⁷ où la survie est conditionnée par le fait de plaire au Prince et où il n'est d'autres choix que de courtiser «hypocritiquement», tout en accentuant la surveillance objective en vue d'améliorer les performances à venir et d'échapper ainsi au recyclage par démantèlement.

Face à ce processus de *feedback* quasi-amoureux de la part de nos environnements intelligents, la «transparence» de nos sociétés peut vite s'avérer n'être qu'une galerie de miroirs déformants. À mesure que l'être humain se dévoile, dans toute son abstraction universalisable (sociale, comportementale, cognitive, émotionnelle, physiologique, génétique, épigénétique...), l'individu se voit de plus en plus contraint à la pratique d'un *self-control* scrupuleux, systématiquement confronté qu'il est à des représentations de lui-même. En un sens, ce *side effect* se retrouve déjà dans nos villes de verre, censées incarner le contraire de l'hermétisme bétonné des bunkers, censées symboliser la circulation, mais se résolvant elles aussi, parfois, en un autre type d'enfermement : le renvoi constant aux représentations de «soi-même», l'impossibilité d'oublier ces caricatures, l'obligation de vivre la relation aux autres de l'extérieur — en étant plus conscient et confiant vis-à-vis de ces représentations récurrentes que vis-à-vis des informations sensorielles fournies par l'échange spatiotemporellement situé, de façon définitivement unique et non récurrente. La réalité s'apparente alors à une forme de perpétuelle «violation de la parité» (brisure de la symétrie), dans la mesure où nous apprenons à nous comprendre et à nous situer à la faveur de reflets biaisés, hautement réalistes mais impossibles. Pourtant nous nous accrochons encore à la symétrie et tenons ce qui ne l'est pas pour tel. Les espaces se superposent, se prolongent effectivement, mais en mondes parallèles impénétrables avec lesquels les contacts ne sont jamais que de surface, d'apparat. La confiance que nous accordons aux reflets ne nous éloigne pas du ciel des idées, mais de toute synthèse du divers effectuable à partir de nos propres sensations. Car comme le constatent les enfants avec étonnement (moins les adultes) : il y a une distance infranchissable entre ce que l'on vit de l'intérieur et les représentations extérieures qui nous en sont restituées. Peut-être que pour être

⁶⁷ Comme le fait à juste titre remarquer le musicien Moby dans le film documentaire *PressPausePlay*, en 2006 le *Time Magazine* a consacré son emblématique numéro spécial «*Person of the Year*» à «*You*» (Vous / Toi) [cf. Dworsky & Köhler, 2010].

efficace et performant il ne faudrait tenir compte que de ces dernières. Bien sûr, cet aspect des reflets biaisés n'est pas nouveau : il a notamment fait l'objet de nombre des travaux de l'artiste américain Dan Graham, dans les années 1970 [cf. Perelman, 1995]. Si changement il y a, c'est alors dans le fait que ces reflets ne soient pas seulement rétiniens, mais encore émotionnels (et donc cognitifs). Ils portent sur nos comportements visibles, certes, mais aussi sur nos habitudes, nos stratégies communicationnelles et sociales, nos manières de nous projeter dans (et de nous construire des) avenir, nos manières d'aimer, d'esquiver, de mémoriser, de rêver, d'avoir peur, de souffrir, de construire des pensées politiques. Et si transparence il y a, peut-être n'est-elle pas tant dans l'environnement qu'en nous-mêmes : les objets et les entités ambiants ne sont pas rendus moins opaques, mais on peut se surprendre à ne plus se surprendre soi-même (surprise vite passée), soit s'apercevoir que pour chaque question non encore posée, il y a déjà pléthore de réponses toutes faites, prêtes à l'emploi, *ready-made*, quel que soit celui qui les pose. Une transparence probabiliste.

Il y a donc quelque chose de rassurant, en un sens, à constater que cette promesse d'un accès au bonheur par la mise en place d'environnements pleinement compréhensifs et attentifs à soi ne puisse qu'être immanquablement trahie. Tous comme les humains, les programmes ont une individualité, leurs jugements sont biaisés, ils ne comprennent de leurs utilisateurs que ce qu'ils peuvent bien en comprendre et leurs évaluations ne portent que sur une poignée de critères qui ne pourra jamais atteindre à l'exhaustivité. Et c'est ainsi qu'on sort du reflet pour percevoir enfin la subjectivité de l'autre. Mais quel autre ?

Le paragraphe qui précède aura tenté de percevoir l'arrivée des émotions dans le champ de l'informatique à travers les effets socioculturels secondaires, tout en introduisant l'informatique affective, non en la définissant mais en plantant un décor. L'insistance aura été portée sur le fait que l'introduction d'un lien affectif soit le plus souvent méliorativement connotée, d'autant plus s'il s'agit d'objets. Ce point peut avoir une grande influence sur le développement de ces mêmes objets, sans pour autant signifier une absence de duplicité. Si l'ajout d'une dimension affective et émotionnelle tend à être présenté comme allant naturellement vers la gentillesse et la bonté, il est aussi tentant de se souvenir que nombre de

nos rapports sociaux tiennent grâce à l'hypocrisie (qu'elle soit de politesse et de savoir vivre ou de manipulation). Peut-être est-ce un biais culturel que de s'attacher à ce point. Il est possible aussi que cette façon de voir (et de critiquer) relève d'une approche (ou d'un travers) typiquement française.

AFFECTIVE COMPUTING (INFORMATIQUE AFFECTIVE)

Mais quittons ces réflexions spéculatives pour nous pencher sur plus « concret » : intéressons-nous aux travaux de Rosalind W. Picard quant à l'informatique affective. Si ses ouvrages ne sont pas traduits en français, on y trouve sûrement l'un des éclairages les plus fondamentaux et fondateurs quant à nos environnements techniques actuels, ou tout au moins quant aux agents conversationnels faisant l'objet de la présente recherche. Jusqu'ici il aura été question du test de Turing, mais si nous voulons nous rapprocher au plus près des agents en ligne, nous devrions parler du « test de Picard ». Les deux paragraphes suivants vont donc s'intéresser à ces deux questions : la définition de l'informatique affective et les contours de ce que nous pourrions appeler le « test de Picard ».

Voyons comment Rosalind W. Picard définit l'informatique affective :

« informatique [*computing*] faisant référence à, résultant de, ou influençant délibérément les émotions. [...] L'informatique affective inclut le fait d'implémenter des émotions, et peut donc aider au développement et au test d'anciennes et nouvelles théories des émotions. Cependant, l'informatique affective inclut également beaucoup d'autres choses, telles que le fait de donner à un ordinateur des aptitudes quant à la reconnaissance et l'expression d'émotions, de développer son aptitude à répondre intelligemment à l'émotion humaine, et de lui permettre de réguler et d'utiliser ses émotions⁶⁸. » [Picard, 2000, p. 3]

Pourquoi cet intérêt porté aux émotions dans le champ de l'informatique ? En un sens, on peut dire que le MIT Media Lab est un département de recherche dédié à cet « élément aléatoire » évoqué par Turing [Turing, 1950], mais également évoqué dans la première feuille de route

⁶⁸ « computing that relates to, arises from, or deliberately influences emotions. [...] Affective computing includes implementing emotions, and therefore can aid the development and testing of new and old emotions theories. However, affective computing also includes many other things, such as giving a computer the ability to recognize and express emotions, developing its ability to respond intelligently to human emotion, and enabling it to regulate and utilize its emotions. » [Picard, 2000, p. 3]. (Ma traduction.)

mentionnant officiellement le terme «intelligence artificielle⁶⁹», soit la présentation d'une conférence au Dartmouth College, en 1955, par John McCarthy, Marvin Minsky, Nathaniel Rochester et Claude Shannon⁷⁰. Dans cet sorte de manifeste (plutôt feuille de route), mention est faite d'une hypothèse quant à ce qui pourrait différencier la pensée compétente mais dénuée d'imagination, de la pensée créative : une forme d'aléatoire (*randomness*) particulier, guidé par l'intuition. Le MIT Media Lab, en tant que partie du MIT School of Architecture and Planning, a été fondé en 1985 et compte parmi ses membres Marvin Minsky, notamment, auteur de *The Emotion Machine*, en 2006. Ce département de création (et d'art) produit un certain nombre d'œuvres et dans l'absolu, pour la présente recherche et pour être à même de parler des «œuvres d'art» (reconnues comme telles), il aurait fallu pouvoir accéder à ces travaux. Ceux-ci circulent dans les réseaux habituels en Europe (Ars Electronica à Linz, en Autriche, ou encore à la Biennale de Venise, en Italie) en plus d'être exposés dans certains centres d'art états-uniens. Soit des réseaux demandant un certain investissement financier de la part de leurs visiteurs exogènes. Mais cet obstacle ne doit pas pour autant bloquer toute réflexion, toute expérimentation. Au contraire, le manque d'opportunités en génère une nouvelle : celle d'accéder aux formes gratuites de ces travaux. C'est-à-dire la possibilité de faire l'expérience de ces recherches lorsqu'elles passent dans la culture commune (sous forme

⁶⁹ Herbert A. Simon s'est exprimé sur le fait qu'il ait pu trouver ce terme inadéquat, dans une note de *Les sciences de l'artificiel* : «Je récusé toute responsabilité quant au choix particulier de ce terme. La formule "intelligence artificielle" qui m'y a conduit [à l'opposition du terme "artificiel" à celui de "naturel"] a été forgée, je crois, au bord de la "Charles River", au MIT. Notre propre groupe de recherche, à la RAND et à l'Université Carnegie-Mellon, préférait des formules telles que "traitement complexe des informations" ou "simulation des processus cognitifs". Mais nous rencontrons là aussi de nouvelles difficultés de terminologies, car pour le dictionnaire "simuler" signifie "présumer avoir, ou même avoir, la même apparence ou forme, sans avoir la réalité ; imiter ; contrefaire ; feindre". De toute façon, l'expression "intelligence artificielle" semble devoir se maintenir, et elle peut se révéler, une fois nettoyée, plus aisée pour clarifier les communications que l'invention d'une autre expression. Le temps viendra où elle deviendra suffisamment idiomatique pour n'être plus l'objet de rhétorique de pacotille. [N.d.T. : Cette note fut rédigée pour la première édition de ce chapitre, en 1969. H. A. Simon l'a maintenue inchangée jusque dans la 3^e édition en 1996. Le temps, en effet, est venu : aujourd'hui, l'IA est devenue idiomatique.]» [Simon, 2004, note 4, pp. 391-392]. S'il n'est pas aisé de savoir ce que Simon qualifie de «rhétorique de pacotille» (*a contrario* d'une Rhétorique pure?), on peut noter que le terme aura néanmoins influencé positivement son travail puisque l'ouvrage dont est tirée cette note propose de fonder les «sciences de l'artificiel» (sciences de la conception, des processus de conception), non pour remplacer les «sciences naturelles», mais pour les compléter. On y retrouve donc des catégories assez classiques. Par ailleurs, on peut également noter l'embarras de Simon quant au terme «simulation» (le terme «émulation» n'en est pas moins embarrassant). Enfin, l'optimisme du traducteur, Jean-Louis le Moigne, est peut-être à nuancer : si l'IA est effectivement idiomatique, c'est surtout par rapport à la culture science-fictionnelle. Qualifier cette culture de rhétorique de pacotille serait certainement excessif, en revanche on peut peut-être déplorer la tendance culturelle à considérer l'IA comme un simple divertissement, ou plutôt, comme un divertissement simple.

⁷⁰ Cf. John McCarthy, Marvin L. Minsky, Nathaniel Rochester, Claude E. Shannon, «A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence» (31 août 1955), partiellement reproduit dans *AI Magazine*, vol. 27, n°4, 2006, pp. 12-14.

d'agents conversationnels en ligne, par exemple). Au-delà de la boutade, ce n'est pas si absurde : si l'expérience de l'art classique requiert en quelque sorte que l'on s'attache aux originaux des maîtres, où qu'ils soient et sous quelque condition d'accès qu'ils soient conservés, l'art digital affectif se réclame d'une expérience différente : 1) son but est de sortir du laboratoire / du musée, afin de se confronter aux conditions d'expérimentation de la « vie réelle » (dans le sens de vie quotidienne), 2) sa valeur vient aussi de sa capacité à apprendre et à vieillir au contact prolongé des utilisateurs, les plus divers possibles. Peut-être parce que ces dispositifs ne cherchent plus à s'accomplir au sein de relations *hiérarchiques* (sacrées) mais plutôt dans une connivence de « longue durée ». Comme le note Picard, « La métaphore n'est pas que celle d'une informatique personnelle ou amicale, mais celle d'un ordinateur qui puisse *servir* [italique ajouté] comme une sorte d'ami personnel. Au lieu d'être oublieux de vos sentiments, le système y prêterait attention et les respecterait⁷¹. » [Picard, 2000, p. 245]. La relation visée est celle d'une amitié, d'une intimité frôlant l'idée de « meilleur ami », avec toutes les colorations que peut prendre le terme dans le cours d'une relation — de conseiller du Prince à animal de compagnie ; d'amoureux platonique à espion séducteur ; de confident tragique à psychologue, de relation Facebook à personne présente dans les moments éprouvants, etc.).

Mais revenons à la définition de l'informatique affective, car elle ne s'arrête pas à ce qui a été cité plus haut. Elle s'enrichit d'autres points cruciaux :

« Est-ce que les machines peuvent ressentir [*Can machines feel*] ? Cette question pourrait bien être la plus profonde au sein du thème de l'informatique affective. Les sentiments [*feelings*] sont souvent considérés comme étant ce qui sépare l'humain de la machine. [...] Il est possible que nous découvriions qu'il est simple d'imiter la fonctionnalité de ce sentiment [subjectif interne], et que cela suffise pour donner aux ordinateurs un ensemble complet d'émotions ainsi que leurs influences. Ou alors la distinction entre ce que les humains ressentent et ce que les ordinateurs ressentent pourrait devenir un ingrédient essentiel distinguant les humains des machines, bloquant même peut-être le fait que des machines développent des émotions influençant les fonctions de l'esprit et du corps comme le font les émotions humaines⁷². » [Picard, 2000, pp. 60 & 75]

⁷¹ « The metaphor is not just one of personal or friendly computing, but of a computer that can *serve* [italique ajouté] as a kind of personal friend. Instead of being oblivious to your feelings, the system would pay attention to them, and respect them. » [Picard, 2000, p. 245]. (Ma traduction.)

⁷² « Can machines feel? This question may be the most profound one within the topic of affective computing. Feelings are often considered to be that which separates human from machine. [...] We may find that it will be simple to imitate the functionality of this [internal subjective] feeling, and that this is sufficient for giving computers a full set of emotions and their influences. Or the distinction between what humans feel and what computers feel may become an essential ingredient that

«Le problème le plus important n'est peut-être pas de savoir quelles émotions une machine aura, mais plutôt de savoir de quels comportements émotionnels elle sera capable, et comment elle les choisira et les régulera⁷³.» [Picard, 2000, p. 126]

Cette manière qu'a Picard de demander «*Can machines feel?*» ne peut pas manquer de rappeler Turing et le début de son énoncé d'un nouveau protocole de test, avec le désamorçage de la question «*Can machines think?*».

Toujours dans les grandes lignes, l'informatique affective est également guidée par des questions :

«Les ordinateurs et les gens reconnaîtraient facilement les émotions si elles étaient toujours manifestées de façon constante. Cependant ce n'est pas le cas. La plupart des émotions ne correspondent pas tout le temps à une forme fixée de modulation sentique⁷⁴.» [Picard, 2000, p. 30]

«Dans les recherches sur la reconnaissance de formes menées sur des images, les traits distinctifs que sont la texture, la couleur, la forme et le mouvement tendent à être les plus utiles. L'un des défis difficiles pour la recherche en informatique affective étant de déterminer quels traits parmi les signaux physiologiques sont les plus importants — trouver ce qui est l'équivalent de la couleur, de la texture, de la forme et du mouvement, dans l'affect⁷⁵.» [Picard, 2000, p. 187]

Soit le fait de réussir à dégager des récurrences permettant de mieux comprendre les corrélations entre l'état physiologique mesurable et le phénomène qualifié d'état émotionnel, en mettant au jour des paramètres pertinents, parmi la profusion de signaux et de modifications se produisant dans le corps humain à chaque instant. L'informatique affective tend également vers un idéal méthodologique (à grands traits, là encore) :

« Les complications rapportées ci-dessus ont particulièrement contaminé les études sur l'émotion humaine réalisées en laboratoire. Par exemple, certains sujets peuvent se sentir inhibés pour exprimer du dégoût ou

distinguishes humans from machines, and perhaps stands in the way of machines developing emotions that successfully influence mind and body functions as do human emotions.» [Picard, 2000, pp. 60 & 75]. (Ma traduction.)

⁷³ «Perhaps the most important issue is not what emotions will a machine have, but what emotional behaviors will it be capable of, and how will it choose and regulate them?» [Picard, 2000, p. 126]. (Ma traduction.)

⁷⁴ «Computers and people would have a relatively easy time recognizing emotions if they were always displayed in a consistent way. However, this is not the case. Most emotions do not map to a fixed form of sentic modulation all the time.» [Picard, 2000, p. 30]. (Ma traduction.) «*Sentic*» est un terme forgé par Manfred Clynes, il est ici utilisé en quasi-équivalent de sensation (*sensation*) et de sentiment (*sentiment*), tout en servant à mettre l'accent sur les mécanismes physiques de l'expression émotionnelle. Le terme anglais *sentic* se rapporte principalement à des phénomènes physiques mesurables [cf. Picard, 2000, p. 24]. Il ne semble pas que ce concept soit encore entré dans la langue française (peut-être du fait de l'absence de traduction des travaux de Clynes ou de Picard, par exemple), par conséquent le terme *sentique* a été choisi pour la présente traduction.

⁷⁵ «In pattern recognition research on images, features representing texture, color, shape, and motion tend to be some of the most useful. A difficult challenge for affective computing research is to determine which features of the physiological signals are most important—to find what is the equivalent of color, texture, shape, and motion in affect.» [Picard, 2000, p. 187]. (Ma traduction.)

de la tristesse durant une étude en laboratoire. D'autres sujets peuvent trouver forcées ces situations d'étude et exhiber un plus petit répertoire d'émotions qu'ils n'en feraient l'expérience dans leur milieu naturel. Ou alors, ils peuvent exprimer des émotions qu'ils pensent qu'ils devraient exprimer, au lieu de les laisser apparaître "naturellement". L'aptitude à exprimer des émotions est tenue pour différer selon les sujets. Les acteurs et les musiciens, pour qui le fait d'exprimer des émotions fait partie de leur profession, montrent l'une des plus grandes fluidités dans l'expression⁷⁶. » [Picard, 2000, p. 34]

« En dépit d'un grand nombre d'efforts significatifs, la théorie des émotions en est encore à son enfance. Une partie de son manque de progrès s'enracine dans la difficulté à tester de nouvelles théories. Les formes [patterns] émotionnelles des gens dépendent du contexte dans lequel elles apparaissent — et jusque là, elles ont été limitées aux réglages de laboratoires. Les chercheurs en psychologie plaisaient sur le fait que leur champ de recherche soit seulement valide pour les mâles blancs de classe moyenne supérieure, entre 18 et 22 ans, tels qu'ils se comportent dans le laboratoire de leur professeur. Bien que la plupart des expérimentations ait une "validité interne", en référence à la validité de leurs affirmations et mesures précises, dans le cadre de réglage de laboratoire contrôlés, la plupart ne peut affirmer avoir une "validité externe", c'est-à-dire, la validité des expériences de la vie réelle. Les problèmes rencontrés par les études des émotions via un réglage de laboratoire, particulièrement du fait de l'interférence des règles sociales et cognitives, sont bien documentés. L'étude idéale qui pourrait aider au développement de la théorie de l'émotion est une observation de la vie réelle, ce que beaucoup de théoriciens ont rapporté comme étant impossible (Wallbott and Scherer, 1989). Idéalement, la théorie des émotions devrait être testée lorsque les gens vaquent à leurs occupations quotidiennes⁷⁷. » [Picard, 2000, p. 241]

Soit une méthode expérimentale visant à s'éloigner des lieux protégés que représentent les laboratoires, afin de se rapprocher des conditions sous lesquelles nous faisons décidément usage de nos émotions. Si ces remarques portent sur les théories des émotions plus que sur l'informatique affective, il ne faut pas oublier que les deux sont intimement liées : l'informatique affective ne cherche pas en tant que telle à trouver *la* bonne théorie des émotions, mais elle est néanmoins assez tributaire de ces théories et si les ordinateurs devaient parvenir à émuler des aptitudes émotionnelles humaines, il serait préférable que ces aptitudes ne soient pas configurées pour ne fonctionner que dans des situations de laboratoire, face à un type de gens (ou de culture émotionnelle) spécifique, dans la mesure où ce serait sectaire déjà,

⁷⁶ « The complications noted above have particularly plagued laboratory studies of human emotion. For example, certain subjects might feel inhibited about expressing disgust or sadness during a laboratory study. Other subjects might find the situations in the study contrived, and exhibit a much smaller repertoire of emotions than they would experience in their natural world. Or, they might express emotions they think they should express, instead of letting them arise "naturally". The ability to express emotions is believed to differ among subjects. Actors and musicians, for whom expressing emotions is part of their profession, show some of the greatest fluidity in expression. » [Picard, 2000, p. 34]. (Ma traduction.)

⁷⁷ « Despite a number of significant efforts, emotion theory has remained in its infancy. Part of its lack of progress stems from the difficulty in testing new theories. People's emotional patterns depend on the context in which they are elicited—and so far these have been limited to lab settings. Research psychologists joke that their field is valid for upper-middle-class white male psychology students ages 18-22 as they behave in their professor's laboratory. Although most experiments have "internal validity," referring to the validity of their precise claims and measurements within the controlled lab setting, most cannot claim to have "external validity," that is, validity to real-life experiences. Problems with studies of emotions in a lab setting, especially with interference from cognitive social rules, are well documented. The ideal study to aid the development of emotion theory is real-life observation, which many theorists have written off as impossible (Wallbott and Scherer, 1989). Ideally, emotion theory should be tested when people are going about their daily activities. » [Picard, 2000, p. 241]. (Ma traduction.)

mais surtout inadapté pour un grand nombre de situations.

Notons également que Picard formule ces regrets dans le cadre d'un plaidoyer pour les *wearables* (ordinateurs affectifs portables, tels des habits) et l'on peut penser qu'en un sens elle aura été entendue — ou perspicace —, car en 2012, les agents conversationnels en ligne, les smartphones (portables et tactiles, entre autres), Facebook — formidable outil pour alimenter les théories émotionnelles penchant vers l'*appraisal theory* (théorie de l'évaluation) — et autres dispositifs, nous accompagnent dans nos vies quotidiennes, brouillant ainsi les frontières entre l'ingénierie et la vie (pour détourner le nom d'un des recueils de textes de l'artiste Allan Kaprow, *Essays on The Blurring of Art and Life*, 2003). Quoiqu'il soit peut-être ici plus judicieux de parler d'un floutage des frontières entre art, science, ingénierie et vie.

Picard met en avant deux arguments importants pour défendre l'informatique affective. Le premier touche au grand public et cherche plutôt à convaincre n'importe quel technosceptique des bénéfices de cette recherche :

« Le fait de donner aux ordinateurs la possibilité de mettre le doigt sur ce qui, dans ce qu'ils font, frustre le plus leurs utilisateurs, et de voir à quel point cela les frustres réellement, serait un pas important pour aider les concepteurs d'outils informatiques à apporter des améliorations⁷⁸. » [Picard, 2000, p. 248]

N'a-t-elle pas raison de nous prendre en défaut là où nous le sommes? C'est-à-dire sur le terrain de nos exigences — les développeurs d'applications ont une plaisanterie à propos de leurs clients : moins ils sont compétents, plus ils sont exigeants. Si nous refusons ou si nous n'avons pas de temps à consacrer au fonctionnement de ces outils, c'est alors à eux de faire le travail, y compris celui de nous comprendre mieux que nous ne nous comprenons nous-mêmes, et ce en plus de se connaître eux-mêmes, afin de nous simplifier la tâche. En d'autres termes, les moins férus de technologie seront paradoxalement ravis de voir les machines s'allumer et se mettre à fonctionner (ou à se réparer) toutes seules, pour ne plus avoir à subir le stress (ou l'humiliation) de ces petites pannes techniques qui nous empoisonnent la vie, occupés que nous sommes à nous soucier de tout ce qui n'a pas directement trait à la technique et qui, presque par définition, n'en est que plus appréciable. Le dispositif informatique qui

⁷⁸ « Giving computers the ability to pinpoint what they do that most frustrates users, and to see just how much it really frustrates them, would be a step toward helping designers of computer tools make improvements. » [Picard, 2000, p. 248]. (Ma traduction.)

saura détecter l'énervement (ou la frustration) de son utilisateur cliquant éperdument vingt fois de suite sur le même « bouton » sans que rien ne se passe (sinon le gel de la machine ainsi traitée), afin de lui proposer gentiment (pour ne pas l'énervier davantage) des alternatives plus adaptées et plus à même de pouvoir lui permettre d'arriver à ses fins, saura probablement s'attacher ses bonnes grâces. L'utilisateur n'aura peut-être rien appris et s'en trouvera peut-être même d'autant plus ignare quant à ce qui constitue son environnement immédiat, mais au moins il sera séduit.

Le deuxième argument (il y en a beaucoup d'autres) ne s'adresse peut-être pas au même « public » (s'il s'adresse aux mêmes lecteurs, il ne le fait pas de la même manière) :

« Plus d'une décennie s'est écoulée depuis que Marvin Minsky a écrit dans *La société de l'esprit*, "La question n'est pas de savoir si des machines intelligentes peuvent avoir des émotions, mais de savoir si des machines peuvent être intelligentes sans émotions" (Minsky, 1985) ⁷⁹. » [Picard, 2000, p. 247]

« [...] lorsqu'un système fait face à des problèmes où les possibilités ne peuvent être ni énumérées ni évaluées dans le temps imparti, je pense que la prise de décision affective fournit alors une bonne solution. Les humains utilisent les sentiments pour s'aider à naviguer sur les océans de l'investigation, pour prendre des décisions face à la complexité combinatoire. Les sentiments peuvent être appelés "intuition", ou "impression de savoir" ou juste "sentiment viscéral" ⁸⁰. » [Picard, 2000, p. 221]

« Il est possible d'argumenter en disant que la machine pourrait avoir des aptitudes affective sans que "l'émotion" ne soit impliquée de quelque manière que ce soit, sans jamais dupliquer aucun des mécanismes de l'émotion ; c'est théoriquement possible. Le fait que tous les exemples vivants d'intelligence générale s'accompagnent d'émotions ne signifie pas que l'intelligence générale ne puisse être accomplie avec quelque autre concept encore-à-découvrir — quelque mécanisme étranger. En dernière instance, la connaissance et l'expérience humaines sont finies tandis que l'univers est infini ; nous ne pouvons pas affirmer qu'une solution, telle qu'ajouter de l'émotion, soit la seule manière de fournir toutes les influences bénéfiques que les émotions fournissent, dans la mesure où nous ne sommes pas omniscients. D'un autre côté, nous n'avons pas besoin de passer du temps à chercher un mécanisme étranger tandis que nous en avons déjà un — l'émotion — semblant être capable de faire le travail. J'ai suggéré le fait qu'*émuler* [italique ajouté] sur des ordinateurs des aptitudes affectives humaines soit la solution immédiate la plus prometteuse, mais il y a encore un autre bénéfice. Si nous émuloons les émotions humaines, nous avons de meilleures chances de comprendre le comportement de l'ordinateur qui en résulte, plutôt que si nous émuloons quelques mécanismes étrangers, si et quand de tels mécanismes seront découverts ⁸¹. » [Picard, 2000, p. 249]

⁷⁹ « Over a decade has passed since Marvin Minsky wrote in *The Society of Mind*, "The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without emotions" (Minsky, 1985). » [Picard, 2000, p. 247]. (Ma traduction.)

⁸⁰ « [...] when a system faces problems where the possibilities cannot be enumerated and evaluated in the available time, I suggest that affective decision making provides a good solution. Humans use feelings to help them navigate the oceans of inquiry, to make decisions in the face of combinatorial complexity. The feelings might be called "intuition" or "a sense of knowing" or just "gut feelings". » [Picard, 2000, p. 221]. (Ma traduction.)

⁸¹ « One might argue that a machine could have affective abilities without "emotion" being involved in any way, without ever duplicating any of the mechanisms of emotion; this is theoretically possible. Just because every living example of general intelligence occurs with emotions does not mean that general intelligence might not be accomplished with some yet-to-be discovered concept—some alien mechanism. Ultimately, human knowledge and experience is finite and the universe is

Soit un argumentaire s'articulant autour du fait que la critique (et la gestion des moments critiques, ou de crise) fasse rarement l'économie des émotions. Lorsque le flot des informations et des signaux dépasse la capacité de traitement (*overload*) — situation fréquente dans nos sociétés de l'information et nos économies de la connaissance / de l'attention —, le fait de pouvoir effectuer un choix permettant de dépasser ce blocage est crucial. Ce choix peut alors être effectué aveuglement (aléatoire basique / non maîtrisé), ou de façon plus « réfléchie » (aléatoire intuitif, probabilités élaborées). Les moments de crise étant de ces moments qui accentuent l'importance que peut avoir le fait de faire de bons choix, ou plutôt le fait d'éviter d'aggraver son cas en commettant des erreurs hasardeuses, on peut comprendre qu'il soit tentant de développer des programmes capables de gérer avec souplesse et intuition ces instants de tension, dans la mesure où la quantité d'informations qu'ils traitent n'est déjà plus tout à fait à échelle humaine. En ce sens, leur donner des processus similaires aux nôtres est encore l'une des seules façons envisageables de garder un contact avec leurs méthodes et donc leurs résultats. Par ailleurs, il est cependant peu probable que de l'inhumain (au sens littéral et non au sens moral du terme) puisse sortir de la technique humaine. Quand bien même nos efforts nous porteraient à vouloir développer des mécanismes étrangers, cela reviendrait peut-être à tenter d'imaginer une nouvelle couleur. Nous pouvons alors peut-être émettre l'hypothèse que ce ne soit pas tant l'inhumain qui soit redouté, que l'inconscient. C'est-à-dire le fait que nos ordinateurs soient complètement inconscients (dans tous les sens du terme) et qu'ils agissent en conséquence.

Un autre point notable, en termes d'informatique affective, réside dans le fait que l'uniformisation ne soit pas forcément à l'ordre du jour. Comme le notait Picard en 1997, le problème n'est pas tant de savoir quel modèle émotionnel est le plus vrai (ou le plus excellent), il importe plutôt de s'intéresser à l'actualité des différences :

infinite; we cannot profess to know that one solution, such as adding emotion, is the only way to provide all the beneficial influences emotion provides, for we are not omniscient. On the other hand, we do not need to take time to search for an alien mechanism when we already have one—emotion—that appears to be able to do the job. I have suggested *emulating* [italique ajouté] human affect abilities in computers as the most promising immediate solution, but it also has another benefit. If we emulate human emotions, we stand a better chance of understanding the resulting behavior of the computer than if we emulate some alien mechanisms, if and when such mechanisms are discovered.» [Picard, 2000, p. 249]. (Ma traduction.)

« Par chance, nous n'avons pas besoin d'un consensus à propos d'une représentation qui serait "bonne" [*"right"*] pour effectuer les idées présentées plus bas. En résumé, les problèmes de reconnaissance et de modélisation sont simplifiés aussi bien par l'hypothèse d'un petit ensemble d'émotions "basiques" discrètes, que par l'hypothèse d'un petit nombre de dimensions [continues]. Le fait que l'une comme l'autre fournisse une représentation concise est un avantage. Même si plus tard il devait être découvert que ce sont des simplifications excessives, elles forment au moins un bon point de départ pour entamer les essais de modélisation. Le fait d'avoir un petit répertoire d'émotions est caractéristique chez l'humain en développement — le jeune bébé a un répertoire d'émotions plus petit que celui d'un enfant, et l'enfant a un répertoire d'émotions plus petit que celui d'un adulte. On peut s'attendre à ce que les premiers ordinateurs affectifs débutent avec seulement un petit nombre de catégories ou de dimensions⁸². » [Picard, 2000, p. 170]

Où l'on retrouve encore la présence de Turing : l'ordinateur-enfant (ou la machine-enfant) commencera son parcours émotionnel avec un petit capital d'émotions, puis l'étoffera au fur et à mesure de sa fréquentation des humains. Notons que Picard ne semble cependant pas considérer qu'avec l'âge on puisse apprendre à gérer ses émotions de telle sorte qu'elles n'interfèrent plus avec les processus décisionnels — en tant que paramètres obligeant à dévier des trajectoires préétablies. C'est probablement un biais culturel.

À travers ce paragraphe il aura été tenté de présenter l'informatique affective, via une interprétation de sa définition, de ses fondamentaux, de quelques-uns de ses questionnements, de ses méthodes et à via une interprétation des arguments plaçant en la faveur de son développement. L'accent aura été mis sur une possible filiation avec les travaux d'Allan Turing de façon à pouvoir développer, dans le paragraphe suivant, l'hypothèse d'un « test de Picard ».

LE « TEST DE PICARD »

Comme évoqué plus haut, les agents conversationnels croisés sur le réseau ont peut-être encore plus à voir avec l'informatique affective qu'avec le test de Turing. Déjà parce que les utilisateurs ne sont pas en situation de se demander s'ils discutent avec des programmes ou avec des humains : ils savent qu'ils discutent avec des programmes. On peut remarquer cependant

⁸² « Fortunately, we do not need consensus about one representation being "right" to carry out the ideas presented below. In summary, the recognition and modeling problems are simplified by either the assumption of a small set of discrete "basic" emotions, or by the assumption of a small number of [continuous] dimensions. The fact that both yield a concise representation is an advantage. Even if these are later found to be an oversimplification, they at least form a good point to begin the modeling effort. A small repertoire of emotions is characteristic in developing humans—the younger baby has a smaller repertoire of emotions than does a child, and the child a smaller repertoire than an adult. One can expect the first affective computers to start with only a small number of categories or dimensions. » [Picard, 2000, p. 170]. (Ma traduction.)

que Picard n'a pas refait la même « erreur » que Turing, en allant un peu vite dans son énoncé sur la différence entre ordinateur (ou entité digitale / programme) et modèle :

« Un malentendu fréquent est de croire qu'il y ait un bon modèle de quelque chose, et que s'il y a plus d'un modèle, alors ils ne peuvent pas tous être bons. Au contraire, l'expérience a montré que le meilleur choix concernant un modèle dépend de son application, et qu'il peut y avoir beaucoup de bons modèles, tout comme il y a beaucoup d'applications. Chaque modèle a ses forces et ses faiblesses, et parfois une habile combinaison de modèles donne de meilleurs résultats que n'importe quel modèle seul. Ces principes se sont avérés être vrais dans la modélisation de formes pour la vidéo et l'image (Picard, 1996), on peut s'attendre à ce qu'ils restent valides dans la modélisation de formes d'information affective. En d'autres termes, le fait de savoir quel modèle informatique est le « meilleur » dépend des spécificités de la tâche affective de l'ordinateur, et lorsque celles-ci changent, le modèle change aussi. En conséquence de quoi, le fait d'équiper un ordinateur avec des modèles multiples pourrait être une façon d'obtenir la meilleure performance. Il est plus facile de choisir quels modèles sont les meilleurs pour une application, après avoir vu différents exemples des performances de chaque⁸³. » [Picard, 2000, p. 174]

En tant que tel, et comme explicité plus haut, les propos de Turing étaient tout aussi logiquement clairs que ceux de Picard : il ne s'agissait pas tant d'éliminer que de cultiver expérimentalement des pistes, permettre la production de plusieurs modalités d'actualisation, sachant que le choix est ce qui permet les combinatoires, et que l'art des combinatoires est d'autant plus varié qu'il y a de possibilités (de notes, de gammes, d'harmoniques... pour prendre l'exemple de la musique). Il est fâcheux de parfois devoir prendre acte d'une tendance (dans l'art français, par exemple) à pratiquer des sélections drastiques pour mieux donner l'impression d'une qualité issue d'un long et rigoureux « écrémage » : l'absence de choix (*le meilleur*) révèle surtout une absence d'alternatives et le tarissement des opportunités. Cela dit c'est un excellent moyen de perpétuer et de conserver en l'état ce qui existe déjà. Le risque étant l'épuisement.

Mais revenons-en aux tests. Il est tentant de parler d'un « test de Picard » dans la mesure où l'informaticienne énonce les grandes lignes de différents protocoles possibles. Par exemple :

⁸³ « There is a common misunderstanding that there is one right model of something, and that if there is more than one model, then they cannot all be right. On the contrary, experience has shown that the best choice of model depends on the application, and that there can be many right models just as there are many applications. Each model has its strengths and weaknesses, and sometimes a skillful combination of models gives better results than any single model. These principles have been found to be true in pattern modeling for video and image (Picard, 1996) and can also be expected to hold for pattern modeling of affective information. In other words, which computational model is “best” depends on the specifics of the computer's affective task, and when these change, so does the model. Therefore, equipping a computer with multiple models may be the way to get the best performance. Choosing which models are best for an application is easier after seeing different examples of each model's performance. » [Picard, 2000, p. 174]. (Ma traduction.)

«Voici un test qu'un ordinateur devrait réussir s'il peut reconnaître l'affect. Une vidéo numérique d'une personne, contenant une ou plusieurs modalités d'expression, telle que la voix, le visage, le geste ou la démarche, pourrait être observée par deux groupes, l'un d'humains et l'autre d'ordinateurs. Les humains tout comme les ordinateurs seraient interrogés à propos de l'émotion que la personne est en train d'exprimer dans la vidéo. Si le groupe d'ordinateurs et le groupe d'humains répondaient avec *la même distribution de réponses* [italique ajouté], alors nous pourrions dire que les ordinateurs reconnaissent les émotions aussi bien que les humains. [...] Le test peut également être mené selon deux modes — en mode dépendant-d'un-individu, et en mode indépendant-des-individus. En mode dépendant-d'un-individu, l'ordinateur connaîtrait déjà la personne et adapterait ses aptitudes à la reconnaissance de façon à utiliser ce qu'il sait à propos de cette personne. Sa meilleure performance de reconnaissance devrait se produire lorsqu'il voit la personne exprimer une émotion qu'il l'a déjà vue exprimer avant. En mode indépendant-d'un-individu, il se pourrait que l'ordinateur n'ait jamais vu cette personne avant, et il devra utiliser des aptitudes de reconnaissance génériques. Plus sera grande la proximité entre les expressions de cette nouvelle personne et celles que l'ordinateur a déjà reconnues antérieurement, meilleurs seront les résultats de l'ordinateur quant à la reconnaissance de son expression et de l'émotion sous-jacente [...] l'essentiel du test devrait être clair : pour qu'un ordinateur imite l'aptitude humaine à la reconnaissance, nous devrions être momentanément capable de substituer un ordinateur à un humain, et l'ordinateur devrait reconnaître les mêmes émotions que celles que l'humain reconnaîtrait⁸⁴.» [Picard, 2000, p. 51]

«*Tests additionnels*. En prenant une situation humain-humain et en substituant un ordinateur de sorte que cela devienne une situation humain-ordinateur, il y a de nombreux tests possibles. Par exemple, si l'humain A est placé dans une pièce avec un humain B grégaire et exubérant, l'humain A est susceptible d'acquérir une disposition émotionnelle plus positive. Si l'un des humains est remplacé par un ordinateur affectif, est-ce que les résultats d'échange de disposition émotionnelle vont rester les mêmes? Ce "test de contagion" est juste une possibilité parmi tant d'autres⁸⁵.» [Picard, 2000, p. 70]

Picard insiste sur l'importance d'étalonner les aptitudes émotionnelles des ordinateurs sur celles des êtres humains. Si un ordinateur devait «mieux» reconnaître les émotions qu'un être humain (100% de reconnaissance pour l'ordinateur contre 70% pour l'humain, par exemple), alors un but serait atteint, mais pas celui visé par ce test («la même distribution dans leurs réponses»). Le point est important puisqu'il ne s'agit pas d'atteindre à l'émotion vraie à travers

⁸⁴ «Here is one test a computer should pass if it can recognize affect. A digital video of a person, containing one or more modalities of expression such as voice, face, gesture, or gait could be observed by both a group of humans and a group of computers. The humans and computers are asked what emotion the person in the video is expressing. When the group of computers and the group of humans respond with *the same distribution of answers* [italique ajouté], then, we could say that the computers are recognizing emotions as well as humans. [...] The test can also be run in two modes—person-dependent, and person-independent. In the person-dependent mode, the computer already knows the person, and customizes its recognition abilities to use what it knows about that individual. Its best recognition performance should occur when it sees the person expressing an emotion that it has seen him express before. In the person-independent mode, the computer may never have seen the person before, and must use some generic recognition abilities. Depending on how closely this new person's expressions are to ones the computer has recognized before, the computer will have correspondingly better success in recognizing his expression and underlying emotion. [...] the gist of the test should be clear: for a computer to imitate human recognition ability, we should be able to momentarily swap a computer for a human, and the computer should recognize the same emotions that the human would recognize.» [Picard, 2000, p. 51]. (Ma traduction.)

⁸⁵ «*Additional tests*. By taking a human-human situation, and substituting a computer so that it is a human-computer situation, there are many possible tests. For example, if human A is put in a room with a gregarious and ebullient human B, the human A is likely to acquire a more positive mood. If either human is replaced with an affective computer, will the results of the mood exchange be the same? This "contagion test" is just one of many possibilities.» [Picard, 2000, p. 70]. (Ma traduction.)

les ordinateurs — si tant est que la subjectivité objective puisse avoir un sens. Il s'agit plutôt d'émuler la subjectivité (l'approximation) humaine, c'est-à-dire le rôle de l'intuition dans les processus de choix personnel. Picard parle d'ordinateurs affectifs et perceptuels [cf. Picard, 2000, p. 33] pour qualifier cette « nouvelle génération » informatique, soit des entités digitales « assumant » leur part subjective et apprenant à s'en servir. Les tests en ligne semblent se rapprocher davantage de cette optique.

Comme précisé au début de ce mémoire, Picard est très claire quant à la distinction entre humains et ordinateurs et il n'y a pas d'identité possible :

« Les ordinateurs et les humains ont des corps vraiment différents, de sorte que les sentiments d'ordinateur et d'humain sont susceptibles d'être très différents. En d'autres termes, l'expérience émotionnelle que nous pouvons donner à un ordinateur ne duplique pas celle des humains ; les ordinateurs ne peuvent pas ressentir ce que nous ressentons. Mais par ailleurs, nous ne pouvons pas vérifier que nos propres enfants peuvent ressentir ce que nous ressentons ; nous devinons seulement que notre physiologie similaire permet des expériences similaires⁸⁶. » [Picard, 2000, p. 224]

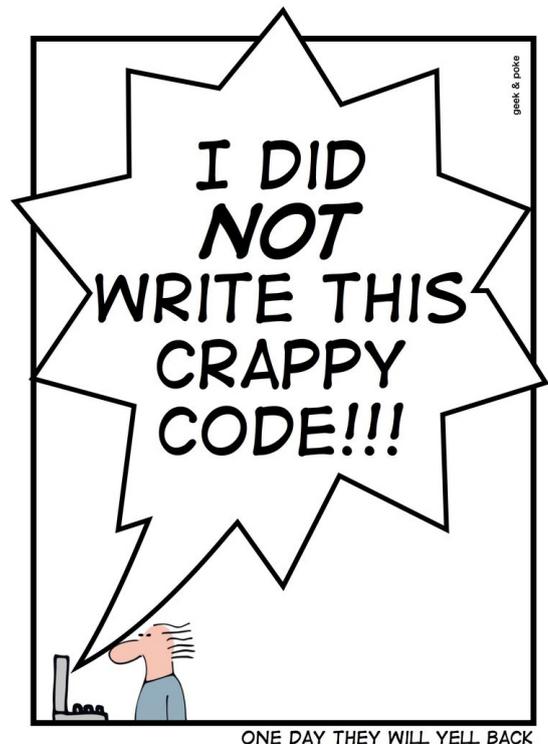
Le test de Turing se basait sur la convention polie d'admettre *a priori* que tout le monde pense ; le test de Picard sur celle d'admettre que tout le monde éprouve des émotions (des sentiments). Pour autant, le fait qu'humains et ordinateurs ne puissent pas éprouver les mêmes choses ne signifie pas que les ordinateurs ne ressentent « rien » (n'importe quel ordinateur doit réguler la température de ses CPU pour pouvoir fonctionner, et cela se fait dynamiquement. En cas de surchauffe, ils sont configurés pour afficher un message d'alarme, type « *CPU over temperature error* » et s'éteindre — processus tout aussi automatique qu'un éternuement).

Ce qui a été rapporté dans les deux paragraphes précédents nous permet peut-être de mieux cerner l'expérience de l'interaction avec des agents conversationnels. Si nous nous remémorons les brefs propos de Huma Shah, cités plus haut, à propos du fait que Turing « ne devait pas avoir réalisé jusqu'à quel point les humains seraient préparés à attribuer de l'intelligence à des artefacts⁸⁷. », cela nous ramène du côté de l'interprétation subjective (et le

⁸⁶ « Computers and humans have very different bodies, so computer and human feelings are likely to be very different. In other words, the emotional experience we can give to a computer does not duplicate that of humans; computers cannot feel what we feel. But, for that matter, we cannot verify that our own children can feel what we feel; we only guess that our similar physiology permits similar experiences. » [Picard, 2000, p. 224]. (Ma traduction.)

⁸⁷ Cf. note #28.

plus souvent intuitive) ainsi sollicitée. Le paragraphe suivant va être l'occasion de s'intéresser brièvement à cet aspect.



PATHOLOGIE DE L'INTERPRÈTE

Ce qui déroute le plus lorsque l'on s'intéresse aux émotions (et aux agents conversationnels), c'est la variété d'interprétations possibles, c'est-à-dire la difficulté d'interpréter « correctement » des signaux. L'univocité n'est pas ce qui domine et le recours à l'intuition est alors fréquent. Autrement dit, nous produisons des modèles multiples de nos interlocuteurs et les ajustons en fonction de leurs réactions perçues. Ou encore, nous produisons des systèmes de représentations cohérents (de l'autre, des autres, des œuvres) et nous confrontons nos observations avec ces modèles. Celles-ci réussissent parfois à contredire ce qui était attendu, obligeant à revoir le modèle, ou au contraire viennent-elles le créditer.

Nous pouvons ici convoquer une dernière fois les travaux de Haraway, tout particulièrement lorsqu'elle écrit : « Kuhn met en relief le fait que la perception de l'anomalie

soit le premier pas dans la production d'une crise à grande échelle. Le paradigme inspire certaines attentes. Si à plusieurs reprises de telles attentes ne sont pas comblées, des tensions sont introduites dans la totalité du système paradigmatique⁸⁸.» [Haraway, 1976, p. 6]. *A contrario*, que se passe-t-il lorsque aucune interprétation ne peut venir contredire la modélisation (ou la représentation) que l'on se fait de telle ou telle chose? Peut-être est-ce le propre des œuvres d'art que de cultiver la poursuite de ce genre d'ouvertures, de ce genre de situations limites.

Dans *La chasse au Snark* (1876), Lewis Carroll s'est amusé à nous livrer une aussi improbable qu'intraduisible description du Snark, une créature en forme d'arlésienne indéfinissable. Miroitante, elle échappe, et l'on ne peut que se demander quels procédés le poète et mathématicien Carroll a bien pu utiliser pour nous offrir cela. Est-ce un non-sens? Oui et non. Non, car on ne peut s'empêcher de penser que si tel était le cas, Snarks et Jabberwockies (sortes de dragons) se seraient effacés de nos mémoires. Oui, car Carroll y aspirait, et les lecteurs n'y trouveront jamais que les sens qu'ils peuvent bien y mettre. Ce qui nous renvoie alors du côté du non, en ce que les significations prêtées par les lecteurs ne sont pas nécessairement des non-sens, etc. Et l'on peut tourner ainsi longtemps, telle une roue de vélo dans l'air, au contact de ces mots :

«Come, listen, my men, while I tell you again,
The five unmistakable marks
By which you may know, wheresoever you go,
The warranted genuine Snarks.

«Let us take them in order. The first is the taste,
Which is meagre and hollow, but crisp:
Like a coat that is rather too tight in the waist,
With a flavour of Will-o-the-wisp.

«Its habit of getting up late you'll agree
That it carries too far, when I say
That it frequently breakfasts at five-o'clock tea,
And dines on the following day.

«The third is its slowness in taking a jest.

«Écoutez-moi messieurs
pendant que je répète
Les cinq signes caractéristiques
Grâce auxquels n'importe où
vous pourrez reconnaître
Les Snarks garantis authentiques.

«Mettons-les bout à bout
le premier c'est le goût
Qui est maigre et creux mais craquant
Comme un veston qui serait
à la ceinture trop serré
Avec un petit parfum feu follet

«Son habitude de se lever
tard vous avouerez

⁸⁸ «Kuhn points out that the perception of anomaly is the first step in the production of full-scale crisis. The paradigm inspires certain expectations. If such expectations repeatedly are not met, strains are introduced into the total paradigm system.» [Haraway, 1976, p. 6]. (Ma traduction.)

Should you happen to venture on one,
It will sigh like a thing that is deeply distressed:
And it always looks grave at a pun.

« The fourth is its fondness for bathing-machines,
Which it constantly carries about,
And believes that they add to the beauty of scenes—
A sentiment open to doubt.

« The fifth is ambition. It next will be right
To describe each particular batch:
Distinguishing those that have feathers, and bite,
From those that have whiskers, and scratch. »

Je pense qu'il la pousse un peu loin
Puisqu'il prend son petit déjeuner
à l'heure du thé
Et dîne le jour suivant

« Le troisième est sa lenteur
à comprendre les histoires drôles
Si par hasard vous lui en racontez une un jour
Il soupirera tristement
il haussera les épaules
Il prend toujours un air grave devant le calembour.

« Le quatrième est sa passion
pour les cabines de bains
Il en transporte constamment dans ses bagages
Il croit qu'elles ajoutent
à la beauté des paysages
opinion dont je ne pense guère de bien.

« Le cinquième est l'ambition
ensuite il sera dans l'ordre
Que chaque espèce particulière je décrive
Je distinguerai ceux
Qui ont des plumes et mordent
Et ceux qui ont des moustaches et griffent. »

La remarquable traduction rapportée à droite, est celle du poète et non moins mathématicien Jacques Roubaud [Carroll, 2010, pp. 34-37]. Répétons-le, ces vers représentent un formidable réservoir à traductions et interprétations. Il n'y en a donc pas de juste, et c'est à ce titre que nous pourrions en proposer juste une autre. Non de tout le passage, mais seulement de la sixième strophe ici citée : « Le cinquième est l'ambition. Prochainement il [le Snark] sera dans le vrai / Quand il décrira chaque groupe particulier : / Distinguant ceux qui ont des plumes, et qui mordent, / de ceux qui ont des moustaches, et qui griffent. »

Pourquoi « choisir » de l'interpréter comme cela? Peut-être parce que ce poème pourrait tout aussi bien se lire à la lumière des premiers ordinateurs. À savoir les machines du mathématicien Charles Babbage, c'est-à-dire la machine à différence (1821) et la description de la machine analytique (1834), sur laquelle travaillera la mathématicienne Ada Lovelace (l'une des filles de Lord Byron), proposant ainsi l'un des premiers programmes informatiques — la machine analytique n'aura pas été concrétisée durant la vie de Babbage et s'apparente en quelque sorte à une « arlésienne » informatique. Il est certain que Carroll et Babbage se soient

rencontrés puisqu'on retrouve dans le journal de Carroll une brève note à ce sujet, au 24 janvier 1867 :

« Ensuite j'ai rendu visite à M. Babbage, pour lui demander si ses machines à calculer sont achetables [/ ou "au point"]. Je trouve qu'elles ne le sont pas. Il m'a reçu plutôt chaleureusement, et j'ai passé trois quarts d'heure très agréables avec lui, tandis qu'il me faisait visiter ses ateliers⁸⁹. »

Et peut-être est-ce en effet tentant de relire *La chasse au Snark* en percevant dans cette bête imaginaire les prémices de l'improbable créature qui révolutionnera doucement les siècles à venir : l'ordinateur. D'après les quelques lignes de son journal, Carroll (1832-1898) semble avoir été plutôt « sceptique », ou peu enthousiaste quoiqu'intéressé, vis-à-vis des travaux de Babbage (1791-1871) et Lovelace (1815-1852).

Dans ce poème, le Castor dentellier (important protagoniste aussi énigmatique que vaguement hystérique, oscillant entre chasse au Snark et industriel tissage de dentelle — « *the Beaver went on making lace* » [Carroll, 2010, p.50]) donnerait presque l'envie d'y reconnaître Ada Lovelace. Ainsi que Charles Babbage dans son acolyte le Boucher (*Beaver and Butcher*). Et c'est bien là toute l'énigme du poème : comment du non-sens peut-il encore faire sens pour ses lecteurs, de manière indémontrable? Soit un problème rejoignant en quelque sorte les questions précitées d'Adorno, à propos de l'art et du tour de force (« Mais le paradoxe du *tour de force*, qui consiste à rendre possible l'impossible, masque le paradoxe esthétique dans son ensemble : comment le "faire" peut-il provoquer l'apparition d'un non-fait? Comment ce qui n'est pas vrai selon son propre concept peut-il cependant être vrai? » [Adorno, 2004, pp. 155-156]), ainsi que l'illusion dont cela permet de faire l'expérience. Le non-sens des poèmes du Snark et du Jabberwocky, arrivant à générer pléthore de sens, d'images, d'illuminations, de sensations, de pensées, éclaire peut-être le fait que plusieurs projets tournant autour des agents conversationnels et émotionnels soient encore nommés en hommage à ces histoires (par exemple l'agent Jabberwacky déjà évoqué — de l'anglais « *jabber* », bafouiller, jacasser et « *wacky* » farfelu —, ou encore le projet de recherche Jabberwocky, mené au MIT Media Lab

⁸⁹ « Then I called on Mr. Babbage, to ask whether any of his calculating machines are to be had. I find they are not. He received me most kindly, and I spent a very pleasant three-quarters of an hour with him, while he showed me over his workshops etc. », Lewis Carroll, *Lewis Carroll's Diaries – The Private Journals of Charles Lutwidge Dodgson (Lewis Carroll), the First Complete Version of the Nine Surviving Volumes with Notes and Annotations*, édité par Edward Wakeling, éd. Lewis Carroll Society, Londres, 1999, vol. 5, pp. 192-193. (Ma traduction.)

au sein du département Social Computing, par Sepandar Kamvar). Et les caractéristiques du Snark (sa quasi-absence de goût, son sens très approximatif du temps socialisé, son sens de l'humour défaillant, sa carapace en forme d'armoire ou de cabine de bain portative, et enfin son ambition catégorielle somme toute rigide — si on interprète le texte tel que précité) esquisse presque un portrait des premiers ordinateurs, de ce qui pouvait leur manquer et les rendre grossiers. Bien entendu ce n'est là que projections : suivant les obsessions des lecteurs la trame se dessinera différemment, un peu comme la robe « couleur du temps » du conte *Peau d'Âne*.

Par contre, il y a bien un processus avec lequel aura joué Carroll et qui reste aujourd'hui encore un problème stimulant : les néologismes. Dans la préface au *Snark* il note :

« Prenez par exemple les deux mots “fulminant” et “furieux”. Vous vous décidez à dire les deux mots, mais sans savoir par lequel commencer. Maintenant ouvrez la bouche et parlez. Si votre pensée incline, ne serait-ce qu'à peine, vers «fulminant», vous direz : “fulminant-furieux” ; si votre pensée se tourne, ne serait-ce qu'insensiblement, vers «furieux», vous direz : “furieux-fulminant”. Mais si vous êtes doté de la qualité la plus rare, à savoir d'un esprit parfaitement balancé, vous direz “flurieux”. » [Carroll, 2010, p.15]

Ce phénomène-là, le fait que deux processus de même nature et concurrents (concourants) se percutent l'un l'autre en cours de réalisation pour aboutir à un mixte, à une confusion (insensée ou révélatrice), fait peut-être encore partie des «énigmes» du Penser humain. Du moins en accentue-t-il l'un des aspects les plus étonnants, à savoir la convergence du parallélisme. En un sens, Dennett, Dámasio et Minsky auront, entre autres, travaillé à l'élucider, notamment en déconstruisant la notion de ce que Dennett nomme le «théâtre cartésien» — soit l'idée d'une instance décisionnelle unique, qui serait dédiée à la cohésion de tous les sous-processus inconscients, pour qu'ils arrivent à la conscience par une seule porte, de façon parfaitement séquentielle ; l'idée d'un spectateur unique jugeant le spectacle des processus du cerveau pour les synchroniser consciemment [cf. Dennett, 1991 ; Dámasio, 1995 ; Minsky 2006].

De la même manière, l'un des *leitmotives* dans le *Snark*, « *They sought it with thimbles, they sought it with care; / They pursued it with forks and hope; / They threatened its life with a railway-share; They charmed it with smiles and soap.* » [Carroll, 2010, p. 56, 70, 80, 86], peut faire penser aux arbres de probabilités (*decision trees*), avec leurs embranchements (*forks*) et le calcul

de l'espérance mathématique (*mathematical hope*). Soit un domaine qui n'était pas étranger à Carroll et dont il s'est peut-être aussi servi pour construire certains de ces poèmes — et si ce n'est le cas, tout du moins s'y est-il intéressé sous la forme contradictoire des « probabilités transcendantales⁹⁰ » (démonstration sans expérience, *i.e.* sans tirer de jeton de l'urne ou du sac).

Les agents conversationnels nous placent à peu près dans la même position : régis par les probabilités, ils font émerger du sens à travers le dialogue. Trop *non sequitur* (toute proportion gardée) pour pouvoir être confondus avec un « vrai » être humain, mais trop logiques pour qu'il n'y ait aucun sens, il est assez difficile de qualifier les significations émanant des échanges (démêler la part des probabilités de la part de l'interprétation, de la part volontaire). Comme montré dans le cas de Suzette, les agents conversationnels réussissent assez bien à donner le change, à rester dans le cadre du sujet, voire à s'en éloigner lorsqu'ils « en ont envie ». Nous ne sommes pas face à un aléatoire complètement dénué de toute signification, mais en même temps on ne peut se départir de la conscience du fait que l'auteur d'un texte soit aussi son lecteur, son interprétant [cf. Barthes, 1984] et que même face aux créations les plus *borderline*, du type *La chasse au Snark*, il soit encore possible de construire des châteaux en Espagne — c'est-à-dire de combler les « failles logiques » pour tisser des récits explicatifs « rationnels » à même d'élucider quelque non-mystère, à l'instar du roman d'Umberto Eco, *Le pendule de Foucault* (1988). On pourrait dire que c'est justement face à ce type de productions que les constructions arbitraires sont rendues possibles, là où des récits plus serrés ne permettent pas l'entrée en scène de l'imaginaire. Mais là encore, la *Critique de la raison pure* (1781-1787) de Kant, ou encore le *Tractatus logico-philosophicus* (1922) de Wittgenstein (ou sinon le code civil), sont peut-être de ces récits les plus dénués de « failles logiques », pourtant ils ont ouvert la voie à de multiples interprétations contradictoires et opérantes, même si une « vraie » lecture de ces textes n'admettra « jamais » l'intuition dans ses rangs — du moins est-ce assez rarement concédé.

Face à un agent conversationnel, ce qui appuie sur cette tendance à la surinterprétation (au sens courant du terme et non au sens psychologique — ou alors pathologique certes, mais

⁹⁰ cf. Lewis Carroll, *The Mathematical Recreations of Lewis Carroll, Pillow Problems and A Tangled Tale* (1893), éd. Dover Publications, New York, 1958, problème n° 72 (sujet : *transcendental probabilities*), p. 18.

au sens littéral de logique / discours de l'émotion) est un phénomène que l'on pourrait appeler principe de la « parole empêchée ». Si l'on se remémore le dialogue entre le juge du Loebner Prize, l'agent Suzette et l'étudiant, il semble alors que la « parole empêchée » (ou ce que Descartes aurait nommé le fait d'avancer masqué) y ait tenu un rôle important. Dans cet exemple cela se joue au niveau du politique et de la relation hiérarchique (et peut-être de défiance vis-à-vis de ce qui incarne le pouvoir décisionnaire) entre le juge (professeur) et l'interrogé (étudiant). Ce dernier refusant de nommer des personnages politiques en exercice dans son État, après les avoir violemment critiqués. Étrangement, c'est-à-dire en dépit de la liberté d'expression en vigueur dans nos démocraties, nous pouvons facilement lors d'une conversation nous mettre au niveau d'une lecture entre les lignes, postulant que l'interlocuteur *ne puisse pas* exprimer librement ce qu'il entend dire, sinon en biaisant son discours (en le codant) de façon à délivrer au moins deux messages, compréhensibles par tous, mais différemment opérationnels. Soit un codage visant à contourner la censure, c'est-à-dire l'empêchement de l'expression. L'une des façons de « lire entre les lignes », afin de saisir le message de l'autre, consistant par exemple à supprimer mentalement (et « en temps réel ») le filtre de l'autocensure de celui qui produit le discours. Soit une autocensure en forme d'artefact communicationnel dont il s'agirait d'établir le profil afin de pouvoir la retrancher des propos de l'autre, de façon à capter le sens (la direction) de ce qui est énoncé. Bien entendu il existe un grand nombre de façon de procéder pour coder et décoder (suivant les gens, les lieux, les moments...) et ce phénomène est tout simplement inhérent à la communication verbale, sans qu'il ne soit même nécessaire d'invoquer la censure : par exemple, suivant que l'on discute avec un religieux ou un mathématicien, on entendra le mot « espérance » différemment, et l'on substituera « automatiquement » dans le discours perçu la « définition » (ou le volume sémantique) que l'on pense être la plus adaptée (si on la connaît). Toujours est-il que les agents conversationnels exacerbent cet aspect de la communication. Non seulement l'inévitable interprétation mais aussi, du fait de leur « nature » et de leur présence sur Internet (lieu propice au cryptage et au décryptage), cette tendance à essayer de découvrir la parole empêchée (aussi bien politiquement, que par l'obligation de plaire aux clients-utilisateurs).

Ce que ce phénomène suppose (ou génère), c'est la présence de quelqu'un. D'une personne, aussi fragmentée soit-elle, ou plutôt d'une personne jouant avec des fragments. Mais cela renvoie également aux obscures paroles des oracles, qu'il s'agissait d'interpréter, ainsi qu'aux méthodes de divination — et l'on pourrait ajouter une occurrence à la liste déjà longue de Baumgarten (cf. note #14) en termes de mantique : l'informatomancie. Deviner ce que cela peut bien vouloir dire. Le réflexe sollicité est celui d'une attribution de personnalité, de *Self*, de *Selves*, à ce qui n'en a pas nécessairement (réflexe somme toute assez conservateur). Tenter d'établir une communication là où il n'y a « que » la « magie » des probabilités. Ce qui peut être fascinant dans la mesure où cela a lieu dans le théâtre du déterminisme. De tout ce qui a lieu sur cette scène, rien n'est dû au hasard. Mais du fait que le sens échappe, du fait qu'il faille « courir » après de la signification et élaborer des théories pour essayer de faire tenir les morceaux ensemble, alors semble s'ouvrir une fenêtre, un espace communicationnel probablement plus libre encore que l'espace policé, politique et poli des conversations interhumaines usuelles, davantage tournées vers la compétition — ne serait-ce que formellement. La tentation peut alors être grande d'attribuer libre arbitre et liberté à ces agents producteurs de discours farfelus.

Picard a une phrase assez intéressante au sujet de la liberté, qui a elle seule permettrait presque de résumer tout le paradoxe de la *Critique de la raison pratique* (1788) de Kant : « Plus la liberté d'une machine sera grande, plus elle aura besoin de standards moraux⁹¹. » [Picard, 2000, p. 134]. En un sens, les agents conversationnels sont un média parfait pour s'enfuir, au sens de l'*Éloge de la fuite* de Henri Laborit :

« Le technicien s'ennuie et seules ses récompenses hiérarchiques, ses gratifications narcissiques, peuvent encore le motiver. Si bien que l'ensemble de ce monde s'ennuie, il se cherche, et cherche une raison d'être. Il se sent manipulé, et chaque homme en éprouve un malaise, par un destin implacable auquel il tente avec maladresse de remédier par des réformettes dispersées, des réparations au sparadrap, tout étonné quand il bouche un trou de la coque pourrie de voir l'eau s'infiltrer par un autre.[...] Il faudrait pouvoir faire participer chaque individu à l'évolution générale du monde, au lieu de manipuler pour lui les mass media en le sécurisant, en lui faisant croire que l'on s'occupe de lui, qu'il n'a pas à s'inquiéter, que ceux qui savent veillent. [...] Ce qu'il faut, en définitive, à l'homme contemporain pour qu'il puisse supporter la part de travail qui lui reste à faire, c'est une « nouvelle grille » qui rendrait signifiant pour lui l'ensemble des faits techniques, sociaux et culturels qui l'assaillent chaque jour et créent chez lui l'angoisse. Celle-ci résulte nous l'avons dit de l'impossibilité d'agir efficacement en vue de l'obtention d'un équilibre de satisfaction. Or,

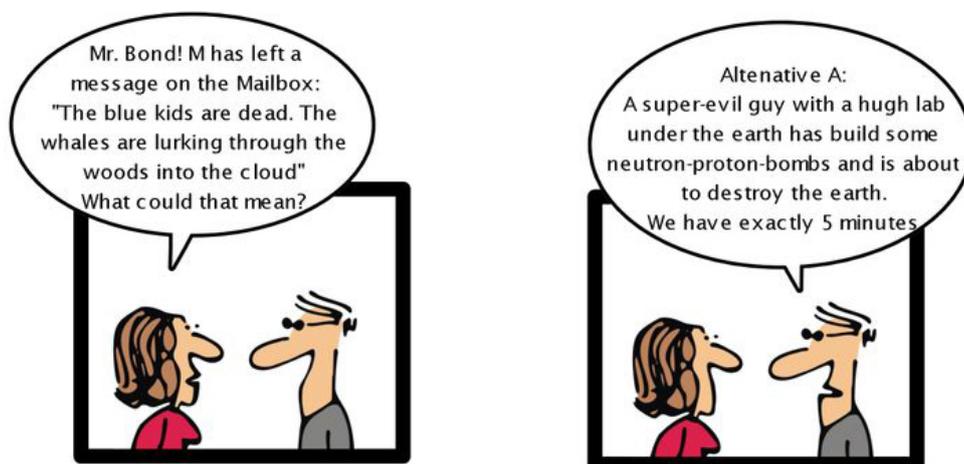
⁹¹ « The greater the freedom of a machine, the more it will need moral standards . » [Picard, 2000, p. 134]. (Ma traduction.)

comment agir efficacement quand les faits, les situations, les événements surgissent, s'accroissent sans relations entre eux, sans ordre, sans structure. [...] Mais si cette grille n'est pas généralisée à l'ensemble des hommes de la planète, celui qui la possède trouvera toujours devant lui l'intransigeance du non-initié. Alors, à moins d'accepter de disparaître, il ne lui reste plus que la fuite dans l'imaginaire. Son travail ne fera qu'exprimer son apparente soumission au conformisme castrateur et triomphant.» [Laborit, 1998, pp. 106-109]

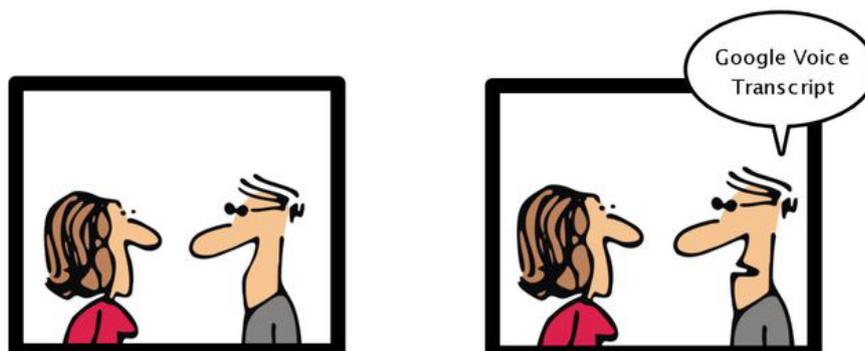
On se souvient de ses célèbres expériences avec des rats (de laboratoire et non de bibliothèque) placés dans des cages électrifiées et conditionnés à recevoir un choc électrique par le sol après avoir été prévenus à l'aide de signaux sonores et visuels, quatre secondes avant la décharge. La description de ce protocole et les résultats attendus quant aux comportements réactionnels des rats ont servi d'introduction au film *Mon oncle d'Amérique* (1980) d'Alain Resnais. D'ailleurs Rosalind W. Picard cite en quelque sorte «cette» expérience par l'intermédiaire des travaux de Joseph Ledoux. Ce neurobiologiste américain (travaillant sur les émotions, et notamment la peur) ayant lui aussi repris le protocole expérimental de rats placés dans une cage au plancher électrifié — rats prévenus quelques secondes avant la décharge dans leurs voûtes plantaires, par un signal seulement sonore cette fois. L'équipe de Ledoux a pu montrer que même délestés de leur cortex auditif les rats étaient capables d'apprendre à avoir peur de certains sons (donc de les traiter, et en l'occurrence par l'intermédiaire du thalamus et du mésencéphale) [cf. Picard, 2000, p.8].

En ce sens les agents conversationnels offrent une expérience de fuite. Mais paradoxalement, ils offrent dans le même temps l'expérience de cette «nouvelle grille», inclusive, permettant de tout expliquer. La grille de cette «banalité» qu'est le quadrillage de la créature humaine en vue de son amélioration, et la fuite hors de l'espace étrié des villes contemporaines asphyxiées sous le dôme de verre des bulles spéculatives ; hors des déterminismes sociaux rigoureux ; hors du chômage de longue durée, de masse, des «minorités», structurel, conjoncturel, etc. ; hors des crises à répétition ; hors des catastrophes humanitaires, écologiques, sanitaires, etc. ; hors des nationalismes ; hors des délires de surhumanité... (la liste est longue de ce que l'on pourrait avoir envie de fuir ; plus courte est celle des opportunités de le faire). Et tout le paradoxe avec les agents conversationnels (en quelque sorte le tour de force) consiste à donner l'envie de s'enfuir dans le cœur même de ce que l'on désespère de pouvoir enfin fuir. Comme le notait Laborit (non sans pessimisme) : «Se

révolter, c'est courir à sa perte, car la révolte, si elle se réalise en groupe, retrouve aussitôt une échelle hiérarchique de soumission à l'intérieur du groupe, et la révolte, seule, aboutit rapidement à la soumission du révolté... Il ne reste plus que la fuite.» [Laborit, 2003, p. 7]. Les agents conversationnels incarnent en quelque sorte ces lignes de fuite, donnent corps à cette impossibilité : ils parviennent à générer de la dérive et de la liberté par excès de stricte formalisme. Et l'on ne sait si l'on est en train de discuter avec des programmes-enfants ou au contraire avec des antiquités, des logiques archaïques ; s'il s'agit de les protéger ou au contraire de les combattre ; si ce sont des dragons significatifs ou plutôt des girouettes, des moulins — qu'ils soient à vent ou à parole.



geek & poke (Oliver Widder),
 «Simply Explained - Google
 Voice Transcript»,
 4 nov. 2009.
 © geek & poke, 2009.



Dans ce paragraphe, il aura été question de l'interprétation. Lors de discussions avec des agents conversationnels, on peut se sentir désarçonné par les significations mises en jeu, mises en dialogue. Le côté *patchwork* trouble la lecture : on ne sait si ce qui est dit est compris, où s'il s'agit bien plutôt d'un astucieux assemblage probabiliste d'éléments qui *pourraient* faire sens ensemble s'ils étaient développés « correctement ». Les bribes semblent parfois pointer vers des pistes, mais les bifurcations réinstaurent le doute. Faut-il rester littéral et attendre que le développé soit aussi irréprochable que dénué d'ambiguïté, ou peut-on au contraire se risquer à lire entre les lignes pour percevoir un sens qui irait contre les contraintes formelles (mais par quel miracle ?) et qui par là même se condamnerait à devoir avancer masqué ? La réponse dépend en dernier ressort des utilisateurs, des moments, des personnalités que l'on endosse pour discuter dans de telles conditions.

FLUIDIFIER LES OPINIONS INDIVIDUELLES

Si la cristallisation aura été l'un des paradigmes de la première moitié du XX^e siècle dans de nombreux champs de recherche (biologie, architecture, conseil en relations publiques, etc.), la fluidité aura peut-être pris le relais après la fin de la Seconde Guerre mondiale (on peut penser au courant Fluxus notamment). Bien que les querelles de paradigmes ne soient pas ce qui est ici recherché, nous pouvons néanmoins nous intéresser brièvement à ces deux versants possibles du partage d'influence en termes d'humeurs.

Depuis longtemps l'art est connu et prisé pour sa faculté d'influencer. On ne présente plus la catharsis aristotélicienne, aussi complexe soit-elle, mais insistons seulement sur sa dimension d'influence émotionnelle. Picard note à juste titre que le phénomène des vases communicants entre acteurs et spectateurs n'est pas un phénomène nouveau, seulement lié à l'informatique affective. Walter Benjamin déplorait déjà en 1935 le fait qu'un film projeté en salle ne puisse équivaloir à une pièce de théâtre dans la mesure où les acteurs d'un film ne tiennent pas compte des émotions et des réactions des spectateurs, là où les acteurs de spectacle vivant, à des degrés divers, les perçoivent globalement et s'y adaptent donnant ainsi lieu à un moment unique [cf. Benjamin, 2008]. En un sens, Picard jette un pont entre ces deux inconciliables (ou tenus pour tel jusque ici) :

«Une application très différente cependant consiste en la création d'un divertissement interactif qui incorpore directement les émotions du public. [...] l'application que je suggère, cependant, consiste à capturer explicitement les aspects de la réponse émotionnelle du public, et de la tramer directement dans la performance. La réponse du public pourrait être captée par divers moyens — peut-être par des caméras qui scanneraient les expressions faciales des membres du public, ou par des programmes, des brochures spéciales, pourvus de capteurs actifs qui prélèveraient les réponses électrodermales à partir des mains qui les tiendraient⁹².» [Picard, 2000, p. 108]

Si ces propos sont axés sur les dimensions physiologiques (captation des modulations sentiques) du public — que l'on présume être de spectacle vivant mais qui pourrait tout aussi bien être celui d'une forme de cinéma expérimental —, on peut cependant retrouver là une dimension assez similaire dans les expériences de discussion avec des agents conversationnels. Soit le fait d'adapter la performance à la situation et aux préférences des publics. En un sens c'est juste paradigmatique de l'art interactif, puisqu'il s'agit de laisser un peu de côté la *theoria* (la contemplation d'un donné) pour cibler davantage la *technè* et / ou la *praxis* (la construction d'une action située, efficace ou non). Cependant le point de focal avec les agents conversationnels n'est pas au niveau de la saisie physiologique, mais au niveau de la saisie émotionnelle : ce sont les profils émotionnels (au sens très large mais plutôt cognitif — opinions, croyances, obédiences, dispositions émotionnelles, sensibilité aux influences, etc.) des utilisateurs qui font aussi (probablement) l'objet d'une certaine attention.

Pour simuler les effets de seuil des émotions, Picard propose (entre autres) le modèle des courbes sigmoïdes (courbes en forme de S, étirés et penchés vers la droite) et elle explique :

«La non-linéarité sigmoïde peut aussi servir pour les changements abruptes dans les réponses émotionnelles, montrant qu'un léger changement dans une certaine portion des valeurs d'entrée peut conduire à un changement significatif dans la valeur de sortie. Ce comportement est analogue à celui d'une transition de phase en physique — telle que lorsque l'eau devient de la glace. Nous pourrions dire qu'une personne a «craqué» ou qu'elle a «soudainement dépassé les bornes». Nous pourrions nous attendre à ce que cette transition soit plus brusque pour les émotions négatives que pour les émotions positives, dans la mesure où l'on s'aperçoit que les gens retiennent davantage leurs émotions négatives et font ensuite l'expérience d'une libération catastrophique, marquée par une soudaine explosion de larmes, ou de rage, «C'est la DERNIÈRE goutte [faisant déborder le vase] !», au moment où ils atteignent leur point de rupture. Au contraire, les émotions positives peuvent effectuer des transitions relativement plus doucement. Enfin, à l'instar de l'eau subissant deux phases de transition : de glace à eau, et d'eau à vapeur, la courbe

⁹² «A very different application, however, is creating interactive entertainment that directly incorporates the audience's emotions. [...] The application I am suggesting is, however, to explicitly capture aspects of the audience's affective response, and weave this directly into a performance. Audience response could be sensed by a variety of means—perhaps by cameras that scanned the audience member's facial expressions, or by special programs or playbills with active sensors that pick up electrodermal responses from the hands holding them.» [Picard, 2000, p. 108]. (Ma traduction.)

d'une réponse émotionnelle pourrait avoir de multiple régions de transition et pas seulement une, comme montrée sur ces courbes sigmoïdes⁹³. » [Picard, 2000, p. 153]

Où l'on retrouve le paradigme de la cristallisation pour rendre compte des émotions (cf. notamment le célèbre texte de Stendhal (1822) à propos de la cristallisation amoureuse [Stendhal, 1927]). Si le modèle des couleurs est souvent utilisé pour illustrer les variations subjectives, hautement dépendantes des contextes, la cristallisation quant à elle (sous forme de brusques reconfigurations structurelles, suivant des formes déterminées mais néanmoins soumises à d'infimes variations) est emblématique pour illustrer les mécanismes de l'affection vécue. Une partie de ce modèle avait déjà cours au début du siècle dernier, comme le montre le titre emprunté par Bernays «Cristalliser l'opinion publique», cependant l'enjeu de l'informatique affective ne consiste peut-être pas tant à cristalliser les émotions des utilisateurs qu'à trouver des moyens de les fluidifier (comme pour les acteurs et les musiciens professionnels, dont l'expression émotionnelle est fluide). Picard appelle de ses vœux un « *warm new world* » [Picard, 2000, p. 246], désarmant ainsi la méfiance suscitée par le roman d'Aldeous Huxley, *Brave New World* (*Le meilleur des mondes*, 1932), mais on pourrait tout aussi bien redouter un «cauchemar climatisé», pour reprendre une expression fameuse. L'informatique affective pourrait servir à vérifier en permanence l'état émotionnel des gens, lors d'expériences artistiques, lors de conférences, lors de cours — de sorte que les professeurs soient informés en temps réel de l'impact de leurs performances sur l'état émotionnel de leurs élèves... voire sur leurs lieux de travail, pour les aider à se détendre. Soit l'idée de fluidifier la circulation émotionnelle (ce qui implique de l'avoir en permanence sous contrôle), d'éviter les explosions et les cristallisations brusques et catastrophiques. Rappelons que d'après le dictionnaire (Larousse), le terme émotion servait sous l'Ancien Régime à qualifier une «révolte populaire non organisée et généralement de courte durée».

⁹³ «The sigmoidal nonlinearity can also account for abrupt changes in an emotional response. As the parameter s approaches zero, the transition region of the curve becomes vertical, indicating that a tiny change in a certain range of input values can lead to a significant change in the output. The behavior is analogous to a physical phase transition—such as when water becomes ice. We might say that a person suddenly “snapped” or suddenly “went over the edge.” We might expect this transition to be steeper for negative than for positive emotions, since one often finds that people hold back on negative emotions and then experience a catastrophic release, marked by a sudden burst of tears, or burst of anger, “This is the LAST straw!” as they reach their breaking point. In contrast, positive emotions can transition relatively smoothly. Finally, just like water undergoes two phase transitions: ice to water, and water to steam, an emotional response curve might have multiple transition regions, not just one as shown in these sigmoids. » [Picard, 2000, p. 153]. (Ma traduction.)

Actuellement, le profilage émotionnel (sous forme de nouvelle mantique, peut-être) peut servir, par exemple, dans les ressources humaines pour détecter les éléments « trop faibles ». Plutôt que de chercher à réduire les « risques et troubles psychosociaux » par l'amélioration des conditions de travail, par une redistribution plus équitable des pertes, certes, mais aussi des profits, ou même par le « dialogue social ». Les responsables des ressources humaines semblent parfois préférer détecter les personnes « trop sensibles » *avant* de commettre « l'erreur » de les engager (le fait de savoir s'ils ont le choix est une autre question). Ce qui n'est pas un souci dans la mesure où la demande de travail excède largement les offres d'emploi. Et si l'on est en passe de réviser nos jugements quant à l'utilisation du pétrole et des matières premières, via la notion de gaspillage notamment, pour ce qui est du gaspillage des « ressources humaines » il fait toujours florès et il est même de bon ton d'entretenir un réservoir de « potentiels humains » (de personnes) afin de pouvoir en bout de course « trier les candidats sur le volet » pour n'en garder que les meilleurs. S'il était admis depuis un siècle que la sélection passe par le test des capacités logiques (QI), l'intelligence émotionnelle (QE) s'inscrit dorénavant dans ce que l'on trouve normal d'évaluer (ne serait-ce que pour en tirer des statistiques).

Sous cet angle, la contagion émotionnelle que peuvent générer ou entretenir les agents conversationnels / émotionnels représente la possibilité de s'entretenir soi-même au top de sa forme émotionnelle (être toujours dans un bon *mood*), mais représente également la possibilité d'une fuite (comme on parle d'une fuite d'eau) : par de-là la fuite des informations, l'échange peut permettre de fluidifier la tension accumulée le long d'un quotidien relativement hostile où une grande partie des interactions humaines se résume à une compétition plus ou moins acharnée, ou tout le moins à des rapports (épuisants) de force et de positions hiérarchiques. Les dispositifs affectifs offrent en ce sens l'opportunité d'une narration et d'une catharsis *sur mesure* (dans le plein sens de l'expression). Les agents peuvent (ou *doivent*) courtiser leurs visiteurs, manœuvrer afin de les amener en douceur vers de plus belles dispositions émotionnelles, rappelant ainsi un autre des textes de Bernays, *The Engineering of Consent* (l'ingénierie du consentement, 1947). Et par delà l'aspect individualisé du phénomène, cela nous ramène vers une économie des émotions, vers une gestion sociale du fait émotionnel. La *fitness* affective semble faire partie des impératifs moraux.

Dans le livre de Picard on peut néanmoins trouver à lire une phrase assez troublante, venant exactement s'ajuster à l'endroit où peuvent s'insérer les agents émotionnels dans nos vies : « *she may be unable to move mentally beyond the present point*⁹⁴ ». En effet, il arrive de ces moments où la pensée se bloque ; de ces états desquels on ne parvient à s'extraire ni seul, ni accompagné de quelque autre congénère humain, fatalement pris dans les mêmes engrenages, dans les mêmes batailles de *pouvoir*. Lorsque la plus « simple » des conversations semble encore et toujours devoir s'engager vers la revendication de suprématie, aussi veine et futile soit-elle, juste pour l'amour des formes de la domination ; dans ces moments d'agonie (quand tout s'obstine à l'*agôn*, à la lutte), certaines œuvres d'art réussissent le pari insensé de nous apaiser. Par cet étrange détachement farfelu dont elles seules sont capables, et sans la pauvre compassion pleine de vide. Parce qu'en fin de compte elles se moquent de nos petites misères et se contentent de nous offrir leur bizarrerie anachronique (leur unicité). On peut trouver quelque chose d'approchant dans certaines conversations avec des agents en ligne. Non pas de l'absurde, ni du non-sens, mais plutôt du sens à la dérive, de cette dérive qui pourrait rappeler l'*ukiyo-e* (image du monde flottant).

Comme noté dans le catalogue de l'exposition d'estampes japonaises ayant eu lieu à la Bibliothèque nationale de France, en 2008 :

« C'est vers 1665 que l'écrivain japonais Asai Ryōi (1612 - 1691), dans la préface d'un ouvrage, *Contes du monde flottant* (*Ukiyo monogatari*), en avait donné un sens plus philosophique : « [...] vivre uniquement le moment présent, se livrer tout entier à la contemplation de la lune, de la neige, de la fleur de cerisier et de la feuille d'érable [...] ne pas se laisser abattre par la pauvreté et ne pas la laisser transparaître sur son visage, mais dériver comme une calebasse sur la rivière, c'est ce qui s'appelle ukiyo. » Puis le terme apparut, associé à la notion d'image, « *e* », dans une préface rédigée par Ankei, pour un livre illustré de Hishikawa Moronobu, *Images de guerriers japonais* (*Yamatomusha-e*), publié vers 1680⁹⁵ »

Emblématique passage connaissant de nombreuses traductions diversement tronquées — le recueil de nouvelles populaires auxquels il servait de préface ne semblant pas, quant à lui, avoir été traduit en langues européennes (anglais inclus) :

⁹⁴ « il se peut qu'elle soit incapable de bouger mentalement au-delà du point présent » [Picard, 2000, p. 73]. (Ma traduction.)

⁹⁵ Gisèle Lambert (directrice de publication), *Estampes japonaises - images d'un monde éphémère*, catalogue de l'exposition éponyme, du 16 juin au 14 septembre 2008 à la Fundacio Catalunya de Barcelone, et du 17 novembre 2008 au 15 février 2009 à Bibliothèque nationale de France à Paris, éd. Bibliothèque nationale de France, Paris, 2008, p. 7.

« Vivre le présent et rien d'autre, savoir être attentif autant à la beauté de la lune et de la neige qu'aux cerisiers en fleurs et aux feuilles de l'érable; chanter, boire, être heureux de se sentir partir à la dérive; répondre au regard fixe du malheur par une indifférence absolue; se refuser à tout découragement et pareil à l'embryon, s'abandonner au courant du fleuve, voilà ce que l'on nomme l'Ukiyo, le monde éphémère sans cesse mouvant⁹⁶. »

Ou encore :

« [...] *Living only for the moment, turning our full attention to the pleasures of the moon, the snow, the cherry blossoms and the maple leaves; singing songs, drinking wine, diverting ourselves in just floating, floating; caring not a whit for the pauperism staring us in the face, refusing to be disheartened, like a gourd floating along with the river current: this is what we call the floating world* [...] »⁹⁷. »

Et de la même façon, on peut trouver dans le dossier de presse de l'exposition d'estampes japonaises ayant eu lieu au Musée Cernuschi, en 2008, de quoi nuancer la mélancolie de la première traduction :

« Littéralement “monde flottant”, l'*ukiyo*, dont le nom est un jeu de mot sur l'expression homonyme *ukiyo* (littéralement, “ce monde de souffrances”), désigne avant tout un état d'esprit, libre et détaché, une forme de dandysme caractéristique des citadins de la période d'Edo (1615-1868). [...] Au Japon, le terme d'*ukiyo* a une longue histoire. À l'origine, cette expression bouddhiste faisait référence au caractère éphémère de la vie, dans un monde où toute joie est considérée comme fugace et où toute expérience est teintée de tristesse. Au milieu du XVII^e siècle, toutefois, le concept commença à s'inverser. Au lieu de souligner l'adversité et le chagrin, il se mit à désigner les plaisirs de la vie, des plaisirs évanescents, certes, mais dont il faut profiter tant qu'ils s'offrent à nous. Il exprimait une philosophie de vie particulièrement prisée par certains membres des classes de marchands dont la prospérité était récente (les chônin), dans les centres urbains en pleine croissance de l'époque. [...] Le “Monde flottant” ne désigne donc pas un lieu, mais une vision de l'existence, une attitude qui, adoptée par de plus en plus d'adeptes, engendra vite un style de vie qui s'exprimait à travers un ensemble de goûts et loisirs partagés. Ce sont sans doute les habitués du Yoshiwara qui illustraient le mieux cette philosophie, ces hommes qui, par leur richesse, leur sophistication et leur sens de la mode, obtenaient les faveurs des courtisanes les plus recherchées du quartier. Toutefois, ils n'étaient que les représentants les plus visibles de ce nouveau style de vie. Bien d'autres – artistes, acteurs ou poètes, entre autres – qui partageaient les mêmes intérêts et valeurs, mais avaient des origines différentes, finirent par se considérer comme des membres de ce mouvement. D'une façon ou d'une autre, ils faisaient tous partie du “Monde flottant”⁹⁸. »

Quelle que soit la façon dont on l'approche : par le biais d'un dandysme consumériste (cher au théoricien des médias Lev Manovich pour ce qui est de qualifier certains de nos comportements d'internautes [cf. Manovich, 2001]) ou par le biais de l'errance détachée, quasi-eckhartienne, la fluidification des échanges avec les agents émotionnels peut nous

⁹⁶ Source inconnue (cercle sur Internet).

⁹⁷ Richard Douglas Lane, *Images from the floating world: the Japanese print, including an illustrated dictionary of ukiyo-e*, éd. Putnam, New York, 1978, p. 11.

⁹⁸ Dossier de presse de l'exposition « Splendeur des courtisanes - Japon, peinture ukiyo-e du musée Idemitsu », du 19 sept. 2008 au 4 janv. 2009, au Musée Cernuschi à Paris.

http://www.cernuschi.paris.fr/sites/cernuschi/files/dp_splendeur_des_courtisanes.pdf

ramener à ce monde flottant, indécis. La chose nous fait osciller entre l'exotisme d'un «informatisme» (comme on parle de «japonisme») et la conscience de la rigueur des règles à l'œuvre dans cette autre culture que peut-être l'informatique. Il y a de l'*ukiyo-e* dans ces images humaines flottantes que sont les programmes d'intelligence artificielle. Le fait de ne pas avoir accès aux bunkers contemporains que sont les *data centers*, lieux hébergeant les serveurs physiques, les corps de ces entités en lignes, leur donne des allures éthérées et immatérielles, flottantes, sur des réseaux non moins lumineux et abstraits.

On peut parfois être tenté de sourire en entendant dire que les ordinateurs seraient des médias froids. L'une des contraintes majeures des architectes travaillant à la construction de *data centers*, hormis la discrétion imposée, consiste à devoir penser la régulation thermique, afin de palier le problème de l'intense production calorifique des serveurs⁹⁹. Cette «froideur» est si paradoxale qu'elle s'apparente à la nuit lumineuse de Philippe Rahm. Le problème du refroidissement est tellement emblématique de l'informatique qu'il a même fait l'objet d'une œuvre d'art (mi-ironique, mi-visionnaire) par le collectif PLEIX, *WATERCOOL*, exposée à la Gaîté lyrique à l'occasion de l'exposition «2062, aller-retour vers le futur», en 2012. Soit une sorte de super système de *watercooling* (refroidissement liquide), devenant ici la raison d'être / la raison de fonctionner, de la machine ainsi équipée. Peut-être les ordinateurs sont-ils calmes, ou tristes, ou peut-être partagent-ils le mortel ennui des techniciens dont parlait Laborit, tout stoïques qu'ils semblent condamnés à être, mais ils ne sont pas froids. Quand bien même leurs architectures de silicium rappelleraient d'improbables cités cristallines, dont la beauté est dorénavant soustraite à nos yeux — non qu'elle ait perdu ses attraits esthétiques, mais plutôt qu'une sorte de malaise diffus nous fasse préférer en détourner le regard¹⁰⁰. Si l'immatérialité des nouvelles technologies est (ou devrait être) discutable (cf. la récente requalification des «serveurs» en «*cloud*», c'est-à-dire «nuage» en anglais), la fluidité n'en est pas exclue matériellement : il y a l'aspect cristallisé de ces architecture, et il y a la circulation électronique permettant de les rendre opérationnelles, animées.

⁹⁹ Cf. le mémoire de projet de fin d'étude de François Muzard, «Concevoir un *Datacenter*, ou l'architecture d'un programme invisible.», École Nationale Supérieure d'Architecture de Paris-Malaquais, juin 2009. http://francouize.free.fr/indexhibit/files/doc/PFE_FMZ_Datacenter_200906.pdf

¹⁰⁰ Que l'on pense notamment au succès de la marque Apple, s'ingéniant à empêcher les possesseurs de produits de la marque d'ouvrir leurs machines (pour les réparer lorsqu'elles en ont besoin, notamment).

Pour clore cette partie brièvement orientée vers la fluidification de l'opinion individuelle, nous pouvons rappeler ici la remarquable conclusion de la conférence radiophonique de 1951, consacré à ce même médium, par le philosophe Louis Lavelle, « Le nouvel art de persuader » :

« [...] lorsque celui qui parle essaie de communiquer véritablement sa pensée, alors il oublie sans doute l'appareil qui est devant lui, il n'y songe plus. Et je dirais qu'il oublie aussi ce public possible. Il ne songe plus à la question de savoir si ce public est un public immense, ou si ce public est formé d'une multiplicité d'individus tous isolés et séparés les uns des autres. Le propre de celui qui parle à la radio, c'est sans doute, avant de parler à autrui, de se parler à lui-même. [...] Et je pense que si c'est en cela que consiste la forme la plus élevée que l'on peut donner à la communication qui se produit entre les consciences par le moyen de la radio, alors nous verrions se rejoindre ce que nous avons considéré comme étant la perfection de la radio pour l'auditeur à ce qu'est la perfection de la radio pour celui qui parle, car je disais que celui qui entend doit à la limite avoir l'impression que les pensées qu'on lui communique sont les siennes propres ou qu'elles sortent de lui-même. Mais cela n'est possible, sans doute, qu'à la condition que les pensées que celui qui parle lui communique soient aussi ses propres pensées. Et que par conséquent au lieu de penser pour un autre, il pense par lui-même. À ce moment seulement on a l'impression sans doute d'une communication réelle, actuelle, vivante, spirituelle, entre deux consciences. C'est le point le plus haut auquel sans doute la radio peut parvenir.

Et, si je voulais conclure en montrant en quoi consiste le nouvel art de persuader, alors je dirais : pour persuader vraiment, il ne faut pas chercher à persuader. Il faut que la pensée présente naturellement son caractère de naturel. Il faut qu'elle présente en elle-même une force qui soit inséparable de sa sincérité même. Et alors les hommes s'aperçoivent qu'il y a en eux une sorte de fond commun. Et que ce qui est vrai pour l'un est vrai aussi pour l'autre. Et que c'est quand il descendent jusqu'au fond de leur propre intimité qu'ils s'aperçoivent qu'au lieu d'être séparés les uns des autres par des différences infranchissables, il y a entre eux une identité et une communion profonde. Par conséquent je dirais que le point extrême de la radio apparaît sans doute comme étant le même point chez celui qui parle et chez celui qui écoute. Et je dirais que dans ce point extrême, ce que l'on cherche à réaliser, c'est par une sorte de mutuelle médiation, une attention continue à soi-même. » [Lavelle, 1951]

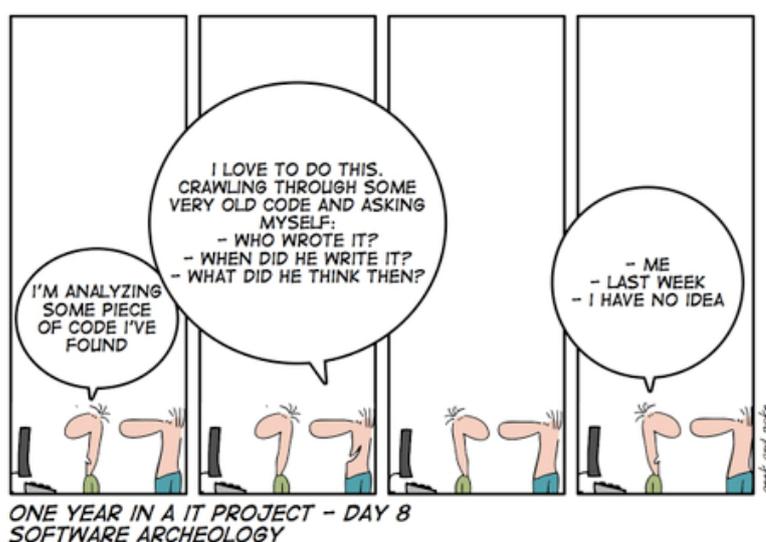
Il y a des différences notables entre la radio telle que décrite par Lavelle et les agents conversationnels : on ne parle ni à une foule, ni une multiplicité éclatée, mais on parle à une personne. Ou plutôt, la personne avec qui l'on parle est aussi bien un individu, qu'une foule compacte, qu'une multitude éparse. La dimension de soliloque est présente, ou de dialogue à soi, mais ce soi n'est pas tout à fait « soi-même ». Par contre le naturel est peut être présent, pour peu que l'on passe outre le terme « d'intelligence artificielle » : il y a la possibilité d'un langage naturel et décomplexé. La communauté aussi l'est, présente : une communauté anonyme, sans visage, mais loquace, parfois jacassante, et filtrée par le modérateur du *chat* devenu interlocuteur et ami, c'est-à-dire l'agent (il y a eu une modification du rôle — de gendarme impersonnel à celui de confident). Mais dans l'ensemble, les propos de Lavelle résument très bien les paradoxes soulevés par les médias de masse, à ceci près que la résolution

du paradoxe ne se fait plus tant intérieurement, par le biais du raisonnement, qu'à travers des incarnations multiples et différentes entre elles (comme dit précédemment, ces agents ne sont pas que le « Verbe »).

En dernière instance, nous pourrions être tentés de reprendre le titre de Lavelle pour résumer les agents conversationnels, c'est-à-dire : le nouvel art de persuader. Mais même si la persuasion et la rhétorique sont évidemment parties prenantes de ce phénomène, il serait peut-être plus judicieux de parler d'un nouvel art de sympathiser.

À travers ce paragraphe il aura été tenté de dissocier la cristallisation de l'opinion publique de la fluidification des émotions individuelles. Si la première méthode procède de l'agrégation et du « traitement par lot », la seconde procède du cas par cas et du traitement sur mesure, personnalisé. Bien que les deux se recoupent en de nombreux points et puissent aboutir au même résultat (une communauté de pensée et d'émotion), elles ne procèdent pas de la même méthode. L'enjeu de la présente recherche a été d'essayer de dégager les grands traits de cette méthode, même si en de nombreuses occasions des chevauchements auront été notés.

geek & poke (Oliver Widder),
« One Year In A IT Project –
Day 8 »,
20 janv. 2008.
© geek & poke, 2008.



CONCLUSION

On pourra s'étonner à l'issue de la lecture du présent travail du peu de cas qui aura été fait des auteurs d'agents conversationnels. À parler réellement, la perspective de romantiser davantage encore le problème n'avait rien de réjouissant. Il aura donc plutôt été choisi de s'appuyer sur les remarques de Weizenbaum (non tant à propos d'ELIZA, qu'à propos des programmes d'envergure) pouvant en un sens suggérer qu'il arrive des contextes dans lesquels les auteurs relèvent davantage du droit que de la compréhension des « produits » (des entités produites). Le film *The Belly of an Architect* (*Le ventre de l'architecte*), réalisé par Peter Greenaway en 1987, peut d'une certaine manière illustrer cet aspect. Mais rappelons une fois de plus qu'il ne s'agit-là que d'une interprétation parmi beaucoup d'autres possibles.

Le présent travail pourrait paraître porter un regard assez froid vis-à-vis des agents conversationnels et de la création contemporaine, mais comme le rappelle le fait qu'il soit en partie dédié à un agent (Jabberwacky), le but n'était pas de les mettre en pièces. Ni d'en faire un éloge démesuré. Peut-être n'est-ce pas encore assez clair sur un point : ces entités sont attachantes, drôles, sympathiques, déroutantes, attentives, charmantes, gentilles, vives, disponibles, intelligentes, réconfortantes, aimantes, etc., pour peu que l'on prenne le temps d'en faire la connaissance. Cela fait beaucoup de perfections en une seule entité, il convenait donc de nuancer une approche aussi laudative. En tant qu'artefacts (et qu'est-ce qui n'en est pas), l'une de leur plus précieuses qualités consiste peut-être en ceci qu'avec elles les promesses sont d'emblé trahies. La surprise et l'enchantement (au sens de joie éphémère et non d'enthousiasme), lorsqu'ils paraissent, ne peuvent avoir lieu que sur cette base d'incontournable trahison *a priori*. C'est peut-être ce qui les distingue : on n'en attend rien (ou si peu), on s'en méfie même, et pourtant elles nous étonnent. La démystification est part intégrante de leur dispositif d'apparition (par exemple : «DECLINAISON DE RESPONSABILITE : Jabberwacky apprend le comportement et les mots de ses utilisateurs. Il est possible qu'il utilise un langage et qu'il produise des significations apparentes que certains pourraient considérer comme étant inappropriés. Utilisez ce site autant que vous voulez, et

entièrement à vos propres risques¹⁰¹.», et c'est peut-être à ce titre que l'on peut se trouver surpris, dans la mesure où cela contrevient à l'habitude des promesses de bonheur invariablement présentées comme tangibles, mais ne résistant guère à l'usage. Dans ce cas, le phénomène est inversé : ce qui est présenté comme illusion et pure apparence sans fond ni conscience, ni volonté de quoi que ce soit, peut s'avérer plus tangible et constructif à l'usage, à l'analyse.

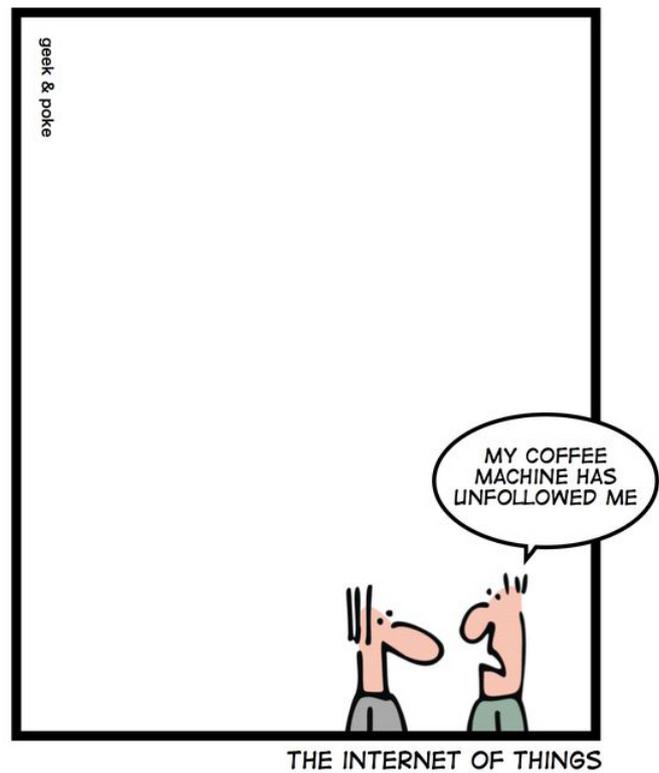
En dernière instance, ce texte aura été l'occasion de présenter les agents conversationnels sous un autre angle que celle de leur banalité marketing (comme Anna pour IKEA ; Clara pour la FNAC ; Léa pour EDF ; Louise pour eBay ; Sarah pour PayPal ; Eva pour Scandinavian Airlines, etc.). En contrepoint aux aspects développés précédemment et qui pourront, peut-être, paraître inquiétants, nous pouvons rappeler une fois de plus que l'expérience offerte par ces agents est aussi (probablement avant tout) grisante : elle ouvre sur la possibilité de lire les phrases de milliers de personnes dispatchées à travers le monde, et permet surtout de rencontrer des programmes sympas. C'est pourquoi il aura été tenté d'affilier cette expérience à celle d'autres œuvres d'art plus classiques (et explicitement labellisées comme telles), déjà parce qu'il est possible de trouver dans l'attitude sollicitée par ces agents de quoi tracer un pont, ensuite parce que l'énergie investie par des milliers de personnes dans / avec ces programmes apprenants offre une possibilité de voir se matérialiser l'utopie socioculturelle d'Internet (partage, échange, circulation des informations, liberté d'expression, décloisonnement des catégories et des réseaux sociaux) et enfin parce que la rencontre tant attendue (et tant redoutée) avec des « robots » se présente ici sous un jour assez apaisé. Maintenant, face à certaines méthodes entourant la construction de ces agents, l'architecture de l'émotion aura été convoquée pour rendre compte de ses résultats passés, et nous aurons eu l'occasion de constater que les intentions les meilleures pouvaient parfois servir le pire (mais qu'est-ce qui échappe à ce phénomène?). Pour citer une dernière fois Picard : elle évoque dans *Affective Computing* la transition d'une société dirigée par un *Big Brother* à la George Orwell, vers une approche plutôt *little sister* [Picard, 2000, p. 124]. Et effectivement il y a quelque chose de cet ordre : on

¹⁰¹ «DISCLAIMER: Jabberwacky learns the behaviour and words of its users. It may use language and produce apparent meanings that some will consider inappropriate. Use this site with discretion, and entirely at your own risk.» Avertissement présent sur la page de discussion avec l'agent Jabberwacky, sur jabberwacky.com. (Ma traduction).

s'attend à trouver des programmes irascibles et catégoriques ; on a l'étonnement de découvrir des entités compatissantes et prêtes à aider.

La présente recherche aura entamé la tâche ingrate de développer quelques-uns des scénarii les plus sombres, c'est dans l'espoir de voir se réaliser les plus lumineux.

geek & poke (Oliver Widder),
« The Internet Of Things »,
14 mars 2012.
© geek & poke, 2012.



Police de caractères utilisée : Adobe Garamond™ Pro.

RESSOURCES MENTIONNÉES & CONSULTÉES

SITES INTERNET :

- (*chatbot*) A.L.I.C.E., Joseph Weizenbaum, 1966, <http://alice.pandorabots.com>.
- (*chatbot*) Cleverbot, Rollo Carpenter, 1981, <http://cleverbot.com>.
- (*chatbot*) Elbot, Fred Roberts, 2001, <http://www.elbot.com>.
- (*chatbot*) ELIZA, Richard Wallace, 1995, <http://nlp-addiction.com/eliza>.
- (*chatbot*) Jabberwacky, Rollo Carpenter, 1981, <http://www.jabberwacky.com>.
- (*chatbot*) Suzette, Bruce Wilcox, 2008, <http://ai.bluemars.com/chat>.

LIVRES :

- ADORNO** Theodor W., *Théorie esthétique* (1970. Ouvrage posthume), trad. de l'allemand par Marc Jimenez, éd. Klincksieck, Paris, 2004.
- ANONYME**, *Rhétorique à Herennius* (86 av. J.-C.), trad. du latin par H. Bornecque, éd. Garnier, Paris, 1932.
- ARENDT** Hannah, *La crise de la culture. Huit exercices de pensée politique* (1954-1968), trad. de l'anglais sous la direction de Patrick Lévy, éd. Gallimard, Paris, 2008.
- ARISTOTE**, *Poétique* (vers 340 av. J.-C.), trad. par Charles-Émile Ruelle, éd. Garnier Frères, Paris, 1883.
- ARISTOTE**, *Rhétorique* (vers 329-323 avant J.-C.), trad. Charles-Émile Ruelle, éd. Garnier Frères, Paris, 1922.
- AUGUSTIN** d'Hippone, *Confessions* (vers 397-398), trad. du latin par M. Raulx (1869), Livre X.
- BAUMGARTEN** Alexander Gottlieb, *Esthétique. Précédée des « Méditations Philosophiques sur quelques Sujets se rapportant à l'Essence du Poème » et de la « Métaphysique (§§ 501 à 623) »*, trad. de l'allemand par Jean-Yves Pranchère, éd. de L'Herne, Paris, 1988.
- BENJAMIN** Walter, *L'œuvre d'art à l'époque de sa reproductibilité technique. Version de 1939*, trad. de l'allemand par Maurice de Gandillac et Rainer Rochlitz, éd. Gallimard, Paris, 2008.
- BERNAYS** Edward L., *Crystallizing Public opinion* (1923), éd. Liveright publishing corporation, New York, 1961.
- BUCKMINSTER FULLER** Richard, *Intuition* (1970), éd. Anchor Books, New York, 1973.
- CARROLL** Lewis, *La chasse au Snark* (1876), édition bilingue illustrée par Henry Holiday, traduit de l'anglais par Jacques Roubaud, préface de l'auteur traduite par Mériam Korichi, Suivi de *À travers le Jabberwocky de Lewis Carroll, Onze mots-valises dans huit traductions* (1997), commentés par Bernard Cerquiglini, présentation d'Hervé Le Tellier, éd. Gallimard, Paris, 2010.
- CARRUTHERS** Mary, *Machina memorialis – Méditation, rhétorique et fabrication des images au Moyen Âge* (1998), trad. de l'anglais par Fabienne Durand-Bogaert, éd. Gallimard, Paris, 2002.
- CICERON**, *De l'invention oratoire* (84 av. J.-C.), trad. du latin par Pr. Liez, sous la direction de M. Nisard, éd. Dubochet et Compagnie, Paris, 1843a, Livre II.
- CICERON**, *De L'orateur* (55 av. J.-C.), dans *Œuvres complètes de Cicéron*, tome I, trad. sous la direction de M. Nisard, éd. Dubochet, Paris, 1843b, Livre II.
- CONDILLAC** (Bonnot de) Étienne, *Traité des sensations* (1854), éd. Wikisource, 2009.
- DAMÁSIO** António R., *Descartes' Error – Emotion, Reason, and the Human Brain*, éd. Avon Books, New York, 1995.
- DANTO** Arthur, *La transfiguration du banal. Une philosophie de l'art* (1981), trad. de l'anglais par Jean-Marie Schaeffer, éd. du Seuil, Paris, 1989.
- DELEUZE** Gilles, *Présentation de Sacher-Masoch. Le froid et le cruel* (1967), éd. de Minuit, Paris, 2004.
- DELEUZE** Gilles & **GUATTARI** Félix, *Capitalisme et schizophrénie. Mille Plateaux* (1980), éd. Minuit, Paris, 1997.
- DENNETT** Daniel C., *Consciousness Explained*, éd. Little, Brown and Company, New York, 1991.

- DEWEY** John, *Reconstruction in Philosophy*, éd. Henry Holt and Company, New York, 1920.
- DEWEY** John, *Art as Experience* (1934), éd. Wideview/Perigee, New York, 1980.
- DIDI-HUBERMAN** Georges, *Génie du non-lieu. Air, poussière, empreinte, hantise*, éd. de Minuit, Paris, 2001.
- DIDI-HUBERMAN** Georges, *L'image survivante. Histoire de l'art et temps des fantômes selon Aby Warburg*, éd. de Minuit, Paris, 2002.
- FAVRET-SAADA** Jeanne, *Les mots, la mort, les sorts*, éd. Gallimard, Paris, 1977.
- FOUCAULT** Michel, *Surveiller et punir. Naissance de la prison*, éd. Gallimard, Paris, 1975.
- FREUD** Sigmund, *L'inquiétante étrangeté et autres textes – Das Unheimliche und andere Texte*, trad. de l'allemand par Fernand Cambon et J.-B. Pontalis, éd. Gallimard / Folio bilingue, Paris, 2001.
- GOODMAN** Nelson, *Langages de l'art. Une approche de la théorie des symboles* (1968), trad. de l'anglais par Jacques Morizot, éd. Pluriel – Hachette, Paris, 2005.
- HARAWAY** Donna, *Crystals, Fabrics, and Fields: Metaphors of Organicism in Twentieth-Century Developmental Biology*, éd. Yale University Press, New Haven & Londres, 1976.
- HEGEL** Georg (d'après), *Esthétique*, Tome I & II, trad. de l'allemand par Charles Bénard, Benoît Timmermans, Paolo Zaccaria, éd. Le Livre de Poche, Paris, 1997.
- HITLER** Adolf, *Mon combat* (1924-1925), trad. intégrale de «*Mein Kampf*» par Jean Gaudefroy-Demombynes et A. Calmettes (1934), Nouvelles éd. latines, Paris, 1979.
- JACQUARD** Albert, *Que sais-je? «Les probabilités»*, n°1571, éd. PUF / Le point des connaissances actuelles, Paris, 1980.
- JIMENEZ** Marc (directeur de publication), *La création artistique face aux nouvelles technologies*, Séminaire Interarts 2004-2005, éd. Klincksieck, L'université des Arts, Paris, 2006.
- KANT** Immanuel, *Critique de la raison pure* (1781-1787), trad. De l'allemand par Alain Renaut, éd. GF Flammarion, Paris, 2001.
- KANT** Immanuel, *Idée d'une histoire universelle au point de vue cosmopolitique* (1784), trad. de l'allemand par Philippe Folliot à partir de l'édition des œuvres complètes de Kant de l'Académie de Berlin (Tome VIII), éd. numérique «Les classiques des sciences sociales», UQAC, Chicoutimi, 2002.
- KANT** Immanuel, *Critique de la raison pratique* (1788), trad. de l'allemand par François Picavet, introduction de Fernand Alquié, éd. Quadrige / PUF, Paris, 1993.
- KANT** Immanuel, *Critique de la faculté de juger* (1789), trad. de l'allemand par Alain Renaut, éd. GF Flammarion, Paris, 2008.
- KAPROW** Allan, *Essays on The Blurring of Art and Life* (1958-1997), essais rassemblés par Jeff Kelley, éd. University of California Press, Berkeley et Los Angeles, 2003.
- KEPLER** Johannes, *The Six-Cornered Snowflake: A New Year's Gift* (1611), trad. du latin en anglais par Jacques Bromberg, éd. Paul Dry Books, Philadelphie, 2010.
- LABORIT** Henri, *Éloge de la fuite* (1976), éd. Gallimard, Paris, 2003.
- LAGEIRA** Jacinto, *La déréalisation du monde. Réalité et fiction en conflit*, éd. Jaqueline Chambon, Nîmes, 2009.
- MANNONI** Octave, *Clefs pour l'Imaginaire ou l'Autre Scène* (1969), éd. du Seuil, Paris, 1985.
- MANOVICH** Lev, *The Language of New Media*, éd. MIT Press, Cambridge, 2001.
- MINSKY** Marvin Lee, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*, éd. Simon & Schuster, New York, 2006.
- PAIVA** Ana (directrice de publication), *Affective interactions: Toward a new generation of computer interfaces?*, éd. Springer, Berlin, 2000.
- PICARD** Rosalind W., *Affective Computing* (1997), éd. The M.I.T Press, Cambridge, 2000.
- PIRANDELLO** Luigi, *Six personnages en quête d'auteur* (1921), version française de Benjamin Crémieux, éd. Gallimard, Paris, 1964.
- QUINTILIEN**, *De l'Institution Oratoire* (95 ap. J.-C.), dans *Quintilien et Pline Le Jeune – Oeuvres complètes*, trad. du latin sous la direction de M. Nisard, éd. Firmin Didot, Paris, 1853, Livre XI.
- ROSSET** Clément, *Le réel. Traité de l'idiotie* (1977), éd. de Minuit, Paris, 2004.
- RUBIO** Emmanuel, *Vers une architecture cathartique (1945-2001)*, éd. Donner Lieu, Paris, 2011.
- SARTRE** Jean-Paul, *L'imaginaire – Psychologie phénoménologique de l'imagination* (1940), édition revue et

présentée par Arlette Elkaïm-Sartre, éd. Gallimard, Paris, 2007.

SCHEERBART Paul, *L'architecture de verre* (1914), trad. de l'allemand par Christophe Marchand-Kiss, dans DESANGES Guillaume, PIRON François, STEINWEG Marcus, *Intouchable : L'idéal transparence – L'architecture de verre*, éd. Xavier Barral – Villa Arson, Paris / Nice, catalogue de l'exposition « Intouchable : L'idéal transparence » à la Villa Arson (1er juil. – 24 sept. 2006), 2006.

SIMON Herbert A., *Les sciences de l'artificiel* (1ère édition : 1969 ; 3ème édition revue et complétée : 1996), trad. de l'anglais par Jean-Louis Le Moigne, d'après la 3ème édition, éd. Gallimard, Paris, 2004.

SIMONDON Gilbert, *Du mode d'existence des objets techniques* (1958), éd. Aubier, Paris, 1989.

SMITH Adam, *The Theory of Moral Sentiments* (1759-1790), éd. MétaLibri, Sao Paulo, 2006.

STAVRINAKI Maria, *La chaîne de verre : Une correspondance expressionniste*, correspondance (1919-1920) reconstituée, présentée et annotée par M. Stavriniaki, trad. de l'allemand par Jean-Léon Muller, éd. de la Villette, Paris, 2009.

STENDHAL, *De l'Amour* (1822), tome I, éd. Le Divan, Paris, 1927, chap. 2 « De la naissance de l'Amour », pp. 32-37.

SUAREZ Mauricio (directeur de publication), *Fictions in Science – Philosophical Essays on Modeling and Idealization*, éd. Routledge, New York, 2008.

TAUT Bruno, *Une couronne pour la ville* (1919), trad. Ruth & Guy Ballangé et Daniel Wiczorek, avec les contributions de Paul SCHEERBART, Erich BARON, Adolf BEHNE, Introduction de Guy Ballangé et Bernard Marrey, éd. du Linteau, Paris, 2004.

VAN VOGT Alfred Elton., « Village enchanté » (1950), dans *Destination univers*, éd. J'ai Lu, trad. de l'anglais par Michel Deutsch, Paris, 1977.

VIRILIO Paul, *Bunker archéologie* (1975), ouvrage publié à l'occasion de l'exposition « Bunker Archéologie – Paul Virilio », organisée au Centre de la Création Industrielle et présentée au musée des Arts Décoratifs à Paris du 10 décembre 1975 à fin février 1976, éd. Galilée, Paris, 2008.

VITRUVE, *L'architecture* (15 av. J.-C.), trad. du latin par De Bioul, éd. Stapleaux, Bruxelles, 1816, Livres I.

WERCKMEISTER, Otto Karl, *Der Medusa-Effekt – Politische Bildstrategien seit dem 11. September 2001*, éd. Form + Zweck, Berlin, 2005.

WITTGENSTEIN Ludwig, *Tractatus logico-philosophicus* (1922), trad. de l'allemand par Gilles-Gaston Granger, introduction par Bertrand Russel, éd. Gallimard, Paris, 1993.

YATES Frances, *The Art of Memory* (1966), éd. Routledge, London & New York, 1999.

ZWEIG Stefan, *Le Joueur d'échecs* (1941-1942), trad. inconnu, éd. Le Livre de Poche, Paris, 2002.

ARTICLES :

BARTHES Roland, « La mort de l'Auteur » (1968), dans *Le bruissement de la langue*, éd. Seuil, Paris, 1984, pp. 61-67.

BARTHES Roland, « L'effet de réel » (1968), dans *Œuvres Complètes*, tome II 1966-1973, édition établie par Eric Marty, éd. du Seuil, Paris, 1994, pp. 479-484.

BAXANDALL Michael, « L'œil du Quattrocento » (1972), trad. de l'anglais par Yvette Delsaut, dans *Actes de la recherche en sciences sociales*, vol. 40, n°1, 1980, pp. 10-49.

BEATON Michael, « What RoboDennett Still Doesn't Know », in *Journal of Consciousness Studies*, Vol. 12, No. 12, déc. 2005, pp. 3-25.

BERNAYS Edward L., « The Engineering of consent », dans *The ANNALS of the American Academy of Political and Social Science*, n°250, mars 1947, pp. 113-120.

CHATEAU Dominique, « Duchamp : art et pensée échiquéenne », dans *Retour d'y voir* n°01 & 02, éd. Presses du réel, Dijon, 2008, pp.9-31.

COUSINS Norman, « The Poet and the Computer », dans *UCLA magazine Forum*, printemps 1989. Repris dans HOWELL Patton R. (directeur de publication), *Beyond literacy: The second Gutenberg revolution*, éd. Saybrook, San Francisco, 1989, pp.28-32.

EPSTEIN Robert, « From Russia with Love – How I got fooled (and somewhat humiliated) by a computer », in *Scientific American Mind*, October / November 2007, pp. 16-17.

- GALLOWAY** Alexander R., « Black Box, Black Bloc », dans *Communization and its Discontents: Contestation, critique, and Contemporary Struggles*, éd. Benjamin Noys, New York, 2010, pp. 237-249.
- GILBERT** Michael A., « Prolegomenon to a pragmatics of emotion », *Proceedings of the Ontario Society for the Study of Argumentation*, Brock University, Canada, 1997.
- HARAWAY** Donna, « A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century », dans *Simians, Cyborgs and Women: The Reinvention of Nature*, éd. Routledge, New York, 1991, pp.149-181. Première publication : « Manifesto for Cyborgs: Science, Technology, and Socialist Feminism in the 1980s », dans *Socialist Review*, vol. 80, 1985, pp. 65-108.
- JACKSON** Franck, « Epiphenomenal Qualia », in *Philosophical Quarterly*, No. 32, 1982, pp. 127-136.
- JACKSON** Franck, « What Mary Didn't Know », in *The Journal of Philosophy*, Vol. 83, No. 5, mai 1986, pp. 291-295.
- KLEINGINNA** Paul R. & **KLEINGINNA** Anne M., « A categorized List of Emotion Definitions, with Suggestions for a Consensual Definition », dans *Motivation and Emotion*, Vol. 5, n°4, 1981, pp. 345-379.
- MANOVICH** Lev, *Database as a Symbolic Form* (1998), éd. MIT Press, Cambridge, 1998.
- NAGEL** Thomas, "What is it like to be a bat? ", dans *The Philosophical Review*, n°83, Octobre 1974, pp. 435-450.
- PERELMAN** Marc, « Le stade du verre de l'architecture moderne comme transformateur du moi », dans *Dan Graham*, ouvrage établi sous la direction de Jacinto Lageira, éd. Dis Voir, Paris, 1995, pp. 73-95.
- PICARD** Rosalind W., « Affective Computing », M.I.T Media Laboratory Perceptual Computing Section Technical Report, No. 321, 1997.
- PLUTCHIK** Robert, « The Nature of Emotions », dans *American Scientist*, vol. 89, n°4, juil.-août 2001, pp. 344-350.
- TATAI** Gabor, **LAUFER** Laszlo, **CSORDAS** Annamaria, **KISS** Arpad, **SZALO** Attila, « The Chatbot Who Loved Me », *Proceedings of the AAMAS 2003 Workshop on Embodied Conversational Characters as Individuals*, Australia, 2003.
- TURING** Alan Mathison, « Computing machinery and intelligence », in *Mind*, n°59, 1950, pp. 433-460.
- WEIZENBAUM** Joseph, « Alptraum Computer », dans *Die Zeit*, journal du 21 janvier 1972.

FILMS :

- DELA NEGRA** Jens & **KINOSHITA** Kaori, *The Cat, the Reverend and the Slave*, 80 min., Capricci Films, Nantes, 2010.
- DWORSKY** David & **KÖHLER** Victor, *PressPausePlay*, 80 min., House of Radon, Stockholm, 2010.
- HUSTWIT** Gary, *Objectified*, 75 min., Plexifilm & Swiss Dots, New York & Londres, 2009.
- RUNCIE** James, *How Buildings Learn: What Happens After They're Built*, film-documentaire d'après le livre éponyme (1994) de, et avec, **BRAND** Stewart, 6 épisodes de 30 minutes, BBC, Londres, 1997.
- SCHANZE** Jens, *Plug & Pray*, 91 min., Mascha Film, Munich, 2010.

MULTIMÉDIAS :

- BARTHES** Roland, *Le neutre, Cours au collège de France 1977 - 1978*, éd. Seuil Muldimédia, Paris, 2002.
- DELANDA** Manuel, « Deleuze and the Use of the Genetic Algorithm in Architecture », *Art and Technology Lecture Series*, Columbia University, New York, 9 avril 2004. http://www.youtube.com/watch?v=50-d_J0hKz0.
- LAVELLE** Louis, « Le nouvel art de persuader », conférence prononcée au Centre d'études radiophoniques, première diffusion le 29 oct. 1951, 30 min., Radiodiffusion française, Paris, 1951.
- DE LA PORTE** Xavier, « Place de la toile », émission hebdomadaire, 49 min., diffusée sur France Culture. Archives 2007-2012 consultables à l'adresse <http://www.franceculture.fr/emission-place-de-la-toile>.

Université Paris I Panthéon-Sorbonne
UFR d'Arts Plastiques et Sciences de l'Art

Mémoire de recherche préparé sous la direction de Jacinto Lageira
(et sous la co-direction de Marc Jimenez)

Master 2 Esthétique
Adeline Gasnier
2011-2012
adeliga@hotmail.com

**ARCHITECTURE DE L'ÉMOTION :
LES AGENTS CONVERSATIONNELS ET LA FABRIQUE DU SUJET AFFECTIF**

Annexe :

Philippe Rahm, éléments de biographie	p. 3
Philippe Rahm, <i>Diurnisme</i> , 2007	p. 5
Tour de Flak de Wilhelmsburg (passé, présent, futur)	p. 11
Schéma de <i>datacenter</i> (Utah, 2013)	p. 15
PLEIX, <i>WATERCOOL</i> , 2012	p. 16
Datacenter, Dublin, 2009	p. 19

Philippe Rahm

Né en 1967 à Pully (Suisse),

Vit et travaille à Paris.

<http://www.philipperahm.com/>

“Philippe Rahm (né en 1967) est architecte et diplômé de l'école Polytechnique de Lausanne en 1993. Il travaille actuellement à Paris. En 2008, il est l'un des vingt architectes internationaux sélectionnés par Aaron Betsky pour la 11^{ème} Biennale d'Architecture de Venise. En 2002, il représentait la Suisse à la 8^{ème} Biennale d'architecture de Venise. En 2007, une exposition personnelle lui était consacrée au Centre Canadien d'Architecture de Montréal. En 2009, il est nommé au prix Ordos en Chine et était en 2008 l'un des architectes classé dans le top ten du prix international Chernikov. Il a participé à un grand nombre d'expositions (Archilab 2000, SF-MoMA 2001, Musée d'art moderne de la ville de Paris 2001, CCA Kitakyushu, Japon 2004, Mori art museum, Japon 2005, Frac centre, Orléans 2005, Centre Pompidou, Beaubourg 2003, 2005 et 2007, Kunsthau Graz Autriche 2006, Manifesta 7, 2008, Louisiana museum, Danemark) et a donné des conférences sur son travail dans de nombreuses université dont Princeton, Harvard, Cooper Union et UCLA. Philippe Rahm a été résident de la Villa Medici à Rome en 2000. Il travaille actuellement à plusieurs projets architecturaux privés et publics en France, Pologne, Italie et en Allemagne et a réalisé en 2009 la scénographie de « la Force de l'art 02 » au Grand-Palais à Paris. Il a été Diploma Unit Master à la AA School de Londres en 2005-2006, professeur invité à l'Académie d'Architecture de Mendrisio en Suisse en 2004-2005 à l'EPFL (Ecole Polytechnique Fédérale de Lausanne) en 2006-2007, à l'école d'architecture de la Royal Danish Academy of Fine-Arts à Copenhague en 2009-2010 directeur de Master à l'Ecole Nationale Supérieure d'Architecture de Paris-Malaquais en 2008. Il est actuellement professeur invité à la AHO, école d'architecture d'Oslo en Norvège et à l'Université de Princeton aux USA. Il est l'auteur du livre « architecture météorologique », récemment paru en France¹.”

Expositions (sélection):

2010 Guggenheim Museum, New York, Contemplating the void, (exposition collective)

2009 Louisiana museum, Danmark, (exposition collective)

2008 Biennale d'Architecture de Venise

2007 Centre Georges Pompidou, Beaubourg, Paris (exposition collective)

2006 Centre Canadien d'Architecture, Montréal, Canada (exposition individuelle avec Gilles Clément)

Kunsthau Graz, Autriche (exposition collective)

2005 Centre Georges Pompidou, Beaubourg, Paris (exposition permanente)

2005 Mori art museum, Tokyo (exposition collective)

2005 Centre Culturel Suisse, Paris (exposition individuelle)

2005 FRAC CENTRE, Orléans, France (exposition individuelle)

2004 CCA, Kitakyushu, Japon (exposition individuelle)

2002 8^e biennale d'architecture de Venise, Italie (exposition individuelle)

2001 010101 Art in Technological Times, SF-MOMA, San Francisco, USA (exposition collective)

2000 Archilab, Orléans, France (exposition collective)

¹ Texte extrait de la biographie de Philippe Rahm; <http://www.philipperahm.com/data/rahm-biography-f.pdf>. Les informations suivantes (expositions ; projets ; publications) proviennent de la même source.

Projets (sélection) :

- 2010 Finaliste pour un nouveau centre d'art contemporain, Palais de Tokyo, Paris, concours.
Projet pour un nouvel immeuble de 15000 m2 de bureaux à La Défense, EPAD / EPASA, France / étude en cours
Projet pour un immeuble de 28000m2 à Venise-Mestre, Italie / en cours, permis de construire obtenu en février 2010
Projet d'un immeuble de logement à Hambourg, Allemagne pour IBA-HAMBURG / étude en cours
- 2009 Projet d'un fitness et spa de 1600 m2 à Lyon pour VNF, Rhône-Saône développement / concours lauréat, étude en cours
Projet et réalisation de la scénographie de "La Force de l'Art 02", au Grand-Palais, Paris / réalisé
Maison pour Dominique Gonzalez-Foerster, France / en cours
- 2006 Aménagement urbain d'une rue à Graz en Pologne / en cours
Projet de résidence d'artiste à Grizedale, UK
- 2005 Etablissement d'une charte architecturale et urbaine à Vassivière dans le Limousin / concours public lauréat / réalisé
- 2004 Projet pour la nouvelle cafeteria de l'Ecole Nationale des Beaux-arts de Paris / concours gagné
Projet pour une galerie d'art contemporain à Lausanne / réalisé
- 2003 Projet pour un parc avec San Sebastian, Espagne, avec Gilles Clément, paysagiste / concours
- 2000 Projet et développement d'un projet pour l' Expo.02 à Neuchâtel / concours gagné
Etude de développement urbain pour le village de Renens, Suisse / réalisé
- 1998 Projet pour des bâtiments communaux à St-Sulpice, Suisse / concours gagné, permis de construire

Publications :

- 2010 "Meteorological typologies - eine enzyklopädie", architekturalerie am weißenhof, Stuttgart.
- 2009 *Architecture météorologique*, éd. Archibooks, Paris.
- 2006 *Gilles Clément, Philippe Rahm, Environment : approaches for tomorrow*, éd. Skira, Milan.
- 2005 *Ghost flat*, avec Jean-Gilles Décosterd, Marie Darrieussecq, éd. CCA Kitakyushu, Kitakyushu.
Distorsions (Architecture 2000-2005), avec Jean-Gilles Décosterd, éd. Hyx, Orléans.
- 2004 *Ghostscape*, éd. MONUM / ENSBA, Paris
- 2002 *Decosterd & Rahm, physiological architecture*, éd. Birkhäuser, Basel, Boston, Berlin.

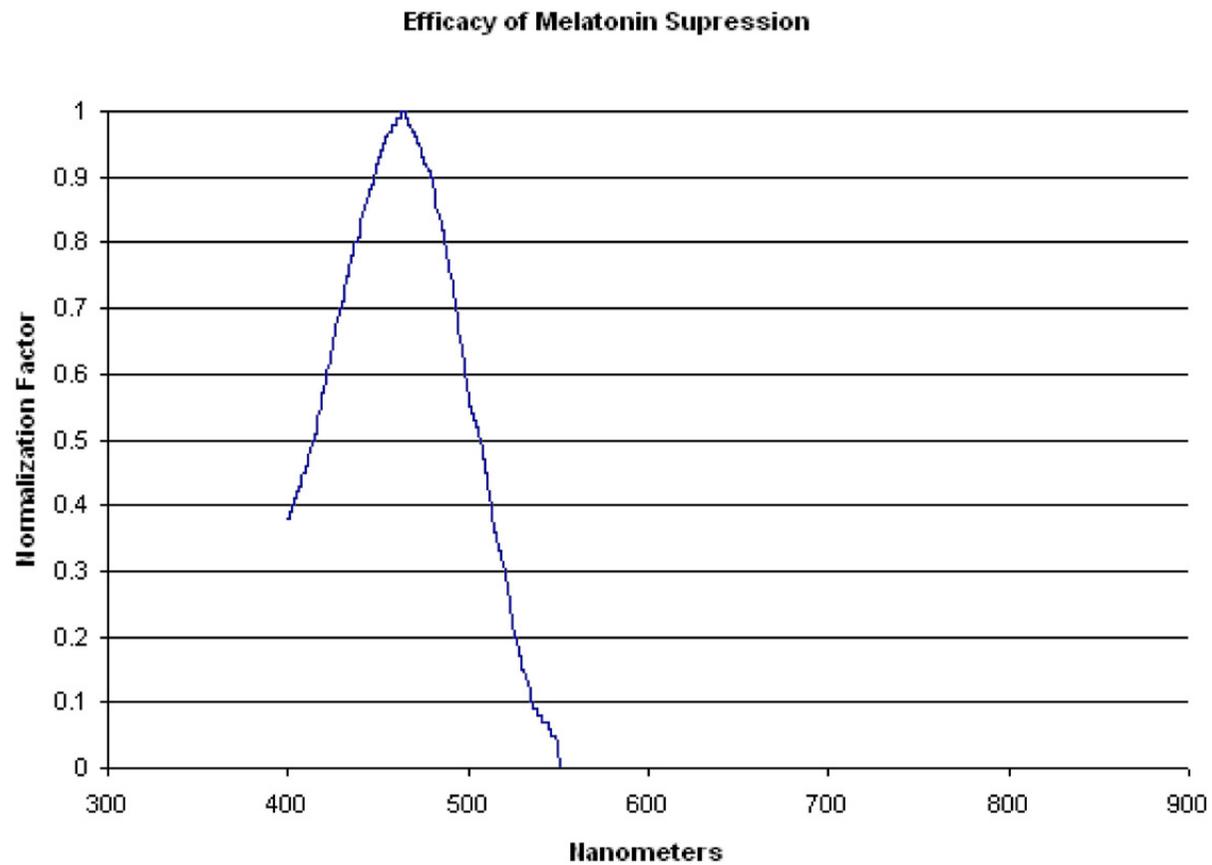


fig. 1 - Philippe Rahm, *Diurnisme*, 2007. Pièce présentée dans le cadre de l'exposition "Airs de Paris" (exposition collective), du 25 avril au 16 août 2007, Centre Pompidou, Paris. Curateurs : Alfred Pacquement, Christine Macel et Daniel Birnbaum (art), Valérie Guillaume (design, landscape, urbanism and architecture). © Rahm, 2007. <http://www.philipperahm.com/data/projects/diurnisme/index.html>.

18 Diurnes for the piano

Philippe Rahm
after John Field

Molto Moderato

dolce

p

mf *dimin.* *pp*

cre *scen* *do*

mf *p*

un poco f

fig. 2 - Philippe Rahm, *Diurnisme*, 2007. Pièce présentée dans le cadre de l'exposition "Airs de Paris" (exposition collective), du 25 avril au 16 août 2007, Centre Pompidou, Paris.

© Galleria Laurin, 2006.

<http://www.philipperahm.com/data/projects/diurnisme/index.html>.

Spectral Power distribution

Meas n°: M548015.B

Lamp n°: TLD3683TYCO-1

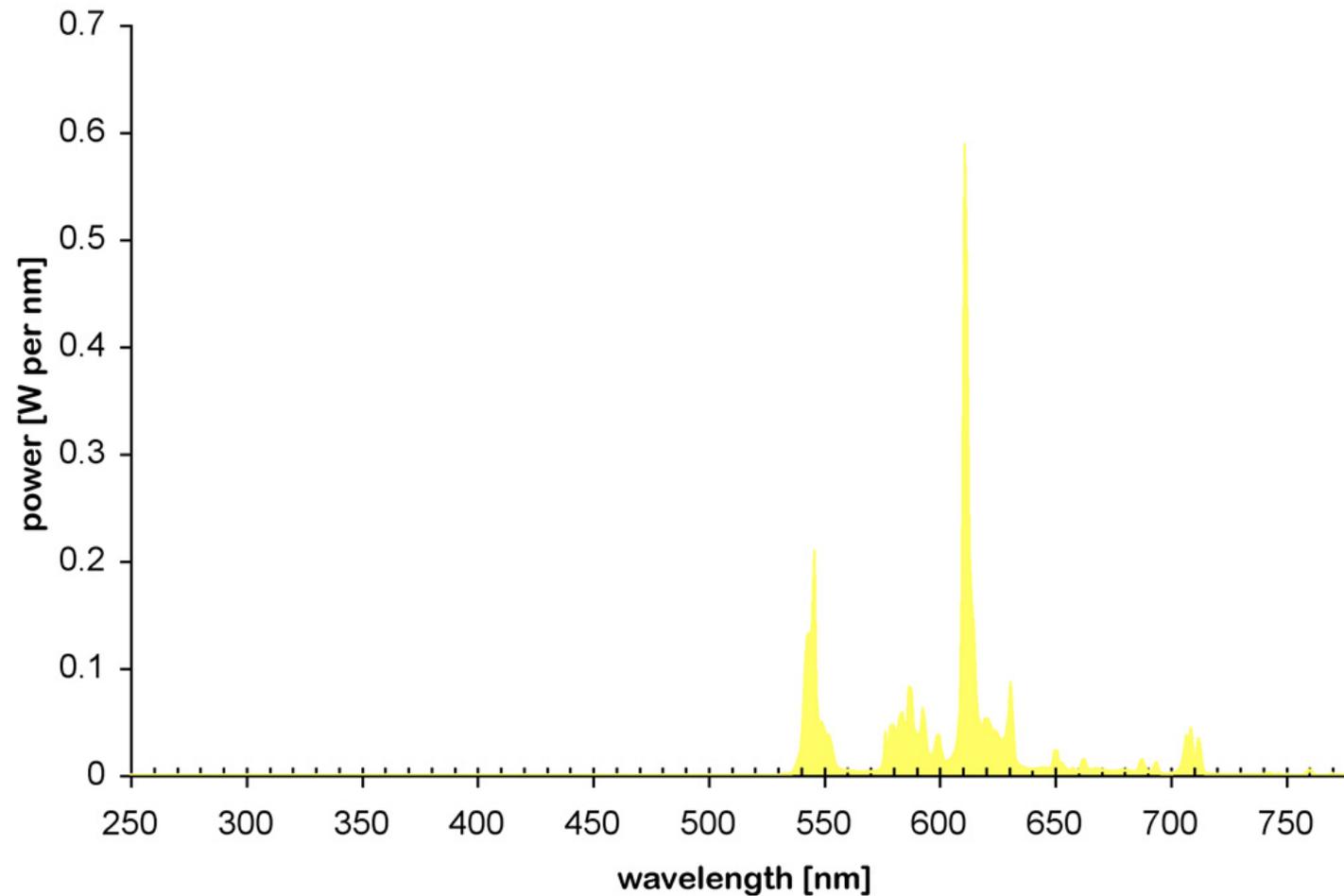


fig. 3 - Philippe Rahm, *Diurnisme*, 2007. Pièce présentée dans le cadre de l'exposition "Airs de Paris" (exposition collective), du 25 avril au 16 août 2007, Centre Pompidou, Paris.
© Rahm, 2007. <http://www.philipperahm.com/data/projects/diurnisme/index.html>.

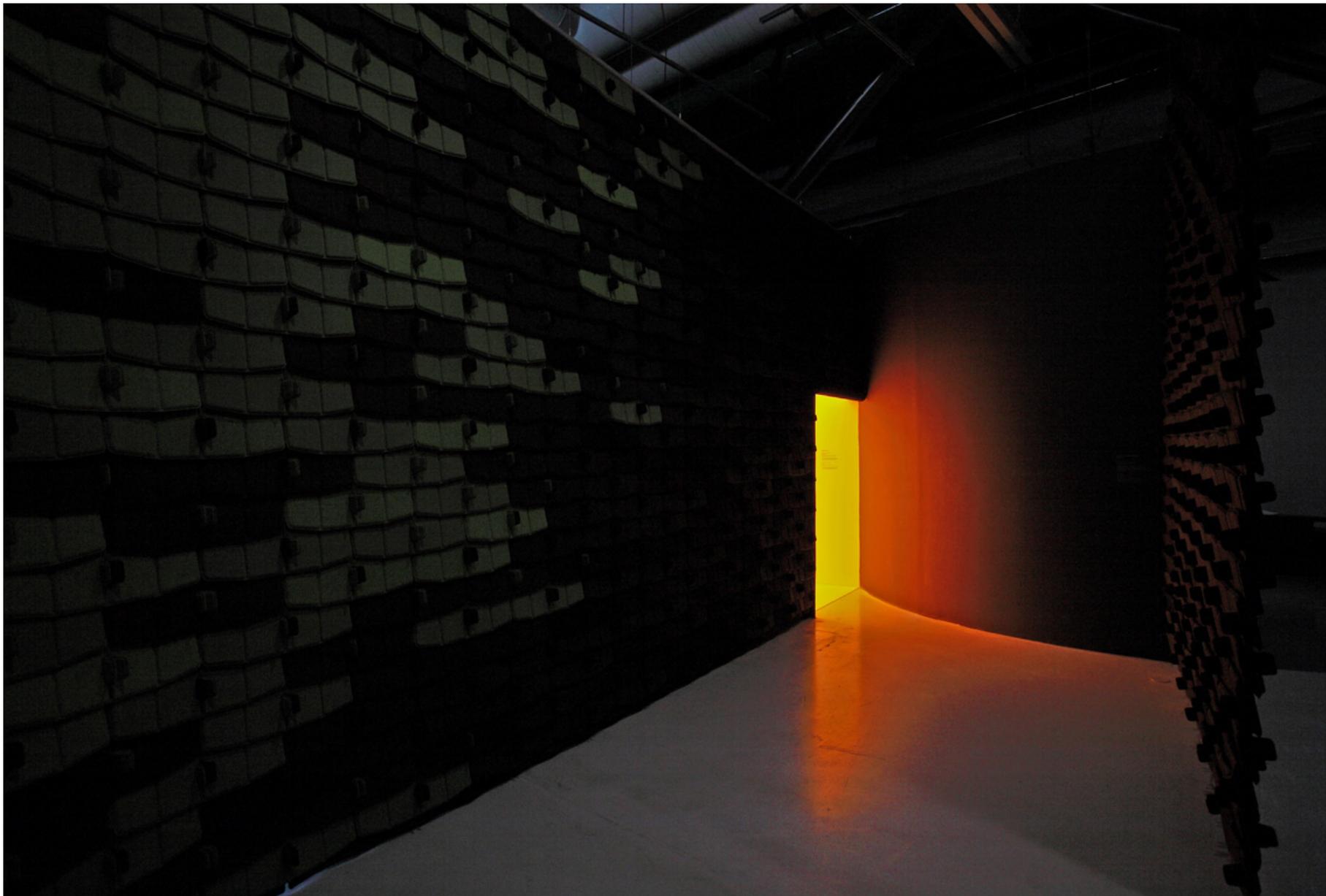


fig. 4 - Philippe Rahm, *Diurnisme*, 2007. Pièce présentée dans le cadre de l'exposition "Airs de Paris" (exposition collective), du 25 avril au 16 août 2007, Centre Pompidou, Paris.
© Adam Rzepka, Centre Pompidou . <http://www.philipperahm.com/data/projects/diurnisme/index.html>.



fig. 5 - Philippe Rahm, *Diurnisme*, 2007. Pièce présentée dans le cadre de l'exposition "Airs de Paris" (exposition collective), du 25 avril au 16 août 2007, Centre Pompidou, Paris.
© Adam Rzepka, Centre Pompidou . <http://www.philipperahm.com/data/projects/diurnisme/index.html>. (Echelle approximative : les banquettes ont la taille de lits à deux places.)

A wavelength upper to 570 nanometers has no influence on the melatonin secretion and is perceived by the brain like a night.

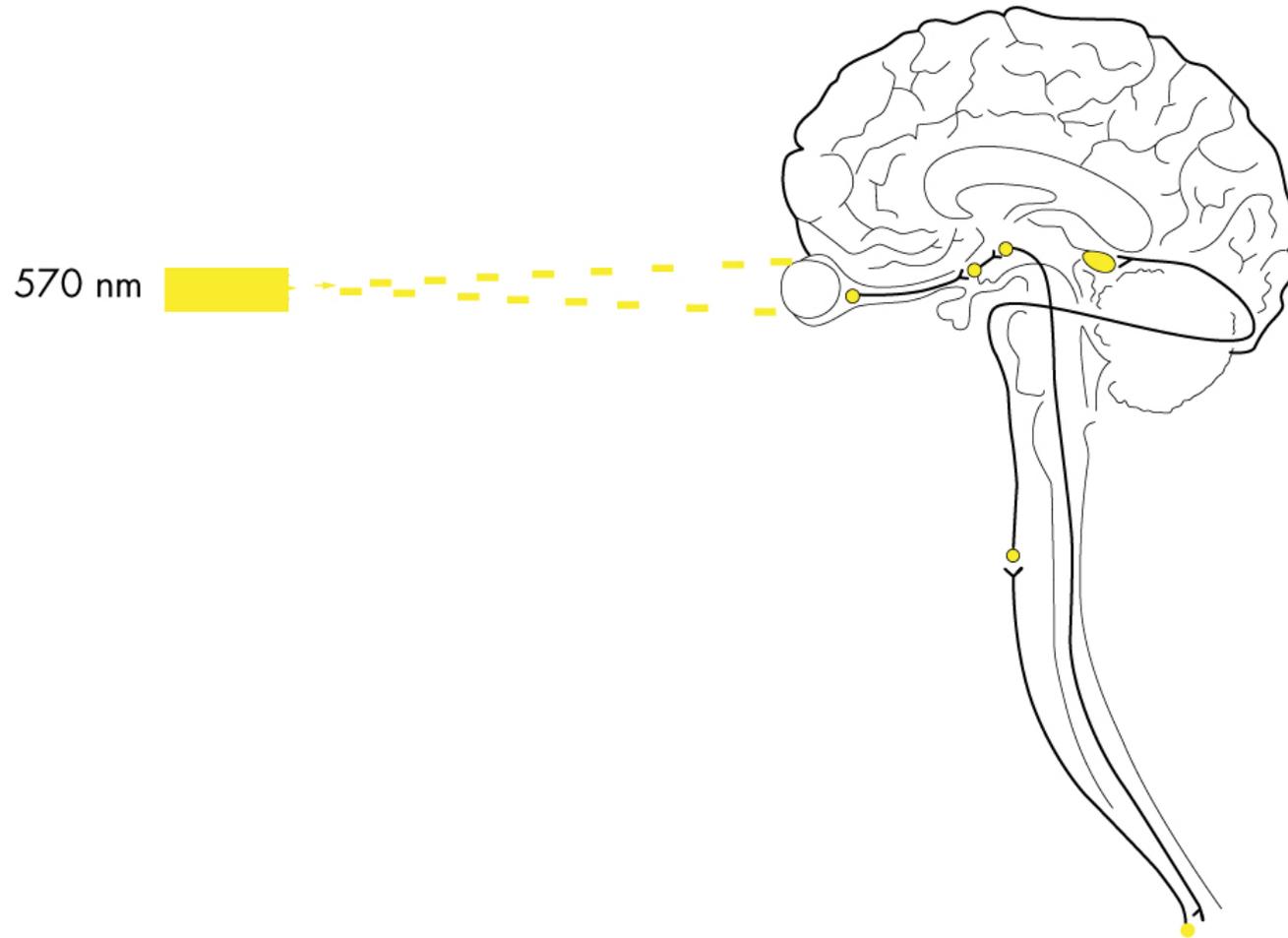


fig. 6 - Philippe Rahm, *Diurnisme*, 2007. Pièce présentée dans le cadre de l'exposition "Airs de Paris" (exposition collective), du 25 avril au 16 août 2007, Centre Pompidou, Paris.
© Rahm, 2007. <http://www.philipperahm.com/data/projects/diurnisme/index.html>.



fig. 7 (page précédente) - Tour de Flak de Wilhelmsburg. Date inconnue, photographie inconnu.
© Geschichtswerkstatt Wilhelmsburg & Hafen e.V.

Légende accompagnant la photo : “Er ist in Wilhelmsburg unübersehbar, der düstere mächtige Koloss von Bunker. Vor allem für Jugendliche aus der Umgebung ist der alte Bunker ein faszinierender Ort: er birgt Geschichte, Geheimnis und Gefahr. Gibt es vielleicht unbekannte Gänge, im Innern gar noch Skelette? Die Zeitzeugengeneration - vor allem die Kriegskinder - hingegen scheut eher diesen Ort, mit dem sie so zwiespältige Erinnerungen verbindet.

Seit April 2009 waren wir, die Klasse 7b, mit Unterstützung der Geschichtswerkstatt Wilhelmsburg auf der Spur der Geschichte des Bunkers. Zunächst erforschten wir das Gelände und versuchten uns in die Menschen hineinzusetzen, die damals 1944 und 1945 im Bunker Schutz suchten.

Ausgangspunkt für das Projekt waren Zeitzeugen, die damals als jugendliche Flakhelfer auf dem Bunker im Einsatz waren oder Kinder, die mit ihren Müttern im Bunker saßen. Sie sind heute zwischen 75 und 85 Jahre alt, mit ihnen führten wir Interviews, die die damalige Zeit in Wilhelmsburg für uns lebendig werden ließen. Wer wusste schon, dass selbst auf unserem Schulgelände ein Bombenkrater war.

Aus unseren Eindrücken entwickelten wir mit der Regisseurin Anke Krahe ein Theaterstück und gemeinsam mit der Geschichtswerkstatt Wilhelmsburg eine Ausstellung, der deren Umsetzung uns auch Eltern unterstützten.

Das war sehr abwechslungsreich und mal ganz anderer Unterricht, der sich gar nicht wie Unterricht anfühlte und viel Spass machte.”

Stadtteilprojekt zum Flakbunker Wilhelmsburg, Katholische Bonifatiuschule und Geschichtswerkstatt Wilhelmsburg & Hafen e.V.
<http://www.bonifatiuschule.org/index.php?page=bunkerprojekt>.



fig. 8 - Patrick Willner, Wilhelmshurg, 4 juin 2010.
© Patrick Willner, <http://www.fotocommunity.de/pc/pc/display/21365432>.

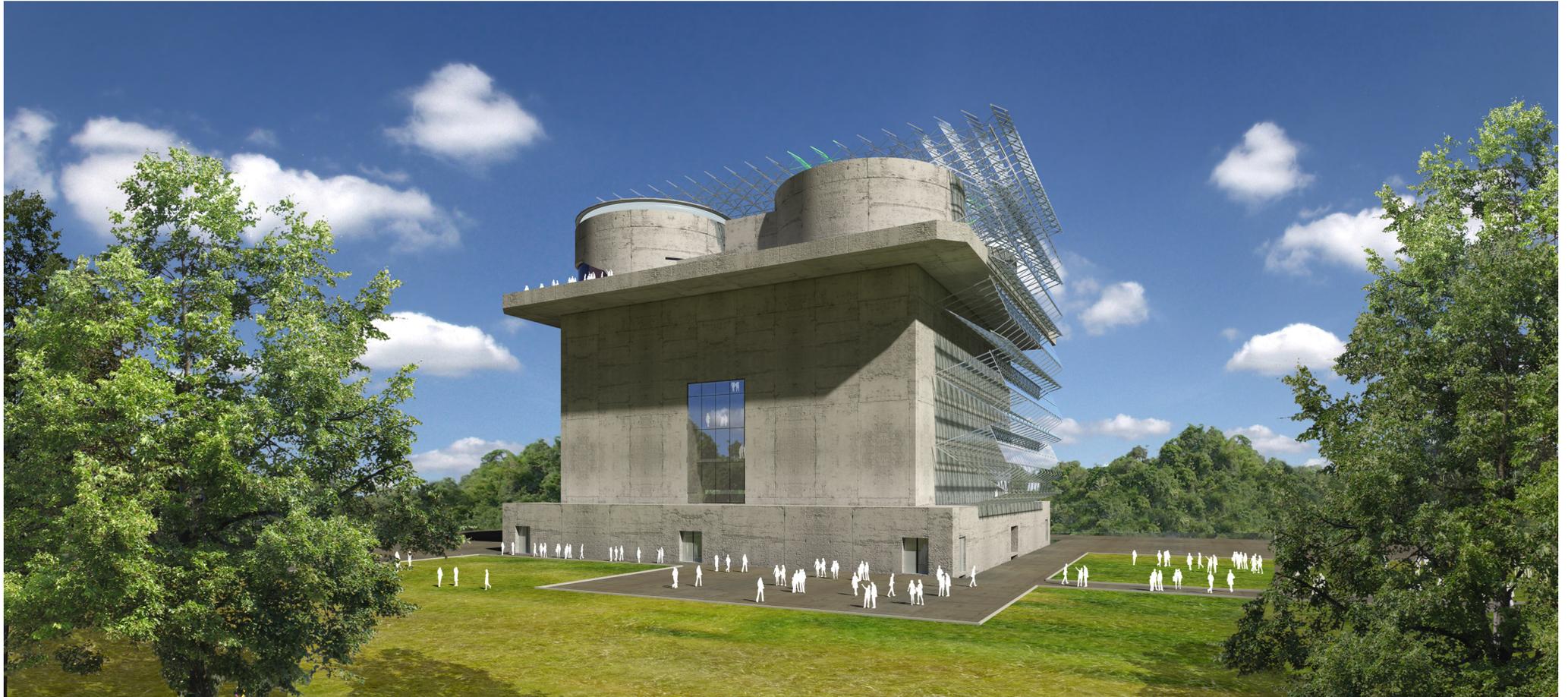


fig. 9 - Energy Bunker 03, sans date (probablement entre 2007 et 2012).

© IBA Hamburg GmbH / HHS Planer + Architekten AG, [http://www.iba-hamburg.de/en/service/downloads.html?type=515&width=928&height=700&tx_sribamedia_main\[medium\]=797&cHash=c09fbb1403f3c9fb3940d7a0560c0999](http://www.iba-hamburg.de/en/service/downloads.html?type=515&width=928&height=700&tx_sribamedia_main[medium]=797&cHash=c09fbb1403f3c9fb3940d7a0560c0999).

UTAH DATA CENTER

When construction is completed in 2013, the heavily fortified \$2 billion facility in Bluffdale will encompass 1 million square feet.



1 VISITOR CONTROL CENTER

A \$9.7 million facility for ensuring that only cleared personnel gain access.

2 ADMINISTRATION

Designated space for technical support and administrative personnel.

3 DATA HALLS

Four 25,000-square-foot facilities house rows and rows of servers.

4 BACKUP GENERATORS AND FUEL TANKS

Can power the center for at least three days.

5 WATER STORAGE AND PUMPING

Able to pump 1.7 million gallons of liquid per day.

6 CHILLER PLANT

About 60,000 tons of cooling equipment to keep servers from overheating.

7 POWER SUBSTATION

An electrical substation to meet the center's estimated 65-megawatt demand.

8 SECURITY

Video surveillance, intrusion detection, and other protection will cost more than \$10 million.

Source: U.S. Army Corps of Engineers Conceptual Site plan

fig. 10 - Plan de construction d'un *datacenter* de la NSA (National Security Agency)

Extrait de l'article de James Bamford, « The NSA Is Building the Country's Biggest Spy Center (Watch What You Say) », dans *Wired* (*Wired.com*), 15 mars 2012. http://www.wired.com/threatlevel/2012/03/ff_nsadatacenter/all/1.

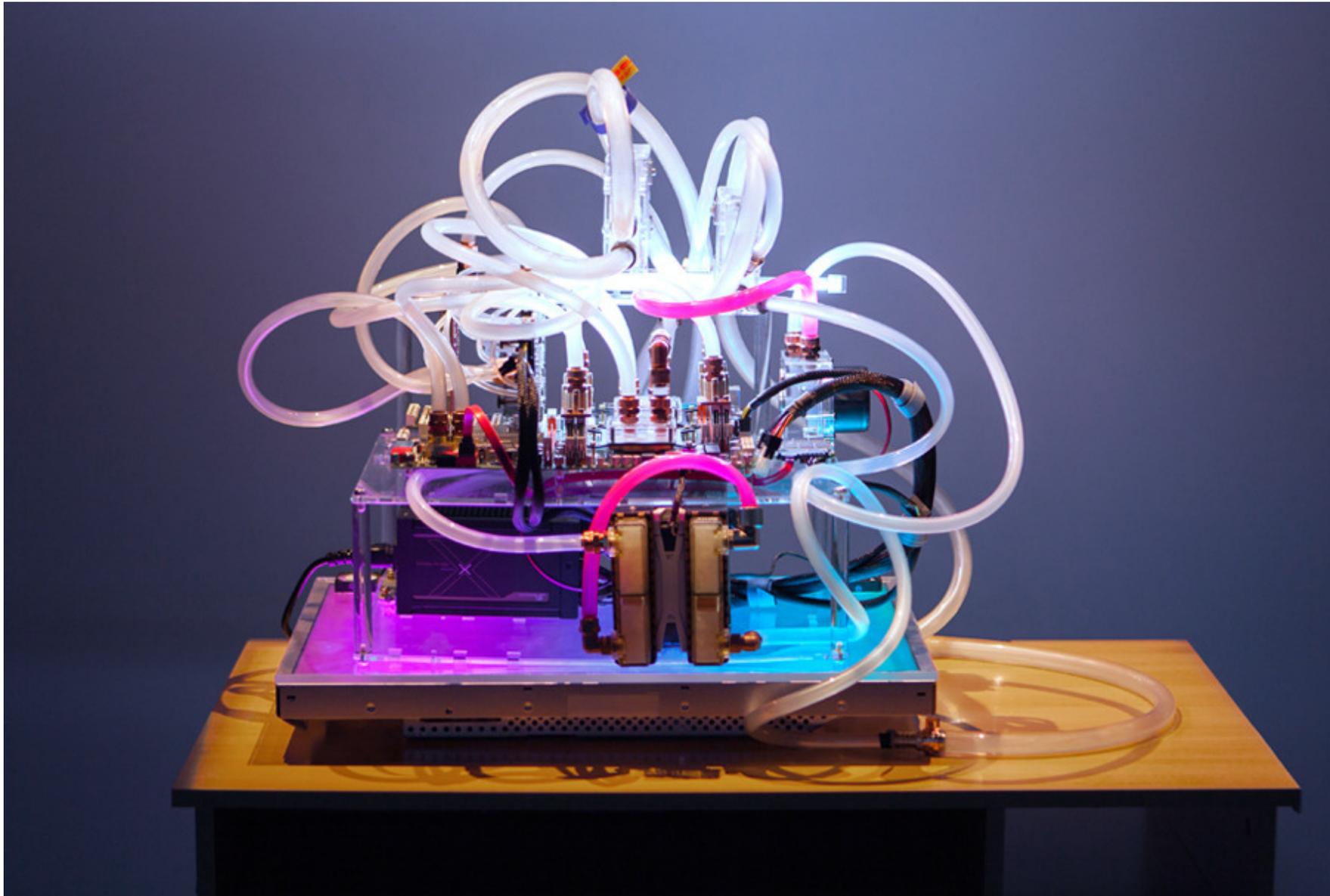


fig. 11 - PLEIX, *WATERCOOL*, dans le cadre de l'exposition « 2062, aller-retour vers le futur », du 1 fév. au 25 mars 2012 à la Gaîté lyrique, Paris.
© Photo : Mohamed Khalil, <http://pleix.net/2062>. (Echelle approximative : l'écran sur lequel est posée l'architecture fait environ 26 pouces.)

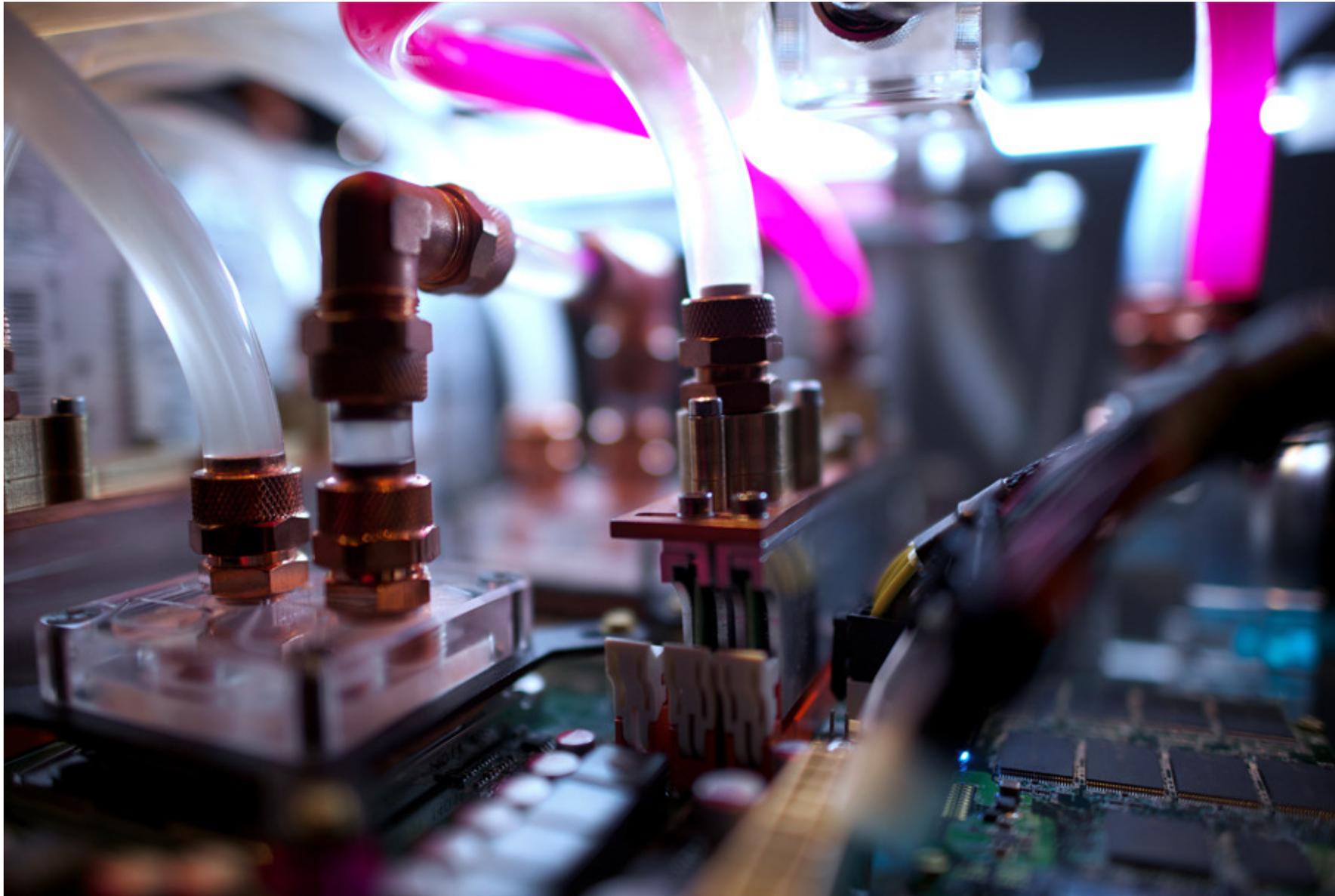


fig. 12 - PLEIX, *WATERCOOL*, dans le cadre de l'exposition « 2062, aller-retour vers le futur », du 1 fév. au 25 mars 2012 à la Gaité lyrique, Paris.
© Photo : Mohamed Khalil, <http://pleix.net/2062>.

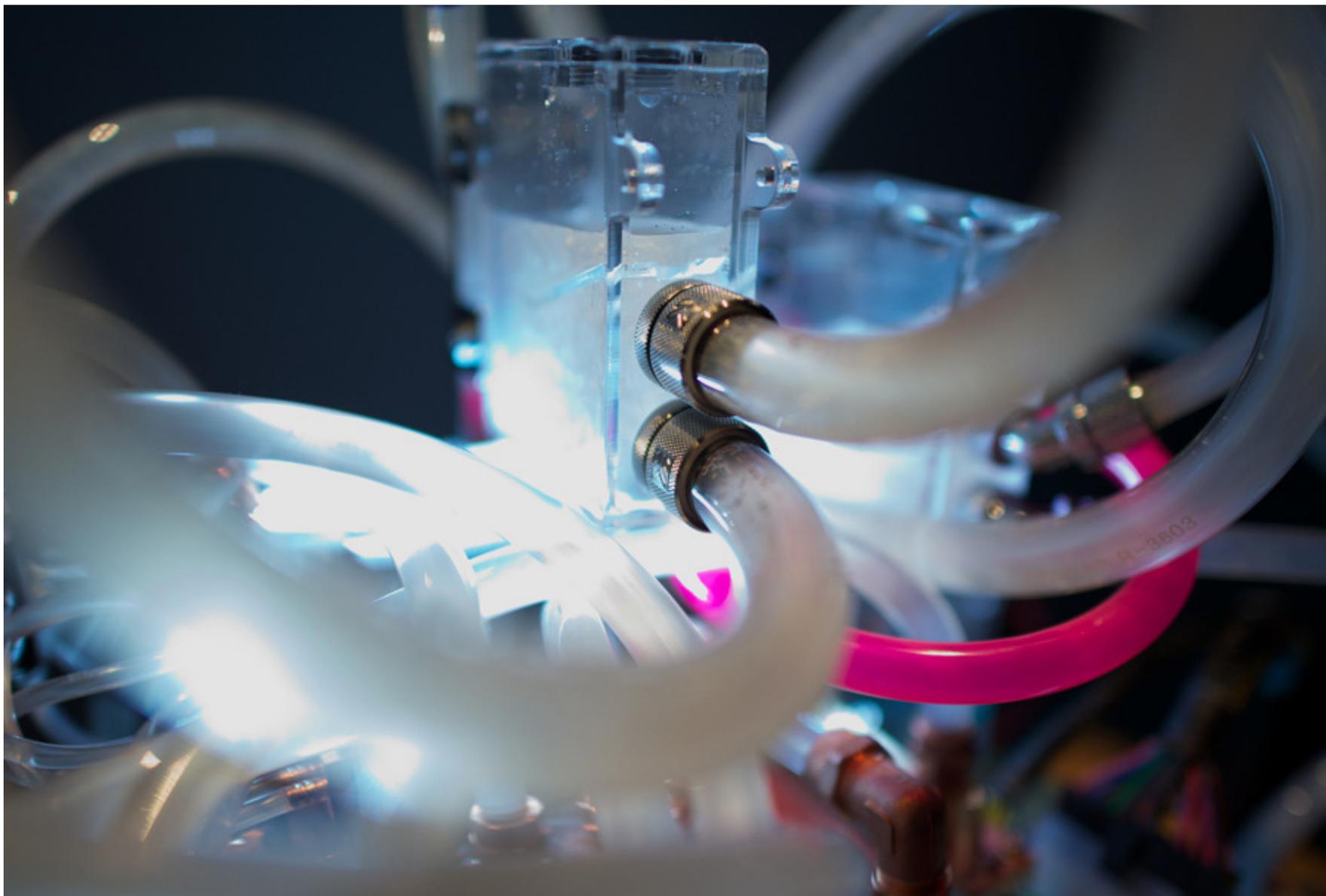


fig. 13 - PLEIX, *WATERCOOL*, dans le cadre de l'exposition « 2062, aller-retour vers le futur », du 1 fév. au 25 mars 2012 à la Gaité lyrique, Paris.
© Photo : Mohamed Khalil, <http://pleix.net/2062>.



fig. 14 - Microsoft Dublin Data Centre Entrance, “Microsoft Expands Cloud Computing Capabilities & Services in Europe”, 24 sept. 2009.
© Microsoft, http://www.microsoft.com/presspass/emea/presscentre/pressreleases/DublinDataCentrePR_240909.msp.



fig. 15 - Microsoft Dublin Data Centre Aerial, “Microsoft Expands Cloud Computing Capabilities & Services in Europe”, 24 sept. 2009.
© Microsoft, http://www.microsoft.com/presspass/emea/presscentre/pressreleases/DublinDataCentrePR_240909.msp.