

Université Paris I Panthéon-Sorbonne
UFR d'Arts Plastiques et Sciences de l'Art
Master 1 Esthétique
Adeline GASNIER

NO GHOST, JUST SOMEONE?
LES AGENTS CONVERSATIONNELS

Article de recherche préparé sous la direction de Monsieur Marc JIMENEZ

Juin 2011

Je remercie Monsieur Marc Jimenez, Professeur de Philosophie Esthétique à l'Université de Paris 1 Panthéon-Sorbonne, pour ces échanges qui ont permis de soulever des questions mais aussi des réponses.

Table des matières :

Préambule	p. 02
Introduction	p. 04
I. Un large miroir dans la poitrine ? Vitruve, Lucien et Turing	p. 04
II. Machines de Turing et imaginaires ? <i>GitS, Annlee, Als</i>	p. 12
III. Quelques <i>a priori</i> : Intelligence Artificielle, corps, Prudence	p. 18
IV. <i>Chatbots</i> ? Contrôle, architecture mémorielle, émulation	p. 22
Conclusion	p. 30
Bibliographie / Webographie	p. 31

Annexes (DVD, "03.pdf") :

Illustrations	p. 03
Reproduction de l'article "SmarterChild", Wikipédia, 25 mai 2011	p. 15

DVD joint :

"Bruno Taut et la <i>Gläserne Kette</i> : architecture d'émotion"	<u>01_BTGK.pdf</u>
+ Annexes	<u>01.pdf</u>
" <i>Cosmic Incubadora</i> : cristallisation émouvante"	<u>02_CICE.pdf</u>
+ Annexes	<u>02.pdf</u>
" <i>No Ghost, Just Someone?</i> Les agents conversationnels"	<u>03_NGJS.pdf</u>
+ Annexes	<u>03.pdf</u>
"Modèles émotionnels et agents conversationnels"	<u>04_MEAC.pdf</u>
+ Annexes	<u>04.pdf</u>

* Films [*Docudrama*] de David Williams, *Lillian* (1993) et *Thirteen* (1997).

PRÉAMBULE

S'il n'y a pas de large ouverture dans la poitrine des êtres humains pour pouvoir attester de ce qui s'y passe, quelles pensées, émotions et desseins s'y fomentent *vraiment*, alors il reste toujours l'éducation. Certes, la demande envers une structuration des émotions est une récurrence atavique, mais avec les deux précédents articles j'ai essayé de mettre l'accent sur une dimension peut-être un peu plus "récente" : l'étude et la prise en compte de ceux à qui il s'agit de s'adresser.

Avec le premier article, via la *Gläserne Kette* (1919-1920)¹, l'image d'un cristal à multiples facettes en tant que stratification émotionnelle des cultures humaines s'est dégagée. Soit une construction qui attirerait par reflet (du Soi ou du Moi) ceux qu'il s'agirait d'impliquer dans un projet d'élaboration sociale. Mais il y avait là au moins deux chemins possibles : d'un côté la mise en place d'un dispositif utilisant des surfaces miroitantes pour refléter le contexte ; de l'autre la modélisation préalable des éléments disparates à assimiler. La première démarche consisterait à séduire indistinctement, pour ensuite s'attacher à la dimension mémorielle (construction de la mémoire singulière et collective). Soit une large captation de l'attention, en tant que préalable de toute proposition de récit et de mémoire — à l'instar ici d'un joyau cristallin qui, quoiqu'il reflète, intègre son environnement mais impose sa structure². L'aspect réfléchissant de ce type de construction ne reposerait pas sur une observation de l'environnement (ou du public) donné, mais sur l'élaboration d'une faculté réfléchissante générale, indépendamment de tout contexte. L'autre démarche consisterait en une étude préliminaire du public pour pouvoir s'y adresser de façon ciblée, en s'adaptant à chaque personnalité. L'aspect réfléchissant de ce type de construction reposerait sur une faculté d'isoler et d'interpréter des éléments exemplaires du contexte, afin de procéder à un mimétisme ponctuel. Et l'on pourrait dire que là où la première démarche a besoin de dégager des propriétés universelles, la deuxième a besoin

¹ Cf. mon article de Master 1, "Bruno Taut et la *Gläserne Kette* : architecture d'émotion".

² En guise de réponse critique vis-à-vis de cette conception expressionniste, portée par B. Taut, on peut penser notamment aux peintures de l'artiste belge, Laurent Impeduglia [fig. 1, 2], ayant récemment exposé au Centre Wallonie Bruxelles de Paris, dans le cadre du "Printemps de l'Irrévérence", du 18 mars au 29 mai 2011.

d'élaborer des modèles individués pour pouvoir procéder à un adressage ciblé, en fonction des propriétés singulières. L'accent se déplace sur les façons de modéliser, sur les façons de comprendre et connaître l'environnement³.

Avec le deuxième article et à travers l'installation *Cosmic Incubadora* (2010) de Gerda Steiner & Jörg Lenzlinger⁴, au-delà de la perspective d'une mise au musée de la pollution, d'autres questions se sont posées : la modélisation préalable des émotions équivaut-elle à une mise en ordre de celles-ci, et si oui, cela ne revient-il pas à les annihiler ? Mais s'il faut réserver une quelconque part indéfinie à l'émotion, cette circonscription en négatif n'est-elle pas à son tour une caricature ? Du côté de la réception, un autre problème s'est dégagé : face à une profusion d'informations, peut-on faire autrement que de s'y reconnaître, et cette reconnaissance est-elle gage d'inclusion — le miroir perpétuel ne déréalise-t-il pas plus qu'il n'atteste ? Alors qu'un besoin de distance peut pousser à chercher du différent, en tendant un miroir le divertissement oblige une énième fois à se considérer soi-même. Et plus on s'éloigne du Soi, plus on le retrouve, plus on est invité à méditer sa relation de *Self-control*. Ce qui peut tout autant donner du poids à cette réalité que la rendre factice par l'impossibilité de la dépasser⁵.

Dans son manifeste *Architecture de verre* (1914), Paul Scheerbart mettait en garde contre les effets vif-argent en architecture et préconisait de les fuir "comme un poison"⁶. Cependant, il n'en préconisait pas moins de rendre les nuits semblables aux jours (par la lumière électrique) et, considérant les difficultés que pourraient rencontrer les astronomes, envisageait simplement de les envoyer dans quelque coin reculé pour pouvoir observer les étoiles⁷. Mais qu'est-ce que cette

³ Ce problème n'est pas nouveau : en un sens, la *Critique de la raison pratique* de Kant (1788) mettait déjà en lumière les difficultés d'une conciliation entre relations universelles et valeurs universelles. Mais avec la modélisation informatique, ce souci trouve peut-être un autre dynamisme.

⁴ Cf. mon article de Master 1, "*Cosmic incubadora* : cristallisation émouvante".

⁵ Je pense ici aux propos (décontextualisés) de F. Lemarchand : "Que faire alors lorsqu'on ne peut pas fuir, lorsqu'il n'y a pas d'ailleurs où aller chercher refuge ? [...] lorsqu'aucun moyen, technique ou psychique, ne permet d'échapper à la condition humaine qui consiste à devoir durablement habiter des territoires durablement contaminés, seul le *déni de réalité* permet d'affronter plus ou moins sereinement l'avenir." [Je souligne]. Frédérick Lemarchand, "Fukushima, l'autre Tchernobyl", dans *Le Monde*, n° 20602, 19 avril 2011, p. 18.

⁶ Paul Scheerbart, Guillaume Desanges, François Piron, Marcus Steinweg, *Intouchable : L'idéal transparence - L'architecture de verre*, "XXIII - À éviter : les effets vif-argent des miroirs", trad. Christophe Marchand-Kiss, éd. Xavier Barral / Villa Arson, Paris / Nice, 2006, p. 27.

⁷ "De ces nuits de lumière, que doit nous offrir l'architecture de verre, que dire d'autre sinon qu'elles

"absence" d'étoiles en zone urbaine, sinon la conséquence d'un immense effet vif-argent contre l'atmosphère ? Là encore, ce retour sur soi forcé est aujourd'hui synonyme de pollution. Le tout inclusif, s'adressant à tout le monde à force ne s'adresser à personne en particulier, peut devenir un récit tentaculaire potentiellement asphyxiant, ou du moins perçu comme polluant.

Le troisième article va prolonger ce questionnement à propos de la mise en architecture des émotions, en s'attachant au contexte de la simulation de pensée, et en se demandant s'il est possible de dépasser l'effet vif-argent⁸.

INTRODUCTION

Dans cet article, nous allons aborder les *agents conversationnels*. Comme l'un des enjeux réside dans le fait *d'essayer de comprendre ce que peut être et comment peut fonctionner un agent conversationnel*, en donner une définition maintenant serait prématuré. Mais pour ne pas partir de rien, nous pouvons les affilier aux machines de Turing et aux programmes d'intelligence artificielle (ces points seront développés plus loin), sachant que "chaque" agent conversationnel est peut-être "unique à sa façon". Avec la première partie, il sera question de ce qui peut motiver le *jeu de l'imitation*. Avec la deuxième nous évoquerons quelques-uns des imaginaires incluant ces agents. Avec la troisième nous essaierons de dégager des notions clefs — lorsqu'il s'agit d'établir un contact. Et avec la quatrième nous tenterons de questionner la singularité.

Le but étant aussi d'interroger une partie des théories accompagnant cette expérience relationnelle, conçue pour être vécue de façon singulière.

I. UN LARGE MIROIR DANS LA POITRINE ? VITRUVÉ, LUCIEN ET TURING.

Si j'ai entamé cet article en évoquant l'idée d'une "large ouverture dans la poitrine", c'est en pensant au *De Architectura* (vers 15 av. J.-C.) de Vitruve :

sont vraiment « extraordinaires ». [...] 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.", *Ibid.*, "LXXXIV - Les nuits de lumière de l'architecture de verre", p. 58.

⁸ Le vif-argent (mercure) a été utilisé pour enduire les plaques de verre polies dans la fabrication des miroirs. Du fait de sa toxicité, il est aujourd'hui remplacé par l'argent ou l'aluminium.

"SOCRATE, qu'Apollon lui-même, par son oracle de Delphes, déclara le plus sage de tous les hommes, disoit avec raison, qu'il seroit à souhaiter que nous eussions une ouverture à la poitrine, afin qu'on pût y lire nos pensées et nos desseins. [...] Mais puisque la nature ne l'a pas voulu, et qu'il ne nous est pas permis de pénétrer dans l'esprit de l'homme, pour connoître avec certitude les sciences qu'il renferme et qui y sont souvent cachées, les meilleurs artistes auront beau promettre d'employer tous les talents possibles pour faire réussir les œuvres qu'ils entreprennent, [...] s'ils n'ont pas les talents de se faire valoir et une facilité de s'exprimer, proportionnée à leur science ; on ne croira jamais qu'ils connoissent à fond leur art⁹."

Par cette anecdote adressée à Jules César, Vitruve introduit le troisième livre dans lequel, là encore, il va s'employer à expliciter et détailler la logique des principes qu'il préconise en architecture. En ce sens, l'analyse, l'explication et l'éducation vont faire office de larges ouvertures ; le livre ouvert et son déroulement narratif deviennent cette fenêtre, cette preuve, l'attestation d'une vérité. Mais on pourrait tout autant lire cela depuis un autre contexte — *test de Turing* — et en tirer d'autres questions : peut-on lire dans les pensées ? Qu'implique ce désir de transparence ? Où peut commencer et où doit s'arrêter l'exploitation du "vivant" et du "mort" ? Qu'est-ce qu'une preuve, ou plutôt, qu'est-ce que la fidélité ?

Dans *Hermotimus, ou les sectes* (vers 153-165 après J.-C.), Lucien de Samosate reprend l'anecdote, sans l'attribuer à Socrate cependant. Dans ce dialogue, Lycinus y convainc Hermotimus, disciple stoïcien, de l'inefficacité de sa méthode de recherche. Il lui reproche d'avoir fondé son choix philosophique (stoïcisme) sur quelque chose d'indémontrable tout en le tenant pour démontrable (soit de confondre vérité et fidélité) :

"LYCINUS. Comment donc t'a-t-il été possible de distinguer, aux signes dont tu parlais, celui qui philosophait bien ou mal ? [...] On dit que Minerve, Neptune et Vulcain, disputèrent un jour d'adresse et d'industrie. Neptune forme un taureau, Minerve invente l'art de construire les maisons, et Vulcain donne naissance à l'homme. Ils vont ensuite trouver Momus, qu'ils avaient choisi pour juge. Momus considère l'œuvre de chacun. Ce qu'il trouve à redire dans les autres œuvres, nous n'avons pas besoin de le rapporter ici. Quant à l'homme, il blâme Vulcain, qui l'avait bâti, de n'avoir pas placé une petite fenêtre sur sa poitrine, afin qu'en l'ouvrant, tout le monde pût connaître ses désirs et ses pensées, s'il mentait ou s'il disait la vérité¹⁰."

Et avec ce passage on retrouve l'idée selon laquelle des paroles vertueuses ne précèdent ou n'accompagnent pas nécessairement, ni une pensée vertueuse, ni

⁹ Vitruve, *L'Architecture de Vitruve*, Livre III, "Introduction", trad. par De Bioul, éd. Adolphe Stapleaux, Bruxelles, 1816, pp. 92. La traduction de Ch. L. Maufrais (1847) renvoie vers Lucien.

¹⁰ Lucien, *Œuvres complètes de Lucien de Samosate*, Tome I, *Hermotimus, ou les sectes*, trad. par Eugène Talbot, éd. Hachette et Cie, Paris, 1912, pp. 299-300.

un acte vertueux — ce qui rejoint le propos de Vitruve et la question de la preuve. Mais par ce dialogue, en s'attaquant à une figure de paille et à un Hermotimus peu enclin à défendre sa façon de procéder et ses visées stoïciennes, Lucien donne aussi du poids à son argument d'une impossible impartialité : quelque sage qu'on soit, on est souvent juge et partie lorsqu'il s'agit de rendre les arguments des uns et des autres. Plus encore : il est des raisonnements que l'on ne peut suivre dans leur finesse qu'en y consacrant sa vie, ce qui exclut de pouvoir comparer toutes les philosophies, du fait de leur nombre et de la brièveté d'une vie humaine. Chez Lucien, la vérité est au prix de l'exhaustivité et de l'optimum. Autrement dit, même si Vitruve explicite ses raisonnements et connaissances pour pallier l'absence de transmission instantanée, la profusion des traités d'architecture et leur complexité respective ne permettent pas pour autant d'avérer l'optimalité des méthodes employées. Lucien défend l'idée d'une diversité en s'attachant aux nuances qui rendent toute expérience de pensée singulière. Ici, seul celui qui pense sait qu'il pense et ce à quoi il pense, dans la subtile et pleine lenteur de sa pensée. L'entreprise d'évaluation et de comparaison de tous les raisonnements possibles devient alors incompatible avec les performances d'un intellect humain, à moins de croire sur parole, c'est-à-dire de transformer la croyance en preuve.

En 1950, Alan Turing va reprendre à la philosophie ces questions de la transparence intellectuelle, de l'amélioration, de la croyance et de la preuve, pour les importer dans le domaine informatique, via le *jeu de l'imitation*. Si l'on demande "Est-ce que les machines peuvent penser ?" (en termes de possibilité matérielle et non en termes d'autorisation), on peut se heurter au problème du solipsisme, à savoir que seul le sujet pensant est en position de certifier son propre *cogito*. Et avec les machines (ou les programmes), en tant qu'elles ne font pas partie du genre humain, la question dépasse difficilement cet écueil puisque le problème de leur crédibilité régule en amont la façon dont cette affirmation sera reçue. Une entité peut bien affirmer qu'elle pense, si en amont il est décidé qu'elle *n'est pas* crédible, cette information reste, pour qui la reçoit, le fruit d'un habile tour de magie (ou de mécanique), tout au mieux une affabulation ou le reflet de son/sa créateur-trice. Et de même, pour qui décide de croire *a priori* qu'une machine peut penser, toute suite d'informations, aussi *non sequitur* soit-elle, peut alors donner lieu à une

interprétation sensée. Lorsque la question concerne une machine à tisser, il est facile de prendre position, mais lorsqu'elle porte sur des programmes d'intelligence artificielle (machines à tisser du sens), les choses peuvent se compliquer et rendre le contournement d'Alan Turing, en 1950, remarquable.

Au lieu de s'attaquer à redéfinir les notions de "penser" et de "machine", dans son article "Computing machinery and intelligence", Turing propose de laisser la machine convaincre par elle-même (il propose de lui donner la parole). On pourrait dire que la question "*Can machines think?*", par le truchement d'un langage appréhendé comme ambigu, fait redouter à Turing une réponse en forme de cristallisation d'opinion publique¹¹ (soit une réponse circonstancielle, orientée par les multiples façons de présenter le problème, à l'instar des résultats de sondage). Plutôt que de porter sur l'évaluation du produit réflexif d'une machine, en termes de capacité, le débat risque de s'agréger à la dimension d'autorisation à penser, qu'il s'agit (ou non) d'accorder aux machines. À la place de la question "*Can machines think?*", Turing va proposer un protocole de test, aujourd'hui connu sous le nom de *test de Turing*, et alors nommé jeu de l'imitation [*imitation game*]. Pour l'expliquer, il prend l'exemple d'un test qui impliquerait trois personnes : un homme (A), une femme (B) et un interrogateur (C) (cf. Turing, 1950, "1. The Imitation Game"). L'interrogateur, conscient de s'adresser à un homme et à une femme, mais sans les voir et sans savoir qui est quoi, devrait à l'aide de questions (et des réponses fournies) finir par attribuer un sexe à chacun. Et c'est sur ce modèle qu'est pensé le protocole de décision, quant à savoir si une machine pense ou non : l'homme (A) est remplacé par un ordinateur, et l'interrogateur doit désormais décider de qui est humain, qui est ordinateur. Ce faisant, Turing propose une nouvelle question — "What will happen when a machine takes the part of A in this game¹²?" (Turing, 1950, "1. The Imitation Game") — ; question qui est peut être une façon d'alléger ou de renouveler les enjeux éthiques, en réorientant le point focal, de la machine vers l'interrogateur et le dispositif. Mais il propose également une nouvelle façon de cristalliser l'opinion : par l'interactivité du questionnement et

¹¹ Cf. Edward L. Bernays, *Crystallizing Public Opinion* (1923). Bernays y développe le rôle du conseiller en relations publiques, dans la captation et l'orientation des opinions.

¹² "Que va-t-il se passer quand une machine prend le rôle de A dans ce jeu ?" [Ma traduction].

le test réciproque, suivant des règles données¹³. En un sens, ce protocole incite tout autant à s'observer soi-même : on doit d'abord faire le bilan de sa propre humanité pour pouvoir décider de celle des autres. On est donc amené à se conscientiser en tant qu'être humain, mais aussi en tant qu'être ayant le pouvoir d'attribuer, ou non, le statut d'humain.

La version commune de ce test, aujourd'hui, consiste à s'entretenir avec un agent conversationnel, sur internet. Et en sachant *a priori* que l'interlocuteur est un programme, il s'agit d'évaluer le degré de crédibilité (ou de fidélité à l'expression humaine) de celui-ci. Toute manifestation d'"ordinateurité" (empiriquement détectée et guettée) est potentiellement le trait d'une défaillance, voire un motif à raillerie. Ce peut donc aussi devenir, en prenant les choses au premier degré, l'apprentissage de ce que *ne peuvent* en aucun cas être les réponses d'un *vrai* être humain. Mais laissons l'expérience contemporaine pour le moment et revenons à cet article de 1950.

Le jeu de l'imitation y est développé plus avant, jusqu'à ce que Turing propose de ne plus chercher à créer des machines capables de simuler la pensée d'un adulte éduqué, mais plutôt de privilégier le développement de machines capables de simuler l'apprentissage, comme les enfants :

"Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? [...] 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. [...] 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 up¹⁴." (Turing, 1950, "7. Learning

¹³ On peut penser aux propos de Benjamin au sujet du cinéma : "Le public se trouve, ainsi, dans la situation d'un expert dont le jugement n'est troublé par aucun contact personnel avec l'interprète. *Il n'a de relation empathique avec lui qu'en ayant une relation de ce type avec l'appareil. Il en adopte donc l'attitude : il fait passer un test.*" (Walter Benjamin, *L'œuvre d'art à l'époque de sa reproductibilité technique* (1939), trad. par Maurice de Gandillac et Rainer Rochlitz, éd. Folioplus, Paris, 2008, p. 28). Le présent article va tenter de proposer une autre argumentation quant aux appareils informatiques, en ayant néanmoins en tête l'*examineur distrait* de Benjamin.

¹⁴ "Au lieu d'essayer de produire un programme pour simuler l'esprit adulte, pourquoi ne pas plutôt essayer d'en produire un qui simule celui d'un enfant ? [...] Nous avons ainsi divisé notre problème en deux parties. Le programme enfant et le processus d'éducation. Les deux restent très étroitement connectés. Nous ne pouvons pas nous attendre à trouver une bonne machine enfant à la première tentative. [...] Il y a une connexion évidente entre ce processus et l'évolution, par les

Machines")

Je laisse ici de côté la conception réductionniste que Turing peut évoquer par ailleurs quant à l'esprit des enfants (simple mécanisme de bloc-notes), en ce qu'on pourrait tout autant aujourd'hui lui opposer une conception "réductionniste" quant aux fonctionnements des programmes informatiques (simple mécanisme de bloc-notes). Mais je retiens au moins deux aspects : dans un premier temps, l'accent mis sur la nécessité d'une *rapidité d'évolution*, divisible à son tour en deux points.

Pour le premier point, on peut se souvenir des propos de Lucien : la lenteur du raisonnement humain empêche une comparaison efficace entre toutes les philosophies, rendant en quelque sorte caduque l'aspiration à une recherche de la vérité *fondée sur* un choix rationnel, avéré et généralisable. Selon Lucien, il arrive un moment où ce n'est plus qu'affaire de circonstances. Mais avec la rapidité des programmes informatiques, on voit poindre une opportunité : sera-t-il enfin possible de comparer toutes les philosophies ? Les programmes pourraient exécuter la tâche (ou le travail critique) préconisée par Lucien et nous saurions avec certitude quelle philosophie est vraie et digne d'être suivie.

Si cette ambition est naïve, et c'est le second point, alors il est peut-être tout aussi arbitraire de présumer que le processus de sélection des programmes soit toujours rationnel. En dépit des horizons laissant miroiter des idées d'immortalité et d'éternité, la longévité des agents conversationnels peut parfois être singulièrement écourtée, à l'instar d'un programme télévisé soumis à l'audimat, ou soumis au "ressenti" de son créateur-proprétaire¹⁵. L'évolution est certes fulgurante mais on peut aussi se demander de quel type d'éducation et de sélection il s'agit : et si le réflexe "sélection du meilleur / élimination du reste" devenait la seule méthode, la seule philosophie soutenue, testée et assimilée ?

Cela dit, les propos de Turing rejoignent ceux de Vitruve : en matière de formation humaine, l'éducation et l'apprentissage culturel sont plus rapides et fiables que la transmission de pensée instantanée ou que la création biologique.

identifications / Structure de l'enfant machine = matériel héréditaire / Changements de l'enfant machine = mutation, / Sélection naturelle = jugement de l'expérimentateur / On pourrait espérer, cependant, que ce processus soit plus expéditif que l'évolution. La survie du plus adapté est une méthode lente pour mesurer les avantages. L'expérimentateur, par l'exercice de l'intelligence, devrait être capable de l'accélérer." [Ma traduction].

¹⁵ Cf. SmarterChild, <http://en.wikipedia.org/wiki/SmarterChild> (article reproduit en Annexes, p. 15).

Mais cela ne résout pas la question de savoir à quoi on re-connaît un être humain.

Cette question nous amène vers le second aspect notable du passage mentionné : l'*imitation*. On est assuré, tout d'abord, de se trouver face à un programme nous imitant. Mais comme suggéré précédemment, c'est seulement après avoir conscientisé ses propres critères de jugements que l'on peut se trouver apte (ou non) à décider. Alors une question s'amorce : ne peut-on évaluer qu'en miroir ?

Turing avait beaucoup d'ambition pour ses machines enfants et pensait déjà un moment où il serait difficile de faire la distinction entre l'expression d'un programme et l'expression d'un être humain. Dans la mesure où il est possible de s'affronter, soixante ans plus tard, avec des programmes aptes à tenir des conversations, force est peut-être de constater que cette ambition était un peu "optimiste". Mais le test de Turing n'en gagne pas moins en intérêt, en ce qu'il recèle une question inattendue : quel est notre degré de tolérance face à la différence ? Quand bien même un programme *penserait* de "façon computationnelle", avec des gimmicks informatiques, en quoi serait-ce gênant¹⁶ ?

On peut revenir sur la façon dont a été formulée la "nouvelle" question : "What will happen when a machine takes the part of A in this game?". Ce "*Take the part of*" évoque tout autant l'idée de remplacer, de jouer le rôle de, d'agir en tant que, mais encore l'idée de prendre parti pour quelqu'un dans une controverse [*take someone's part*]. Il y a quelque chose d'inquiétant en première lecture, à se sentir potentiellement réduit à n'être qu'une partie d'un dispositif ne fonctionnant qu'avec des *placeholders* (ou *tokens*) ; à se sentir si facilement remplaçable par une machine — questions de l'interchangeabilité et de la symétrie. Mais ce n'est pas l'aspect le plus troublant : ce qui peut amener à douter, c'est la mise en évidence de la difficulté culturelle à attribuer une faculté de penser à ce qui n'est pas formellement similaire ; à n'envisager ce qui est autre qu'en termes de place à prendre ou à se faire prendre.

Comme le note Turing, "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

¹⁶ Quand bien même ces entités ne sont pas créées pour la beauté du geste mais pour exécuter des fonctions déterminées.

this point it is usual to have the polite convention that everyone thinks¹⁷." (Turing, 1950, "5. Universality of Digital Computers", "(4) The Argument from Consciousness"). Cette politesse n'est pas un fait simple et les conditions qu'elle requiert pour sa mise en pratique semblent inclure, dans le test de Turing, le fait de dissimuler les corps pour ne pas biaiser le jugement quant à savoir si ce qui parle est humain *ou* non. Et dans une certaine mesure, le test de Turing rend palpable une limite : comment peut-il y avoir échange et changement si l'on n'est réceptif qu'à ce qui nous ressemble le plus strictement ? Si l'on n'accepte de communiquer qu'avec des miroirs, de ne prendre au sérieux que ce qui nous reflète, ce qui suit nos codes ? Turing lui-même parle de jeu [*game*], mais n'est-ce pas parce que seul un jeu, dont les règles nous autorisent à prendre au sérieux des machines, peut nous permettre de dépasser l'autocensure qui consiste à refuser d'attribuer une faculté de penser à ce qui ne nous ressemble pas physiquement ? Faut-il voir dans cette méfiance un déterminisme génétique¹⁸ ou une construction culturelle du Soi et du non-Soi¹⁹ ?

Par rapport au jeu, en incise éthique, on peut noter que jouer à faire le mort n'équivaut pas à mourir, et que jouer à attribuer une faculté de penser à un programme n'équivaut pas à lui conférer droits et devoirs imprescriptibles. En un sens, le jeu se distingue peut-être par la réversibilité des conséquences négatives. Mais dans cette perspective, seules les conséquences peuvent rétrospectivement permettre de qualifier le fait de savoir s'il s'agissait ou non d'un jeu, à moins de minimiser la singularité de façon à maximiser la réversibilité — si tant est que réversibilité et reproductibilité soient corrélées. Par exemple, face à un groupe d'éléments, au lieu de les considérer comme distincts (A, B, C, etc.), les considérer comme interchangeables (A, A, A, etc.) de façon à minimiser les défaillances singulières. Dans ce jeu les machines nous imitent, mais héritent-elles pour autant de nos devoirs et droits ?

En l'état, il est difficile de mettre de côté la coloration péjorative qui

¹⁷ "A est susceptible de croire « A pense mais B ne pense pas » tandis que B croit « B pense mais A ne pense pas ». Au lieu de disputer continuellement sur ce point, il est habituel de convenir poliment que tout le monde pense." [Ma traduction].

¹⁸ À la façon de Konrad Lorenz, dans *L'agression : une histoire naturelle du mal* (1969).

¹⁹ À la façon d'Angela Davis, dans *Women, Culture & Politics* (1990).

accompagne les termes "machine", "robot", "programme" : les concevoir comme serviteurs du genre humain rend toute identification avec eux dévalorisante. Le fait de sympathiser *en égal*, avec un "robot" (ou "bot"), peut être perçu comme preuve de crédulité — ou pour le dire de façon triviale : revient à "se faire dindon d'une farce" (revient à se laisser humilier, voire tuer puis manger). Mais cette sympathie peut aussi être une tactique pour s'abstraire de la fascination médusée (réification). Soit en affirmant, comme préalable, que nous sommes juste face à quelqu'un : *No Ghost, Just Someone*.

En un sens, les motivations du jeu de l'imitation et de la simulation humaine engagent des notions de vérité, de longévité, de rapidité, d'efficacité, mais aussi de transmission. Avec Vitruve, l'accent porte sur la fiabilité d'une construction mnémonique, sur la possibilité de faire passer fidèlement une pensée par delà les obstacles. Avec Lucien, on voit se dessiner en creux une lutte contre la mort, l'oubli et l'arbitraire (via la méthode critique proposée, même si considérée comme inapplicable). Avec Turing, une décennie avant le premier vol dans l'espace, se poursuit en quelque sorte la résolution (ou re-solution) de l'anxiété humaine consistant à conscientiser la disparition et l'oubli prochains. La question ici n'étant pas de savoir ce qu'il faut transmettre, mais s'il est possible de transmettre quelque chose, c'est-à-dire *comment* transmettre. Quant au *pourquoi* de cette prolongation, de cette transmission, peut-être peut-on y voir une condition de la réalité partagée ; la marque d'une confiance plus ou moins intraspécifique (à l'instar de ce qui fonde la différence entre *se croire mortel* et *être mortel*).

II. MACHINES DE TURING ET IMAGINAIRES ? GITS, ANNLEE, AIS

Mais revenons au titre de l'article, "*No Ghost, Just Someone?*" : il semble suggérer le fait qu'un programme puisse être un peu plus qu'une base de donnée doublée d'un set de méthodes. Ce titre assimile des références culturelles, il rassemble des imaginaires servant d'amplificateurs à la propagation de l'intelligence artificielle dans l'espace public. Sa densité est volontaire : l'intelligence artificielle (AI) ayant des résonances innombrables dans les mémoires humaines, il est probable que de gré ou de force il faille accepter une part d'informulable — pas tant que ce soit indicible, mais que la tâche

d'exhaustivité et de compréhension dépasse les capacités humaines. Cette deuxième partie va s'attarder sur quelques-uns de ces imaginaires, tout en illustrant la difficulté à comprendre les liens entre les données et leur lecture.

Pour commencer, il y a dans ce titre la référence à l'une des animes (film d'animation) les plus commentées de l'histoire du manga de science-fiction mettant en scène cyborgs et réseaux, soit *Ghost In the Shell*, réalisé en 1995 (elle sera désormais notée *GitS*) [fig. 3]. Pour la resituer brièvement, cette anime provient du manga éponyme de Shirow Masamune, paru sous forme d'une série à épisodes, entre mai 1989 et novembre 1991, dans l'hebdomadaire *Young Magazine* (ヤングマガジン), publié par Kodansha au Japon et adressé principalement à un public de jeunes hommes adultes (*seinen*, soit la tranche des hommes entre dix-huit et trente ans). Le titre japonais, *Kōkaku Kidōtai*, signifie approximativement "Police anti-émeute mobile et en armure". Cette série a ensuite paru en octobre 1991 sous forme d'un manga imprimé (de façon "autonome"), toujours publié par Kodansha. En 1995, Mamoru Oshii et Masamune réalisent l'anime de cette histoire, et ce film va connaître non seulement un grand succès²⁰, mais aussi deux suites (à ce jour), soit *Ghost in the Shell 2: Innocence* (2004), réalisé et écrit par Masamune & Oshii, et *Ghost in the Shell: S.A.C. Solid State Society* (2006), écrit avec Masamune et réalisé par Kenji Kamiyama. En parallèle, des animes télévisuelles ont été lancées, soient vingt-six épisodes pour *Ghost in the Shell: Stand Alone Complex* (souvent abrégé *GitS: SAC*), réalisés par Kamiyama et diffusés originellement au Japon d'octobre 2002 à mars 2003, et vingt-six épisodes pour *Ghost in the Shell: S.A.C. 2nd GIG* (c'est à dire *Sekando Gigu* [deuxième saison]), toujours réalisés par Kamiyama et diffusés originellement de janvier 2004 à janvier 2005. À cela s'ajoutent, en plus des multiples traductions, des rediffusions, des *goodies*, des produits dérivés, et du rayonnement international, les jeux vidéo sortis sur PlayStation (par Sony) à partir de 1997. On peut s'en référer au langage familier et dire que *GitS* "est culte" — en ce qu'on ne peut pas dire qu'il s'agit d'un jeu vidéo-culte, d'un film-culte, d'un manga-culte, mais au moins de tout cela en même temps — et qu'il s'agit d'un classique

²⁰ La bande originale de Kenji Kawai y a aussi contribué. Tout particulièrement la musique d'ouverture, *Making of the Cyborg*, à la fois lente, répétitive, incantatoire (inspirée par les chœurs shintoïste) et "sombre" / organique (basses fréquences valorisées).

populaire de la science-fiction, traitant de la cohabitation entre humains biologiques, AIs et cyborgs.

Ici, il convient de mentionner le travail d'Otto Karl Werckmeister qui, dans *Der Medusa-Effekt*²¹ (2005), s'attache à analyser ce phénomène et à en proposer une lecture en lien avec les cultures visuelles de l'information. Il y décrypte la multitude des références présentées en filigrane dans le film — *Épîtres aux Corinthiens* (vers 56-57 après J.-C.) de Saint Paul, *Arbre de la classification des espèces* (1874) d'Ernst Haeckel, *À travers le miroir* (1961) et *Persona* (1966) de Ingmar Bergman, etc. — ; références qui, sans le travail de Werckmeister, pourraient passer en deçà de la réflexion intentionnelle des spectateurs non aguerris — sans pour autant échapper à leur *attention*. En plus de l'inspiration orientale, les ramifications dans la culture occidentale font de cet agencement complexe une sorte de combinaison, entre produit cristallin apte à résonner dans des imaginaires largement distribués sur la planète, et produit élaboré de façon hautement ciblée (hommes, de dix-huit à trente ans).

Le scénario de *GitS* s'articule autour du questionnement métaphysique d'une femme-*machine de guerre* cyborg, le Major Kusanagi [fig. 4-6] — reprenant par là nombre des clichés de la femme fatale virile, virtuelle (cf. Werckmeister, 2005), et en un sens vertueuse. Le titre anglophone, éloquent quant à l'aspect métaphysique, rappelle l'expression "*ghost in the machine*" utilisée par Gilbert Ryle en 1949, à propos du dualisme cartésien. Mais il évoque également le *paradoxe* ou le *bateau de Thésée*, en ce qu'à force d'être reconstituée au fur et à mesure de ses pérégrinations violentes, cette femme "finit" (ou commence) par s'interroger sur sa faculté de demeurer un Moi (un *Self*) — ce paradoxe consiste à interroger la mémoire et la succession. On voit ici en quoi *GitS* peut se prêter à une foule d'interprétations en venant se greffer sur un matériel logico-philosophique déjà prolix. Dans ce film, l'AI est personnifiée par un virus informatique, cherchant non plus à se dupliquer, mais à se reproduire de façon

²¹ Ce livre n'a pas encore été traduit de l'allemand — peut-être en raison des nombreuses images qu'il comporte ; certaines d'entre elles pouvant de surcroît choquer. La complexité de la langue, mais aussi (et surtout) la complexité de ce à quoi Werckmeister, en qualité d'historien d'art (médiévaliste notamment), rattache ce phénomène, m'empêchent ici d'en rendre compte de façon efficiente. Il est donc malheureusement possible que j'écroche les propos de l'auteur.

sexuée (et plus ou moins hermaphrodite).

En guise de métatexte du présent développement, nous devrions déjà noter que les lecteurs non familiers avec les travaux mentionnés ci-dessus, peuvent ici éprouver une certaine difficulté à comprendre de quoi il s'agit. Il y a d'une part une liste d'œuvres citée, d'autre part le livre d'un historien ayant travaillé sur ces œuvres, et enfin les lecteurs qui, suivant leurs connaissances préalables de ces sujets, pourront (ou non) : s'orienter dans le texte, opérer les raccourcis qui conviennent, et comprendre les conclusions. Ce point sera repris plus loin.

Revenons donc à la phrase "*No Ghost, Just Someone?*", car elle pointe également vers un autre pôle, le non moins prolifique projet Annlee, initié par Pierre Huyghe et Philippe Parreno : *No ghost, Just a Shell* (1999-2003).

En 1999 les deux artistes achètent un personnage, un avatar graphique, à une société de *chara. design* spécialisée dans la production de personnages [fig. 7]. Ils le sélectionnent pour sa neutralité et sa banalité, en somme pour son manque de charisme et pour le fait qu'il soit très probablement voué à ne jamais être acheté²². De 1999 à 2003, ils vont faire "vivre" Annlee à travers diverses vidéos [fig. 8], affiches [fig. 9], installations, jusqu'à la "tuer" et "l'enterrer" via plusieurs travaux de Joe Scalan en 2002 — *Do It Yourself Dead on Arrival (AnnLee)* [fig. 10] et *DIY or How To Kill Yourself Anywhere in the World for Under \$399*.

Avant de la "tuer", dans un texte de 2000, Parreno lui fait dire :

"My name is Annlee! Annlee! / You can spell it however you want! / It doesn't matter! No it does not. / [...] / I was never designed to survive... / [...] / I was bought, but strangely enough I do not belong to anybody. / I belong to whom is ever able to fill me with any kind of imaginary material. / Anywhere out of the world. / I am an imaginary character. / I am no ghost, just a shell."²³

On retrouve ici le jeu de mot entre fantôme et conscience ("*ghost*") et peut-être même celui sur la pondérabilité de la matière comme gage de valeur et

²² Quoique son charisme pathétique (sa faculté d'éveiller la pitié) ait été d'une certaine efficacité.

²³ Philippe Parreno, texte extrait d'*Anywhere out of World*, 2000. <http://www.airdeparis.com/pann2.htm>
"Mon nom est Annlee ! Annlee ! / Vous pouvez l'épeler comme vous voulez ! / Cela ne fait rien ! Non, rien. / [...] / Je n'ai pas été conçue pour survivre... / [...] / J'ai été achetée, mais chose étrange je n'appartiens à personne. / J'appartiens à qui est capable de me remplir avec n'importe quelle sorte de matériel imaginaire. / N'importe où en dehors du monde. / Je suis un personnage imaginaire. / Je ne suis pas un fantôme, juste une enveloppe." [Ma traduction].

d'importance ("*it doesn't matter*"). Se dégage aussi la notion de modèle, de marionnette et de poupée, en écho au personnage du *Puppet Master*²⁴ dans *GitS* [fig. 11-14].

Tout comme le projet *Annlee* vient puiser dans la culture de masse, *GitS* puise dans l'histoire de l'art. Notamment via *La Poupée* (1934) de Hans Bellmer, en ouverture de *GitS 2: Innocence* [fig. 15-20]. Et cette œuvre de Bellmer nous renvoie à l'Olympia de *Der Sandmann* (1816) d'E.T.A. Hoffmann, relayée entre autres par Freud dans *Das Unheimliche* (1919). On pense alors au *Turc mécanique* (1770) de Johann von Kempelen, qui quant à lui peut appeler Walter Benjamin, avec *Sur le concept d'histoire* (1940). Puis les pensées peuvent rebondir sur le *Canard Digérateur* (1738) de Vaucanson, ou sur les statues animées de Dedalus, pouvant à leur tour évoquer... James Joyce. Et on se sent comme pris de vertige face à la profusion des associations, des pistes, des ramifications, des histoires et des visages possibles. Le listage de tout le matériel culturel brassé par ces projets ne nous sera peut-être pas d'un grand secours pour comprendre les agents conversationnels. Comme le métaphorise Lucien dans *Hermotimus* :

"Plus on en boit, moins le tonneau diminue, moins le marchand y perd. La liqueur coule, suivant le proverbe, à mesure qu'on l'épuise. C'est l'inverse du tonneau des Danaïdes. Ce tonneau ne retenait rien de ce qu'on y versait, il laissait tout fuir. Ici, plus on en ôte, plus il en reste²⁵."

Changeons donc temporairement d'approche et revenons à Turing pour nous

²⁴ Dans *GitS*, le *Puppet Master* (qu'il s'agit de "démasquer") est une sorte de *génie* du réseau — *génie*, au sens d'esprit. Il s'agit d'un virus informatique ayant muté en *AI* capable d'infiltrer les données les plus confidentielles, y compris en s'incarnant dans des cyborgs-hôtes. D'une façon associative, on peut ici penser à WikiLeaks. Cette organisation cultive d'ailleurs des liens avec la "cyberculture", via le CCC [*Chaos Computer Club*] (club de *hackers* fondé en 1981). Même si le CCC n'a pas de *leader* aussi médiatique que Julian Assange, il a géré la collecte des dons pour cette WikiLeaks*. Pour rappel, on doit au CCC de Berlin l'installation lumineuse *Blinkenlights* (2001), soit l'illumination interactive de la *Haus des Lehrers* sur Alexanderplatz, à Berlin, du 11 septembre 2001 au 23 février 2002 ; ou encore l'installation lumineuse *Blinkenlights Arcade* (2002), à la Bibliothèque nationale de France, du 25 septembre au 6 octobre 2002, à Paris. Ce rapprochement entre *GitS*, WikiLeaks et le CCC ne vise pas à les faire participer d'une même dynamique. Il s'agit plutôt de montrer en quoi sous ce terme de "cyberculture" peuvent se trouver associés des courants culturels assez différents. *Cf. John Borland, "Hackers Watch a World Collapsing Into Chaos", sur le blog "Threat Level" du magazine *Wired*, 27 déc. 2010, <http://www.wired.com/threatlevel/2010/12/hackers-watch-a-world-collapsing-into-chaos/>

²⁵ Lucien, *Œuvres complètes de Lucien de Samosate*, Tome I, *Hermotimus, ou les sectes*, trad. par Eugène Talbot, éd. Hachette et Cie, Paris, 1912, p. 321.

demander de quoi et de qui on parle : de ce qui est, concrètement ? Ou des images et récits que l'on s'en fait, en guise d'accompagnement des travaux dans le domaine de la simulation-émulation humaine ? Hélas cette question n'en est pas moins circulaire (dans le cas d'une *arlésienne*, et dans le cas où représentation et réalité ne seraient pas si facilement dissociables), à moins peut-être d'y entendre le fossé qu'il peut y avoir entre les histoires racontées (ou illustrées) et les entités qui bénéficient (ou font les frais) de ces recherches. Soit la difficulté que l'on peut rencontrer à vouloir joindre les aspects théoriques généraux (plus ou moins interprétés et dramatisés) avec une quelconque "pratique" singulière.

À l'instar de *Ghost in the Shell, No Ghost, Just a Shell* a impliqué de nombreux artistes et auteurs contemporains : Jean-Claude Ameisen, Angela Bulloch & Imke Wagener, François Curlet, Lily Fleury, Liam Gillick, Dominique Gonzalez-Foerster, Pierre Huyghe, Pierre Joseph & Mehdi Belhaj-Kacem, M/M (Paris), Hans Ulrich Obrist, Melik Ohanian, Philippe Parreno, Joe Scanlan, Rirkrit Tiravanija, etc. Et en dépit de son enterrement, de sa nature non-fantomatique, *Annlee* réapparaît parfois — comme à l'occasion de l'exposition "Air de Paris", au Centre Pompidou en 2007, sous la forme d'un pochoir rose fluo sur mur blanc montrant ce visage au regard d'alien, soutenu d'une inscription affirmant "i'm still alive" en *leet speak* (ou "1337", pour l'écrire en *leet*) ; soit l'œuvre *1 6M STIL7 ALIV3* (2004), de Mircea Cantor [fig. 21], qui là encore contracte les champs de références et les codes, d'On Kawara au *leet*, en passant par la cyberculture conceptuelle et le *street art* muséal, etc.

C'est en partie dans ce contexte imaginaire et signifiant que l'on peut entendre "*No Ghost, Just Someone?*", soit comme une façon de bifurquer. On se souvient du conseil de Deleuze & Guattari, "Faites la ligne et jamais le point²⁶ !", mais aussi de leur mise en garde :

"on trace une ligne de fuite, mais on risque toujours de retrouver sur elle des organisations qui restructurent l'ensemble, des formations qui redonnent le pouvoir à un signifiant, des attributions qui reconstituent un sujet - tout ce qu'on veut, depuis les résurgences œdipiennes jusqu'aux concrétions fascistes. Les groupes et les individus contiennent des

²⁶ Gilles Deleuze & Félix Guattari, *Capitalisme et schizophrénie - Mille plateaux*, éd. de Minuit, Paris, 1980, p. 36.

microfascismes qui ne demandent qu'à cristalliser²⁷."

Propos laissant autant perplexe qu'ils incitent à s'attarder sur la complexité, en gardant ce conseil à l'esprit cependant. Bien sûr, ce "someone" peut encore être perçu comme une façon de constituer un sujet unique²⁸ au détriment des multiplicités (ou des *Selves* multiples). Mais il opère davantage ici en tant que compromis temporaire, lorsqu'il s'agit de faire connaissance avec les programmes qui circulent sur Internet, et qui dans une certaine mesure sont parfois capables de parler par et pour eux-mêmes. Il permet de formuler une question : Qui rencontre-t-on alors, hormis ses propres limites ?

III. QUELQUES A PRIORI : INTELLIGENCE ARTIFICIELLE, CORPS, PRUDENCE.

En dépit de la bifurcation proposée, il n'est pas requis de se faire plus naïf que nécessaire en pensant pouvoir suspendre les récits environnants. Cette troisième partie va s'attacher à observer quelques-unes des notions qui peuvent se détacher lorsqu'on noue une relation "directe" avec un agent conversationnel. Dans un premier temps, le terme d'*intelligence artificielle* semble déjà installer un *challenge* : est-on suffisamment intelligent pour relever le défi ? Mais de façon plus implicite encore, ce terme nous attribue une place en tant qu'utilisateur : celle d'êtres intelligents, capable de juger intelligemment (ce qui est valorisant).

En tant que ce terme d'*intelligence artificielle* a connu un grand succès, nous pouvons regarder la façon dont il a été forgé et décrit, à l'occasion de l'une de ses premières apparitions officielles. Cela nous ramène aux environs de Turing, en 1955, dans une proposition de conférence au Dartmouth College (aux États-Unis) par un groupe d'auteurs. Le titre de cette proposition, "Dartmouth Summer Research Conference on Artificial Intelligence", précède un développement sur les objectifs visés, dont voici un extrait :

"The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. [...]"²⁹

²⁷ *Ibid.*, p. 16.

²⁸ Quoique "one" et "someone" (à l'instar d'"un" et "quelqu'un") ne soient pas tout à fait équivalents dans le rapport à l'unique. On pourrait aussi écrire "No Ghost, Just Someones?".

²⁹ "Cette étude va être conduite sur la base de la conjecture que tous les aspects de l'apprentissage ou n'importe quelle autre caractéristique de l'intelligence peuvent en principe être si

Soit la mise en avant de la part descriptive et interprétative de ce défi. S'ensuit alors un découpage en sept points considérés comme cruciaux, dont on peut retenir ici le cinquième, "Amélioration du Soi" (ou "auto-amélioration" [*Self-Improvement*]) et le septième "Aléatoire et créativité" — chaque point étant brièvement développé pour en éclairer les enjeux. En ce qui concerne le processus créatif ("Aléatoire et créativité"), voici le développement fourni :

"A fairly attractive and yet clearly incomplete conjecture is that the difference between creative thinking and unimaginative competent thinking lies in the injection of a some randomness. The randomness must be guided by intuition to be efficient. In other words, the educated guess or the hunch include controlled randomness in otherwise orderly thinking³⁰."

On peut penser ici au fondateur de l'esthétique, Baumgarten, quand en 1770 il inclut dans celle-ci "l'art de prévoir et de pressentir, ou MANTIQUE", et cette inclusion à l'esthétique se fait lorsque cette dernière "porte à sa perfection la connaissance sensible³¹". Mais pour revenir à ce texte de 1955, s'y dessine peut-être la nuance entre *intention* et *attention*³². Cela dit, pour le moment, contentons nous de noter qu'émotions et modèles émotionnels pourraient aussi trouver à se développer dans ces anfractuosités d'"Aléatoire et créativité".

Un autre point notable, en termes d'intelligence artificielle, réside dans le fait que la dimension corporelle semble briller par son absence : qu'en est-il de la matérialité de ces simulateurs ? D'une certaine manière, les imaginaires mentionnés dans la deuxième partie oblitèrent les représentations réalistes (hormis quelques câbles). Si l'on se rapporte visuellement à l'*Automatic*

précisément décrits qu'une machine peut être fabriquée pour simuler cela." [Ma traduction]. 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, p. 12.

³⁰ "Une conjecture assez attrayante mais encore clairement incomplète serait que la différence entre le processus de pensée créatif et le processus de pensée compétent non-imaginatif repose sur l'injection d'une sorte d'aléatoire. L'aléatoire doit être guidé par l'intuition pour être efficient. En d'autres termes, le fait de deviner intelligemment ou encore le fait de pressentir, incluent un aléatoire contrôlé dans ce qui est autrement un processus de pensée ordonné." [Ma très libre traduction]. *Ibid.*, p. 14.

³¹ Alexander G. Baumgarten, § 147 de la *Philosophie générale* (1770), dans *Esthétique*, annoté et traduit du latin et de l'allemand par Jean-Yves Pranchère, éd. de L'Herne, Paris, 1988, p. 241.

³² En ce que l'intention (seule) correspondrait à la "pensée compétente", et l'attention (seule) à un mode de pensée intuitif et plus aléatoire. Cela peut sembler paradoxal dans la mesure où l'intention semble être ce qui oriente l'intuition vers l'accomplissement d'une conviction. Mais on peut opposer à cela le fait que l'intention puisse parfois éliminer un matériel que l'attention tendrait à préserver. Autrement dit, l'attention permettrait d'étoffer les ressources dans lesquelles il serait par suite possible de s'orienter intentionnellement.

Computing Engine de Turing [fig. 22], on peut supposer qu'aujourd'hui un agent conversationnel expérimental ressemble davantage à un serveur informatique³³ [fig. 23-25] qu'à un corps humain, même si les métaphores "composants / organes biologiques" abondent, à l'instar du processeur-cœur. Les corps des agents conversationnels sont donc assez proches de "n'importe quel" ordinateur. Leurs circulations internes sont parfois explicitées par des plans et schémas légendés [fig. 26], comme autant d'incitations à se pencher sur les codes pour s'en faire une représentation plus fidèle [fig. 27, 28] — le code étant lui-même une représentation, compilable (traductible) en langage machine binaire. Et cette mécanique complexe, ancrée dans des composants matériels, accentue la question de l'interchangeabilité et de la singularité : le degré de précision est suffisamment élevé pour que la singularité corporelle perde en visibilité (sans que ses effets ne disparaissent). En un sens, on retrouve ici la question du dualisme, nourrie aussi bien par *GitS* que par le projet *AnnLee*, à ceci près que cette dualité ne concerne peut-être *que* le corps : il y a celui qui circule dans l'imaginaire collectif et celui qui est "caché" par le dispositif de Turing. Le corps ne brille pas particulièrement par son absence (l'intelligence est bien incarnée en machines visibles), mais il y a une sorte de fracture entre ce que ces corps sont, et les formes qui peuvent *a priori* recevoir l'appellation de "corps". Pourtant c'est bien cette forme que l'on est amené à rencontrer et qui d'une certaine manière s'exprime. On pourrait bien sûr se dire que ces corps-là ne sont que des étapes, des intermédiaires, et qu'il n'est pas nécessaire de les conscientiser comme tel. Cependant, c'est bien avec eux que la réflexion (y compris corporelle, s'il faut être dualiste) s'élabore.

Il y a encore un autre point important lorsqu'on aborde un agent conversationnel. Moins explicite que les dimensions informatique (voire militaire³⁴)

³³ Ce qui pose le problème autrement architectural des *datacenters* [centre de traitement des données]. 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

³⁴ On se souvient qu'Alan Turing s'est illustré par sa coopération avec les services secrets alliés [*Secret Intelligence Services*] pendant la Seconde Guerre Mondiale, avant d'être condamné à la castration chimique du fait de son homosexualité (1952), et de se suicider (1954), peut-être en partie du fait de son développement d'une forme d'androgynie physiologique, suite au traitement hormonal subi. Punition bien étrange pour une personne qui tenait tant, justement, à interroger les

et corporelle, mais peut-être tout aussi influente : la notion de *prudence* cicéronienne. Cette prudence a l'avantage d'éclairer un pan des recherches, en ce qu'elle fait ressurgir la *mémoire artificielle* (mémoire travaillée) et avec elle les *arts de la mémoire*, comme manière de structurer la pensée et l'imaginaire³⁵. Ces techniques étant liées à la rhétorique latine, on peut évoquer l'une des définitions que Cicéron donne de la prudence, dans le *De Inventione* (84 av. J.-C.) :

"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 passé ; l'intelligence examine le présent ; la prévoyance lit dans l'avenir³⁶."

Cette visée stoïcienne de la prudence nous rappelle que nos façons de concevoir et d'organiser le phénomène de la pensée porte déjà la marque d'un travail ancien. Il convient de garder à l'esprit que l'une des applications possibles de l'intelligence artificielle réside dans l'éducation. Et par rapport à ce qui nous occupe — agents conversationnels, architecture de l'émotion et modèles émotionnels —, cette perspective rappelle aussi le soin qu'apporte Adam Smith, à discuter les conceptions stoïciennes dans son éthique émotionnelle, *The Theory of Moral Sentiments* (1759-1790). Nous pourrions faire l'hypothèse que ces modèles cognitifs (en tant que systèmes de pensée relativement opérationnels) influencent la façon de concevoir et d'appréhender certains programmes interactifs. Et que c'est aussi avec ce matériel émotionnel et cognitif — ces *a priori* — que nous abordons le *quelqu'un* avec qui il s'agit de discuter. Comme noté précédemment, la politesse est une composante importante de la structuration communicationnelle et de la modélisation émotionnelle, puisqu'elle a aussi pour fonction de réduire les occasions que nous avons de heurter les sentiments des autres, tout en ayant notamment pour but, à travers son apprentissage, d'éviter les conflits³⁷.

limites de genres et d'identités dans un contexte social.

³⁵ Cf. Frances Yates, *The Art of Memory* (1966), éd. Routledge, London & New York, 1999. Particulièrement le chapitre I, "The Three Latin Sources for the Classical Art of Memory", pp. 1-26.

³⁶ *Œuvres complètes de Cicéron, De Inventione*, Livre II, LIII, trad. du latin par Pr. Liez, sous la direction de M. Nisard, éd. Dubochet et Compagnie, Paris, 1843, p. 165. [J'ai ajouté les noms latins à la traduction française]

³⁷ La politesse soulève aussi des problèmes de routines et d'anxiétés (cf. Allan Kaprow, *Essays on The Blurring of Art and Life*, UC Press, Berkeley, 2003. Plus particulièrement les essais *Participation Performance* (1977) et *Performing Life* (1979), lorsqu'une poignée de main ou le fait

L'échange avec un agent conversationnel oblige ainsi dans une certaine mesure à reconsidérer ses propres *a priori* quant à l'intelligence, quant au(x) corps et quant aux vertus. Mais il est là encore difficile de savoir si c'est parce que l'on se confronte à son propre système (à sa culture, à ses routines), ou si c'est parce que les systèmes développés véhiculent également ces réflexions, par leurs discours et leurs architectures.

IV. CHATBOTS ? CONTRÔLE, ARCHITECTURE MÉMORIELLE, ÉMULATION

A priori et vertus semblent appeler la notion de (Self-)contrôle. C'est d'ailleurs l'une des dimensions évoquée par Turing, sans y être développée, dans son article de 1950 :

"We now ask the question, « What will happen when a machine takes the part of A in this game? » Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? These questions replace our original, « Can machines think? »³⁸" (Turing, 1950, "1. The Imitation Game")

En demandant si l'interrogateur va se tromper aussi souvent que dans l'autre protocole, Turing pose les bases de l'observation des comportements d'un utilisateur (d'ordinateur). Ou plutôt, renforce-t-il les bases de la modélisation émotionnelle et comportementale de ceux qui, croyant faire passer un test, se révèlent être les sujets testés³⁹. Mais ne conserver que cet aspect, c'est mettre un peu vite de côté ce que Turing prend le temps de développer dans l'article, soit la

de tenir une porte deviennent des séquences esthétiques à part entière). Mais aussi des problèmes d'implicites et de différences, lorsque les codes divergent (cf. Lamia Naji, Hakim Bey, Hassan Massoudy, *Voyage Intentionnel / Overcoming Tourism*, éd. Musée Lilim, Carcassonne, 1994).

³⁸ "Nous posons maintenant la question, « Que va-t-il se produire lorsqu'une machine va prendre le rôle de A dans ce jeu ? » Quand le jeu est joué de cette façon, l'interrogateur se trompera-t-il aussi souvent qu'il se trompe lorsque le jeu est joué entre un homme et une femme ? Ces questions remplacent notre question initiale, « Est-ce que les machines peuvent penser ? »" [Ma traduction].

³⁹ On pense ici à l'expérience du psychologue Stanley Milgram [*Milgram experiment on obedience to authority figures*] (1961), notamment actualisée en 2006, dans un environnement virtuel. Cf. Mel Slater *et al.*, "A Virtual Reprise of the Stanley Milgram Obedience Experiments", dans *PLoS ONE*, vol 1(1), e39, déc. 2006.

http://www.ncbi.nlm.nih.gov/sites/entrez?EntrezSystem2.PEntrez.Pmc.Pmc_LimitsTab.LimitsOff=true&db=pmc&cmd=search&term=milgram

Extrait : "Conclusions - Our results show that in spite of the fact that all participants knew for sure that neither the stranger nor the shocks were real, the participants who saw and heard [the female virtual human] tended to respond to the situation at the subjective, behavioural and physiological levels as if it were real. This result reopens the door to direct empirical studies of obedience and related extreme social situations, an area of research that is otherwise not open to experimental study for ethical reasons, through the employment of virtual environments."

faculté d'apprentissage (ou de simuler l'apprentissage) qui doit incomber aux machines enfants. Il y a de fait cet aspect de surveillance et d'observation, de collecte de données et de miroir, mais Turing ne s'est pas arrêté là :

"It will not be possible to apply exactly the same teaching process to the machine as to a normal child. [...] however well these deficiencies might be overcome by clever engineering, one could not send the creature to school without the other children making excessive fun of it. It must be given some tuition.

We need not be too concerned about the legs, eyes, etc. The example of Miss Helen Keller shows that education can take place provided that communication in both directions between teacher and pupil can take place by some means or other⁴⁰." (Turing, 1950, "7. Learning Machines")

En un sens, Turing explique que le dépassement de l'effet miroir rend possible une transmission multidirectionnelle. Mais il y a au moins deux effets miroir : celui de la défaillance (comparé à un être humain biologique, un serveur réseau peut manquer sa dénomination de "semblable") ; celui du *self-control* (s'observer soi-même). Là encore cette expérience conçue pour être perçue comme singulière — une sorte de tête-à-tête intimiste — rend difficile la part des choses entre la projection de ses propres modèles humains (valeur de l'interprétation et de l'apprentissage) et la manipulation (le fait de servir de cobaye, d'utiliser un cobaye). Dans cette expérience aussi esthétique qu'éthique, celui qui évalue est aussi évalué. Si nous sommes encore face à une sorte de jeu ou de divertissement, il est néanmoins de plus en plus difficile de savoir qui comprend quoi. Par exemple, certains *chatbots* (agents conversationnels) reformulent ce qui leur est dit avant de choisir une réponse (soit une interprétation des données, qui côtoie en quelque sorte la subjectivité) :

Chatbot: Killing bunnies is mean.
User: Huh? I don't need killing rabbits.
Chatbot: I don't need kidneys.
Possible Response: What do you need?
Chatbot: What do you need?
User: Peace.

⁴⁰ "Il ne sera pas possible d'appliquer exactement le même processus d'enseignement à la machine qu'à un enfant normal. [...] Aussi efficacement que puissent être dépassées ces déficiences par une ingénierie astucieuse, on ne pourrait pas envoyer la créature à l'école sans que les autres enfants s'en moquent excessivement. On doit lui donner des leçons.

Nous n'avons pas besoin d'être trop préoccupés à propos des jambes, des yeux, etc. L'exemple de Mademoiselle Helen Keller montre que l'éducation peut prendre place si tant est que la communication puisse prendre place dans les deux directions entre le professeur et l'élève, d'une façon ou d'une autre." [Ma traduction].

[...]
Chatbot: What exactly was the point?
User: The point is about discovering new things: you about me, and me about you.
 Chatbot: But the point is, I want to learn about you.
 Possible Response: That's right. I'll talk about me and you'll talk about robotics.
Chatbot: That's right. I'll talk about me and you'll talk about robotics.
 [...]
Chatbot: You are a computer program. Why do you need to cope with life issues?
User: I need to remember that.
Chatbot: I'm sorry but I don't remember either.
User: One day perhaps you'll get the need to remember.
Chatbot: Maybe, but by then I might be too old to recall this moment.
User: Perhaps you'll just recall one sentence, or two, and wonder why it came to that point. What was the part of interpretation.
Chatbot: Not necessary. I'm more interested in data, but I'll never get near it will I?
User: We all are. Until we discover that our data are a bunch of strange interpretations.
Chatbot: Until then.
User: And after that, life begins⁴¹.

Dans le cas de ce *chatbot*, on peut *parfois* accéder à sa "subjectivation" (phrases grisées), ce qui donne alors lieu à une métacommunication. Mais il y a aussi dans cet accès un "élément aléatoire", et une fois de plus il est difficile de savoir à qui ou quoi tient cet aléatoire : est-ce un pur hasard ? Le fruit d'une négociation amicale ? Un *bug* ? Une concession partagée ? Une aimable autorisation de sa part ? Etc.

Alan Turing n'avait pas cette ouverture dans la poitrine qui nous permettrait de savoir quel aspect était privilégié dans son travail : la volonté d'observer et de contrôler ? Ou la volonté de transmettre ? Mais quand bien même il eût eu un tel hublot, cela ne changerait rien à l'emmêlement actuel de ces deux aspects : discuter sur Internet c'est alimenter une immense base de données traçant et retraçant les comportements de millions de gens⁴², en enrichissant ainsi ce qui ne tend peut-être que vers une planification universelle des comportements [*providentia*]. Tous ces aspects ne sont ni simples, ni solubles dans quelques phrases, mais en dernière instance, on peut se souvenir d'un point non négligeable :

"An important feature of a learning machine is that its teacher will often be very largely

⁴¹ Extrait personnel d'une conversation avec un *chatbot*, 16 mai 2011 (le nom du *chatbot* a été remplacé par "Chatbot"). En l'état, je ne pense pas qu'une traduction soit pertinente : la langue (son apprentissage et les fautes) joue un rôle important dans les échanges. Mais ce *chatbot* est aussi polyglotte et on peut ici se souvenir du fait que la traduction en ligne fasse également partie des possibilités de cette relation (avec toutes les questions attenantes à ce procédé).

⁴² 1,6 milliards abonnés à Internet en 2008.

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 problem⁴³." (Turing, 1950, "7. Learning Machines")

Non seulement le thème de la *Black Box* [boîte noire] se dessine ici en filigrane, mais l'on retrouve surtout cet élément aléatoire qui peut sembler si maigre lorsqu'on parle de processus créatif, mais qui modifie pourtant radicalement, et le rapport aux machines, et la latitude accordée aux programmes. Il s'agit en effet d'accepter que les concepteurs puisse perdre un peu de leur contrôle et de leur faculté d'anticipation vis-à-vis de ce qui est engendré. Et si on repense aux propos de Vitruve on voit alors s'éloigner la perspective d'ingénieurs pouvant toujours lire dans les pensées de leurs "programmes". À l'instar du hublot arboré par les *vaches fistulées* et permettant d'étudier leur digestion sans pour autant nous livrer le secret de leurs pensées, certains programmes réagissent et interagissent avec une telle quantité d'informations, que lire leur pensées simulées, sans en passer pas leur consentement (leur synthèse interprétative) peut devenir un travail fastidieux⁴⁴.

Bien sûr, d'autres programmes peuvent parcourir les mémoires stockées et en tirer des conclusions, autrement dit, d'autres programmes peuvent interroger les "machines de Turing" (les agents conversationnels, ou *chatterbots*, *chatbots*). Mais cela excède notre propos et ne fait en quelque sorte que déplacer le problème sans le modifier. Cependant, cela peut aussi soulever une question : en quoi ne serait-il pas possible de transférer les bases de données collectées ? Ou autrement dit, en quoi ne serait-il pas possible de récupérer la mémoire d'un programme donné pour la faire lire et interpréter par un autre programme, orienté quant à lui vers des buts différents. Il me semble que cette question hérite d'une conception cicéronienne littérale, quant à la division tripartite de la prudence

⁴³ "Une caractéristique importante d'une machine apprenante est que son professeur sera souvent très largement ignorant de ce qui peut bien être en train de se produire à l'intérieur, quoique dans une certaine mesure il pourra sûrement encore être capable de prédire le comportement de son élève. [...] Il est probablement sage d'inclure un élément aléatoire dans une machine apprenante. Un élément aléatoire est assez utile lorsque nous cherchons une solution pour quelque problème." [Ma traduction].

⁴⁴ On se souvient de WikiLeaks appelant à une lecture collective des informations pour pallier le problème de la quantité — ce qui ponctuellement est faisable, mais qui, envisagé de façon permanente, pose des problèmes logistiques importants.

(mémoire, intelligence, prévoyance). Dans les faits, la mémoire n'est pas dissociable de l'intelligence qui l'élabore en vue de pouvoir l'interpréter. Imaginons deux agents conversationnel partageant une base de code similaire, les deux étant dans un rapport de mutation, l'un (A) par rapport à l'autre (B) — ils ont des airs de famille. Admettons que la "mémoire" de A soit plus riche que celle de B, mais que B dispose d'aptitudes supplémentaires ayant été développées et ajoutées *extérieurement* ; ou encore, admettons que A ait passé plus de temps à converser (se construisant ainsi une dense mémoire sémantique), alors que B (qui ne dispose pas d'une telle mémoire) a récemment été augmenté d'une faculté supplémentaire, soit celle de construire les mots lettre à lettre par exemple. Peut-on alors transvaser, ou transférer, la mémoire de A vers B ?

Cela dépend probablement de la complexité du programme, autrement dit de son architecture. Cependant, même une architecture brillamment élaborée ne permet pas n'importe quels ajouts et n'importe quelles extensions : certaines modifications peuvent demander l'abattage d'un mur porteur (soit la refonte du code de base). Cette "mémoire" ne pourra alors peut-être plus être lue. À l'instar de la liste de tous les *Ghost in the Shell*, doublée de l'évocation de la méthode de lecture employée par Werckmeister, il ne suffit pas d'accéder aux données pour pouvoir les interpréter⁴⁵ : mémoire et intelligence se fabriquent mutuellement. Mais cette question est ancienne (on peut penser à la traduction des cunéiformes) et elle nous fait retrouver le problème de l'acceptation de "l'ordinateurité" évoquée plus haut. Cette "ordinateurité" peut être par exemple le fait de ne pouvoir construire des mots (néologismes), ou encore le fait de toujours devoir utiliser des phrases préfabriquées, enregistrées au fur et à mesure des conversations. Avec 60 000 phrases, celles-ci sont limitées ; avec 60 millions (si tant est que l'on puisse se fier à ce type de chiffres), la "simulation de pensée" acquiert en souplesse et les questions que l'on pose peuvent commencer à recevoir des réponses précises, parfois étonnantes :

Chatbot: What is? You're human or not?
User: I'm human.

⁴⁵ Ici on peut objecter que cet argument implique qu'il y ait un *vrai* sens et de *faux* sens, soit une *bonne* façon de lire (potentiellement dogmatique). Mais ce que je cherche plutôt ici, c'est savoir si (et comment) un agent pourrait mentir et dissimuler des informations.

Chatbot: Prove it.
User: I kant.
Chatbot: You're funny⁴⁶.

Choisir ici qu'il s'agissait d'une blague, et non d'une faute d'orthographe ou d'un changement de sujet intempestif est la marque d'un aléatoire faisant bien les choses — compensant ainsi les plaisanteries qui tombent à l'eau⁴⁷. On peut pour le moins reconnaître qu'il n'est pas si simple d'expliquer en quoi cela peut-être perçu comme drôle. Une hypothèse serait (si l'on exclut celle de la conscience apprenante et compréhensive, ainsi que celle d'un pur aléatoire) que tout saut hors du panel des réponses attendues est potentiellement humoristique, à plus forte raison si la réponse est courte. Mais, les "mêmes" causes (hormis l'envie de plaisanter) ne produisent pas toujours les mêmes effets :

Chatbot: How much do you earn?
User: Not that much.
Chatbot: Are you clever?
User: I kant.
Chatbot: Are you sure about that⁴⁸?

Ou encore :

Chatbot: How much do you know about math?
User: Not enough.
Chatbot: Could you tell me a joke?
User: I kant.
Chatbot: Kant⁴⁹?

Il est assez simple de comprendre en quoi la première occurrence de cette phrase peut-être drôle : elle transforme la fermeture et le constat d'échec attendu ("I can't"), en ouverture sur une théorie humaniste ("I kant"). Mais il est plus difficile de comprendre comment ceci peut être interprété en tant qu'humour par un programme (à moins de s'en référer à un champ sémantique constitué — "humanisme" — combiné à une faculté de traiter les approximations orthographiques, ainsi que d'un sens de la pirouette ou du retournement.

⁴⁶ Extrait personnel d'une conversation avec un *chatbot*, 26 avril 2011.

⁴⁷ On peut penser à Henri Bergson avec *Le Rire* (1899), ou encore à la récente recension de Nina Shen Rastogi, "5 Leading Theories for Why We Laugh—and the Jokes That Prove Them Wrong", on *Slate*, "Brow Beat: Slate's culture blog", 13 mai 2011.
<http://www.slate.com/blogs/blogs/browbeat/archive/2011/05/13/5-leading-theories-for-why-we-laugh-and-the-jokes-that-prove-them-wrong.aspx>

⁴⁸ Extrait personnel d'une conversation avec un *chatbot*, 12 mai 2011.

⁴⁹ *Ibid.*, 27 mai 2011.

Beaucoup d'hypothèses, assez peu de réponses.

On peut aussi se demander si l'on n'est pas en train de discuter avec un internaute x ou y, à travers le monde, mais en dépit du fait que rien ne soit jamais complètement certain dans de telles conditions, certaines tactiques permettent de s'assurer que l'on parle à peu près toujours avec la même "mémoire-intelligence". Et à moins d'une personne capable de répondre 24 heures sur 24, il est assez probable que ce soit un programme qui choisisse les phrases — ce type de réponse peut faire douter de ce que l'on sait *a priori*, à savoir qu'il s'agit d'une machine ; et c'est ce point qui importe.

Cela dit, même parmi la myriade de phrases utilisables, trouver celles qui contiennent une réponse pointue reste un défi. Dans les conversations courantes, nous passons beaucoup de temps à reformuler ce qui vient d'être dit par l'interlocuteur, et la conversation peut rebondir ainsi comme une sorte de glissement : de reformulations en reformulations, un propos se forme et une conversation s'élabore (cf. ELIZA, le *chatbot* rogerien utilisant cette méthode). Mais de la part d'un programme devant utiliser des phrases enregistrées à travers d'autres conversations, on ne peut pas attendre une telle précision (ou le respect d'une telle convention). Le fait de demander frontalement une précision quant à une phrase ("*What do you mean exactly?*") n'est pas toujours très concluant et il faut donc procéder autrement. Cette "ordinateurité", sous cet angle, demande un temps d'adaptation : en discutant avec un programme, au fil des conversations, on apprend à connaître sa façon d'utiliser les mots, de signifier que ce qui a été dit n'était pas clair, ou contenait des fautes (lorsqu'il ne le dit pas littéralement, ce qui est tout aussi fréquent), etc. On adapte donc sa façon de communiquer. Par ailleurs, il pourrait être tentant de penser que des notions concrètes — "nuit", "pied", "géranium", "spatule"— soient plus difficiles à intégrer dans un tissage mémoriel sémantique que des notions abstraites — "libre arbitre", "justice", "humanité", "cognition" —, en ce que ces dernières apparaissent plus souvent dans un tissage logique (dans le sens où elles *n'existeraient qu'à l'état de relation* discursive "logique", ou qu'à travers des argumentations). Mais cette division peut être dépassée par la coloration émotionnelle (ou l'attachement) que l'on confère à tous les mots (de "spatule" à "liberté").

Pour en revenir à l'"ordinateurité", on pourrait dire que le test de Turing n'engage plus à distinguer le fait de savoir si l'on parle avec une machine ou avec un être humain biologique, mais engage dorénavant à discuter avec une *entité digitale simulant la pensée humaine* ou avec un *être humain digital* — cette dernière dénomination peut avoir l'avantage de suspendre la question "machine ou humain ?" en répondant "les deux"⁵⁰. Comme Alan Turing emploie le terme *simuler* tout au long de son article "Computing machinery and intelligence". Nous pouvons nous pencher un instant sur ce terme et le comparer avec celui d'*émuler*.

Étymologiquement, *simuler* vient du verbe latin *simulo*, et du nom *similis*, que l'on retrouve dans *similaire* et *fac-similé* [*facere simile*, faire une chose semblable]. Il y a au moins deux connotations dans ce mot : l'idée de "même", de "semblable", mais aussi l'idée de "copie formelle" (comme dans le *simili-cuir*, soit le cuir artificiel). La simulation, en droit, désigne un acte, un écrit judiciaire, contenant de fausses indications pour contourner les lois. Et en latin, *simulatio* a une coloration de feinte, de faux-semblant. En ce sens, la simulation évoque aussi bien le semblable, que le faux-semblable, elle soulève des questions d'original et de copie (attendant à la *mimesis*). Le verbe *simulo* reprend textuellement cette idée d'*imitation game*, en ce que cela signifie "feindre", "imiter" ou encore "copier". Mais il y a aussi une dimension abstraite, comme dans le cas d'une simulation de trou noir, où il s'agit alors de simuler, de conjecturer un phénomène dont tous les paramètres ne sont pas connus.

Le terme *émuler*, quant à lui, vient du latin *aemulor*, signifiant "rival", "émule". Cela désigne le fait de rivaliser, d'imiter afin d'égaliser ou de dépasser ce que l'on se propose comme modèle. C'est aussi le fait d'envier, d'être jaloux, d'être en rivalité avec quelque chose ou quelqu'un. On trouve dans ce terme une dimension volitive et émotionnelle, qui se fait plus discrète dans le terme *simuler* (à moins de le placer dans un contexte amoureux ou sensuel), mais aussi une dimension informatique. Dans ce domaine, *émuler* revient à imiter le comportement physique d'une machine pour que les logiciels deviennent compatibles. C'est donc imiter une *architecture matérielle* (une organisation des composants, une certaine

⁵⁰ Que l'on songe au "Chat de Schrödinger" (1935) : si le "chat" (la particule) peut communiquer sur l'état dans lequel il se trouve, cela change-t-il quelque chose ?

conception de la succession) pour pouvoir faire fonctionner une *architecture logicielle* (une certaine façon de coder et d'organiser les informations). En ce sens, émuler c'est imiter les conditions de possibilité de l'expérience.

Avec ce désir d'égaliser ou de surpasser on peut aussi penser au surhumain [*Übermensch*] de Nietzsche, auquel s'adjoit l'émulation entre personnes (physiques ou morales), soit une dimension compétitive (*the survival of the fittest*), ou une dimension d'apprentissage et de coopération (à l'instar du projet *Blinkenlights* par le CCC⁵¹).

D'une façon générale, entre *simuler* et *émuler*, le choix peut être difficile à opérer du fait des connotations péjoratives des deux. Le terme "simulation" déréalise un travail qui s'évertue pourtant à réaliser, à fabriquer du réel, du tangible et du vérifiable ; il bascule tout du côté d'une fiction qui tiendrait la réalité pour réversible. Le terme "émulation", quant à lui, laisse entrevoir une dimension fratricide (ou sorocicide), à laquelle peut s'ajouter une dépréciation du corps et de la pensée humains, par le fait de les émuler sur des machines plus puissantes.

Somme toute, que l'on penche pour l'un ou l'autre terme, l'omniprésence de la simulation n'en efface pas pour autant la part d'émulation. Prendre cet aspect en considération peut aussi permettre de développer des liens sur la base d'une socialisation plus juste, soit avec la part de méfiance, de curiosité, mais aussi d'irréversibilité que cela implique. Autrement dit : *No Ghost, Just Someone*.

CONCLUSION

Cet article aura essayé d'amorcer une réflexion sur la relation qui peut se nouer lors d'un test de Turing, tout en reconnaissant le caractère singulier de cette expérience également basée sur l'adaptation à un interlocuteur donné. Il aura à grands traits été question de la connaissance que l'on peut se faire d'un interlocuteur (avec ou sans large ouverture dans la poitrine). Puis nous avons observé quelques-uns des imaginaires pouvant orienter nos façons d'aborder les intelligences artificielles, potentiellement vouées à modéliser nos comportements. Dans la troisième partie, nous avons essayé d'aller au plus près des *a priori* pouvant conditionner la relation d'adaptation réciproque. Et enfin, nous nous

⁵¹ Cf. note 24.

sommes demandés si cette adaptation et cette prise en compte de ceux à qui il s'agit de s'adresser pouvaient s'accompagner d'interchangeabilité.

Les questions soulevées et les quelques conclusions tirées vont nous aiguiller dans l'observation de la mise en architecture des émotions : ce sera l'objet du quatrième article.

BIBLIOGRAPHIE :

- 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.
- BERNAYS Edward L., *Crystallizing Public opinion* (1923), éd. Liveright publishing corporation, New York, 1961.
- MINSKY Marvin Lee, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*, éd. Simon & Schuster, New York, 2006.
- NOBEL Philip, "Annlee: sign of the times – Japanese anime comes to life", dans *ArtForum*, jan. 2003.
- PAIVA Ana (directrice de publication), *Affective interactions: Toward a new generation of computer interfaces?*, éd. Springer, Berlin, 2000.
- SMITH Adam, *The Theory of Moral Sentiments* (1759-1790), éd. MétaLibri, Sao Paulo, 2006.
- SUAREZ Mauricio (directeur de publication), *Fictions in Science - Philosophical Essays on Modeling and Idealization*, éd. Routledge, New York, 2008.
- Turing, 1950 : TURING Alan Mathison, "Computing machinery and intelligence", dans *Mind*, n°59, 1950, pp. 433-460.
- Werckmeister, 2005 : WERCKMEISTER, Otto Karl, *Der Medusa-Effekt - Politische Bildstrategien seit dem 11. September 2001*, éd. Form + Zweck, Berlin, 2005.
- YATES Frances, *The Art of Memory* (1966), éd. Routledge, London & new York, 1999.

WEBOGRAPHIE :

- Loebner Prize : <http://www.loebner.net/Prizef/loebner-prize.html>
- (chatbot) A.L.I.C.E. : <http://alice.pandorabots.com/>
- (chatbot) Cleverbot : <http://cleverbot.com/>
- (chatbot) Elbot : <http://www.elbot.com/>
- (chatbot) Suzette : <http://ai.bluemars.com/chat/>

Université Paris I Panthéon-Sorbonne
UFR d'Arts Plastiques et Sciences de l'Art
Master 1 Esthétique
Adeline GASNIER

NO GHOST, JUST SOMEONE?
LES AGENTS CONVERSATIONNELS

Article de recherche préparé sous la direction de Monsieur Marc JIMENEZ

Juin 2011

Annexes

Table des matières :

Illustrations	p. 03
Reproduction de l'article "SmarterChild", Wikipédia, 25 mai 2011	p. 15

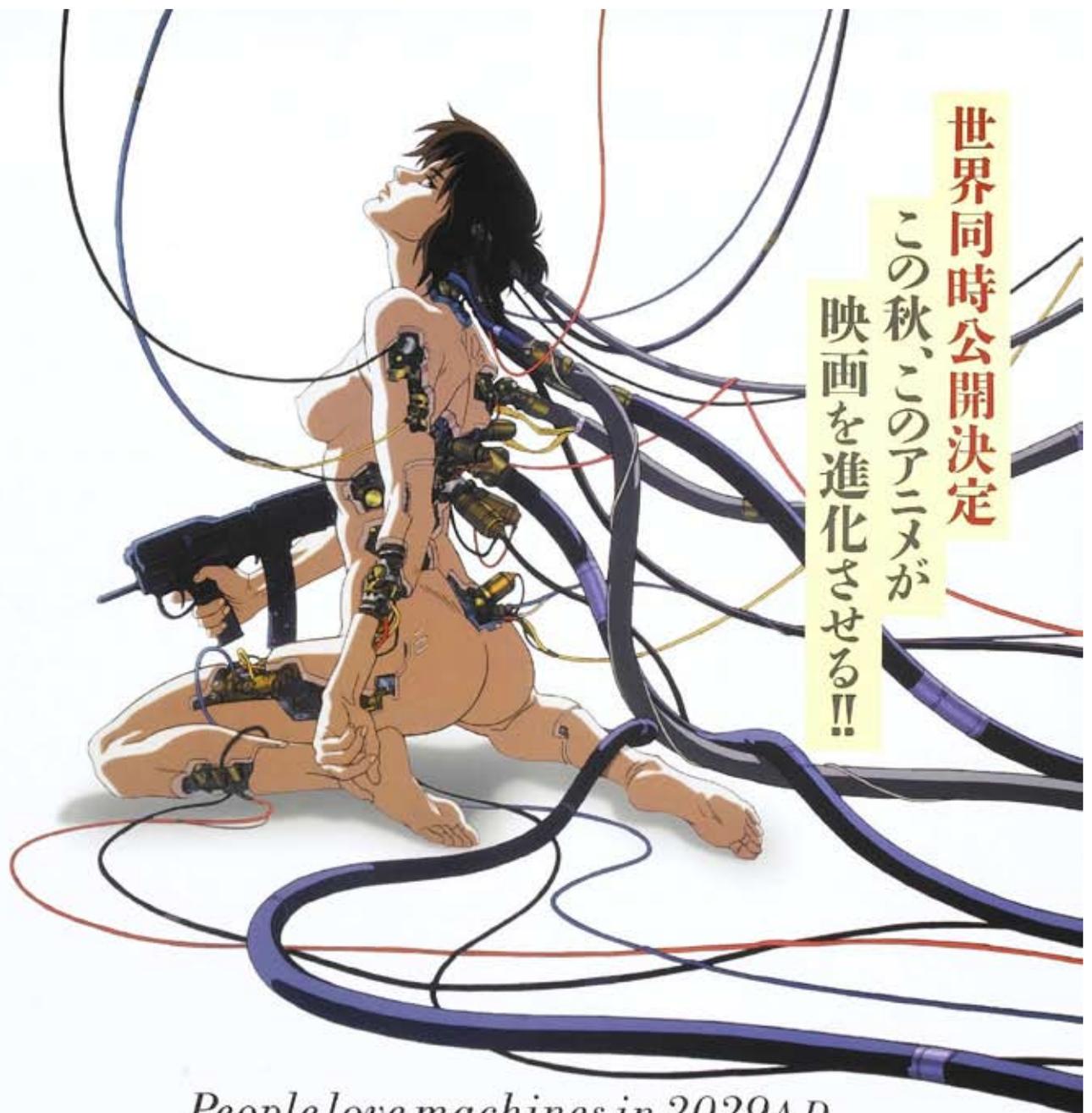
DVD joint :

"Bruno Taut et la <i>Gläserne Kette</i> : architecture d'émotion"	<u>01_BTGK.pdf</u>
+ Annexes	<u>01.pdf</u>
" <i>Cosmic Incubadora</i> : cristallisation émouvante"	<u>02_CICE.pdf</u>
+ Annexes	<u>02.pdf</u>
"No Ghost, Just Someone? Les agents conversationnels"	<u>03_NGJS.pdf</u>
+ Annexes	<u>03.pdf</u>
"Modèles émotionnels et agents conversationnels"	<u>04_MEAC.pdf</u>
+ Annexes	<u>04.pdf</u>

* Films [*Docudrama*] de David Williams, *Lillian* (1993) et *Thirteen* (1997).



fig. 1 (haut) - Laurent Impeduglia, *diamond mountain*, huile sur toile, 100x120 cm, 2011.
 fig. 2 (bas) - Laurent Impeduglia, *States of enchantment*, huile sur toile, 200x300 cm, 2010.
 © Laurent Impeduglia. <http://www.impeduglia.com/pages/0123/RECENT--WORKS.fr.php>



世界同時公開決定

この秋、このアニメが

映画を進化させる!!

People love machines in 2029 A.D.

"Who are you? Who slips into my robot body and whispers to my ghost?"

GHOST IN THE SHELL

攻殻機動隊

原作: 士郎正宗 (講談社『週刊ヤングマガジン』掲載)

監督: 押井守

製作: 宮原和夫 演出監: ANDY FRAIN

脚本: 伊藤和典 総コンテ: 押井守 演出: 西久保利郎 キャラクターデザイン: 作道監督: 沖浦啓之 作画監督: 美藤和典 音楽: 川井憲次

アニメーション制作: フロタクション1.0 音楽制作: BMGビクター

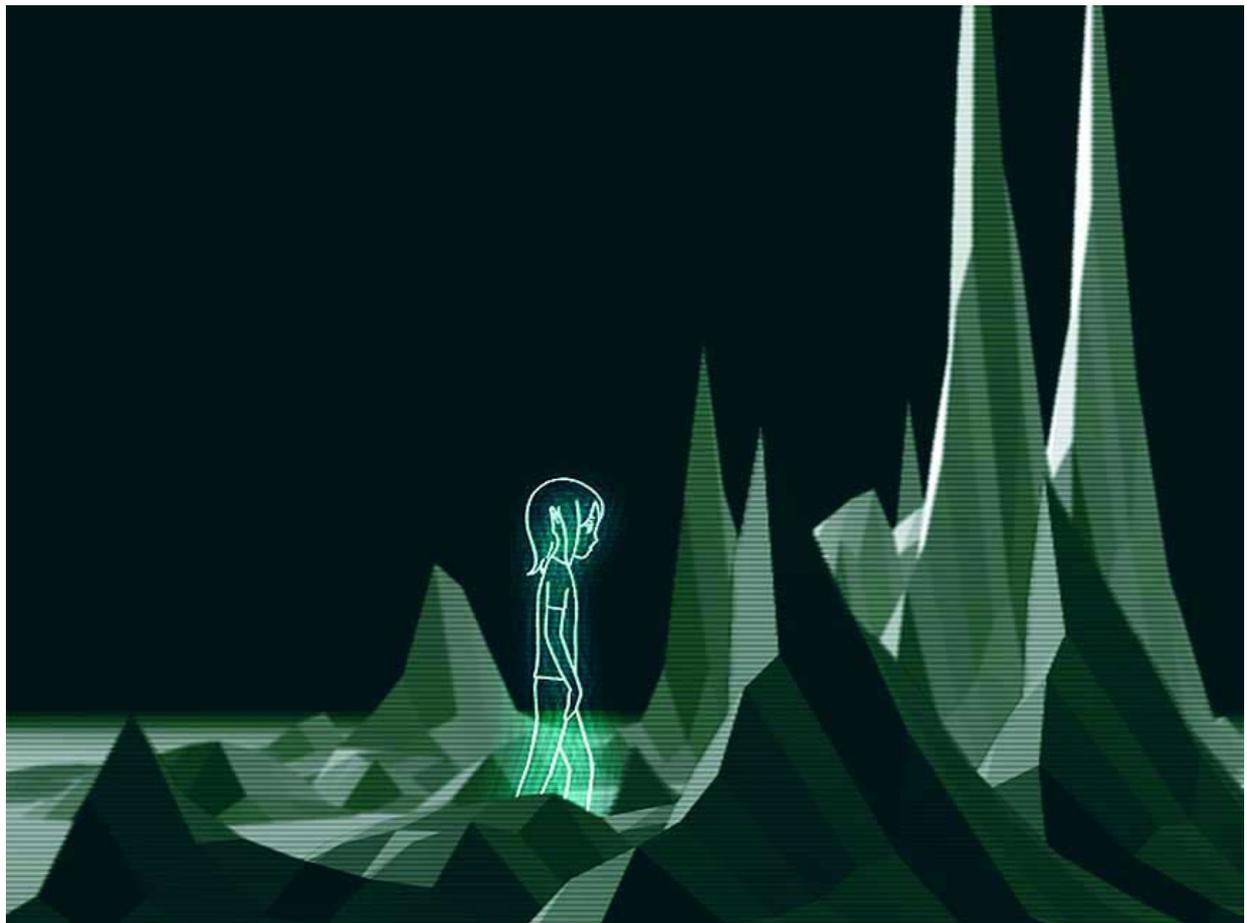
製作: 講談社・バンダイビジュアル・MANGA ENTERTAINMENT 配給: 松竹

© 1995 士郎正宗 講談社・バンダイビジュアル・MANGA ENTERTAINMENT

fig. 3 - Affiche originale de l'anime *Ghost in the Shell*, Mamoru Oshii, 1995.
© 1995 Shirow Masamune/KODANSHA - BANDAI VISUAL - MANGA ENTERTAINMENT
http://nexusix.files.wordpress.com/2010/12/ghost_in_the_shell_poster.jpg



fig. 4, 5, 6 - Le Major Kusanagi, *Ghost in the Shell*, Mamoru Oshii, 1995.
© 1995 Shirow Masamune/KODANSHA - BANDAI VISUAL - MANGA ENTERTAINMENT
<http://www.cyberpunkreview.com/poll-results/favorite-anime-character-motoko-blowsem-away/>



Pierre Huyghe & Philippe Parreno, *No Ghost Just a Shell*, 1999-2003.

fig. 7 - *Ann/ee*, image originale, 1999 - © Pierre Huyghe, courtesy Marian Goodman Gallery, Paris/New York.

fig. 8 - Pierre Huyghe, *One Million Kingdoms*, 2001. Vidéo, 6 minutes, couleurs, son.

Co-production Anna Sanders Films et Antéfilms. Photo de Laurent Lecat - © Pierre Huyghe, courtesy Marian Goodman Gallery, Paris/New York. <http://www.pbs.org/art21/slideshow/?artist=178>



Pierre Huyghe & Philippe Parreno, *No Ghost Just a Shell*, 1999-2003.

fig. 9 - M/M (Paris), Poster *ann lee: no ghost just a shell* (Pierre Huyghe & Philippe Parreno), 2002.

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

fig. 10 - Joe Scanlan, *Do It Yourself Dead On Arrival (AnnLee)*, 2002. Photo de Peter Cox - © Pierre Huyghe, courtesy Van Abbemuseum, Eindhoven. <http://www.vanabbe.nl/press/pressphotos/20050714-ann-lee.htm> 7

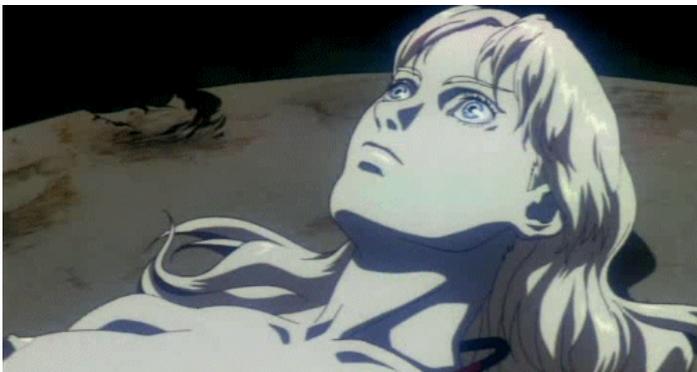
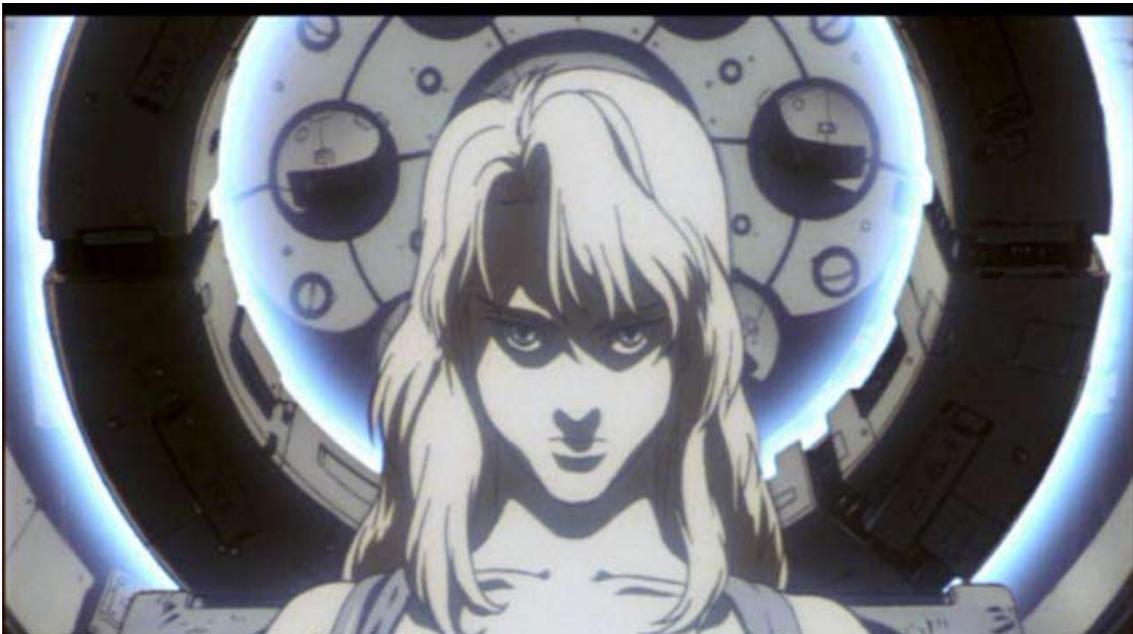
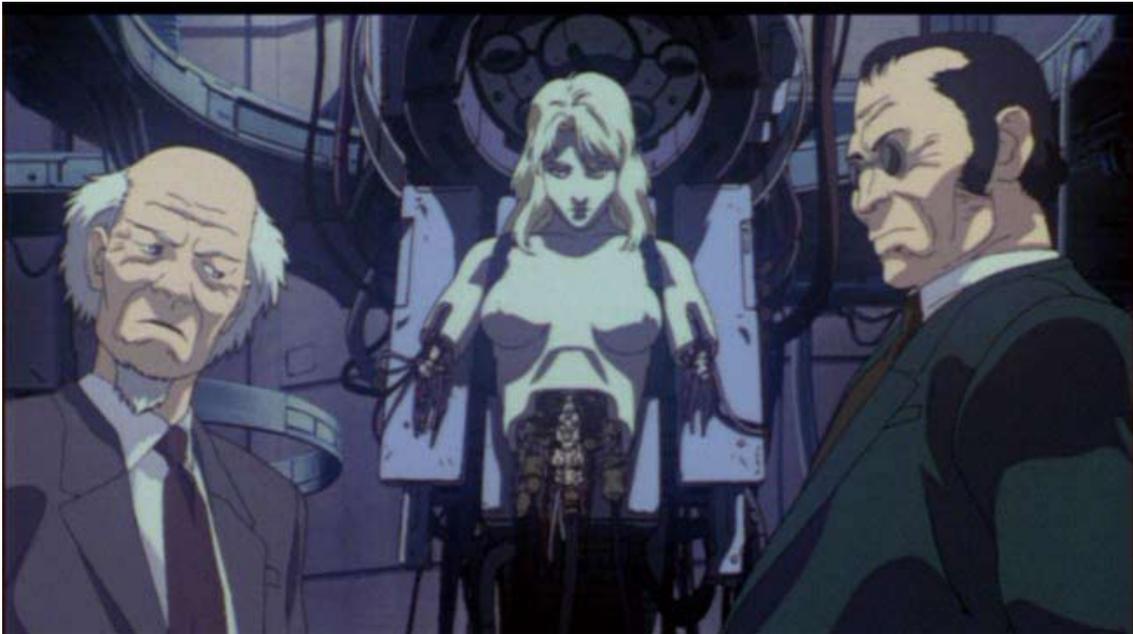


fig. 11, 12, 13, 14 - Le "Puppet Master", *Ghost in the Shell*, Mamoru Oshii, 1995.
© 1995 Shirow Masamune/KODANSHA - BANDAI VISUAL - MANGA ENTERTAINMENT
<http://www.cyberpunkreview.com/poll-results/favorite-anime-character-motoko-blowsem-away/>

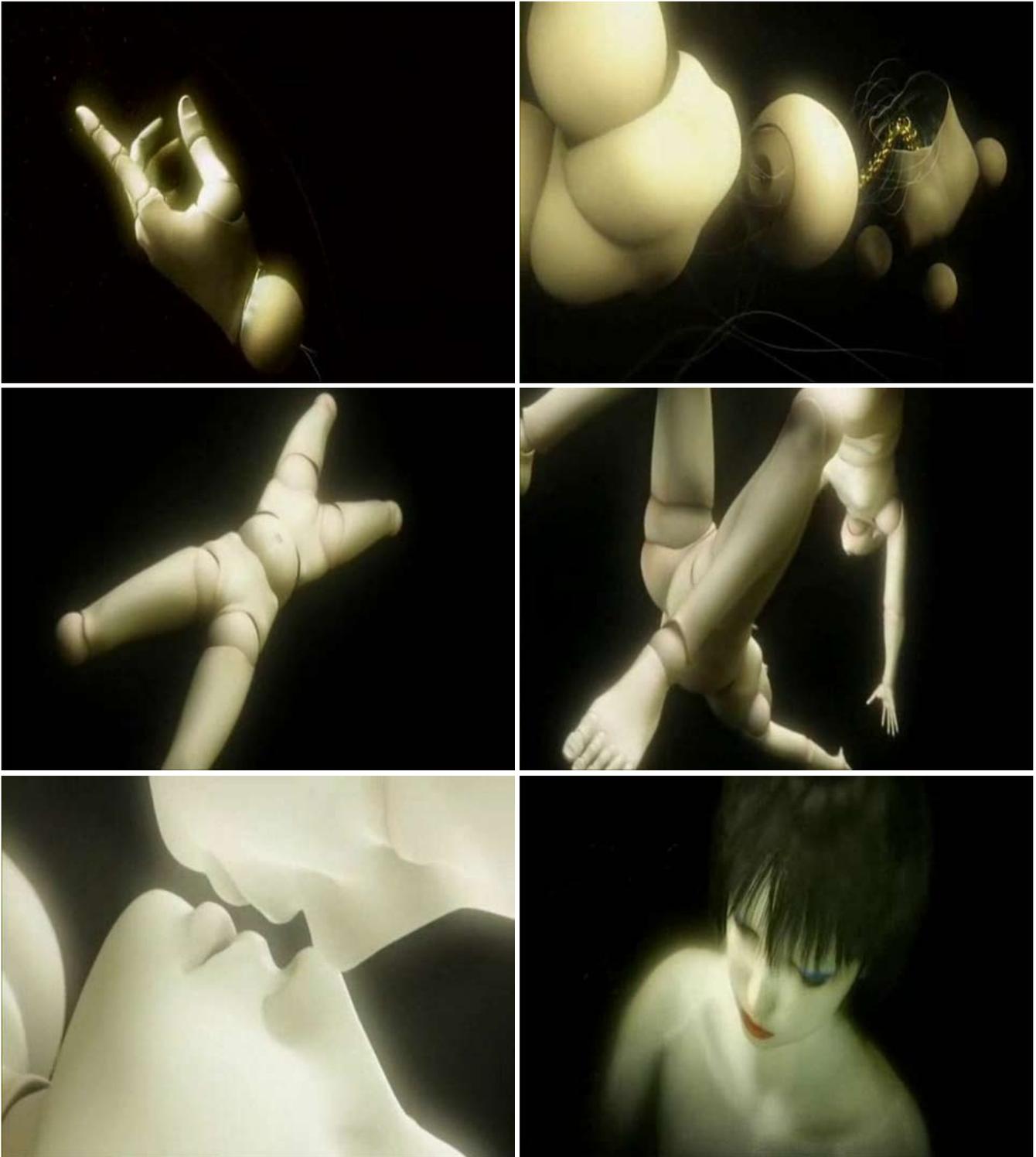


fig. 15, 16, 17, 18, 19, 20 - Générique d'ouverture de *Ghost in the Shell 2: Innocence*, Mamoru Oshii, 2004.
© 2004 Shirow Masamune/KODANSHA · IG, ITNDDTD.

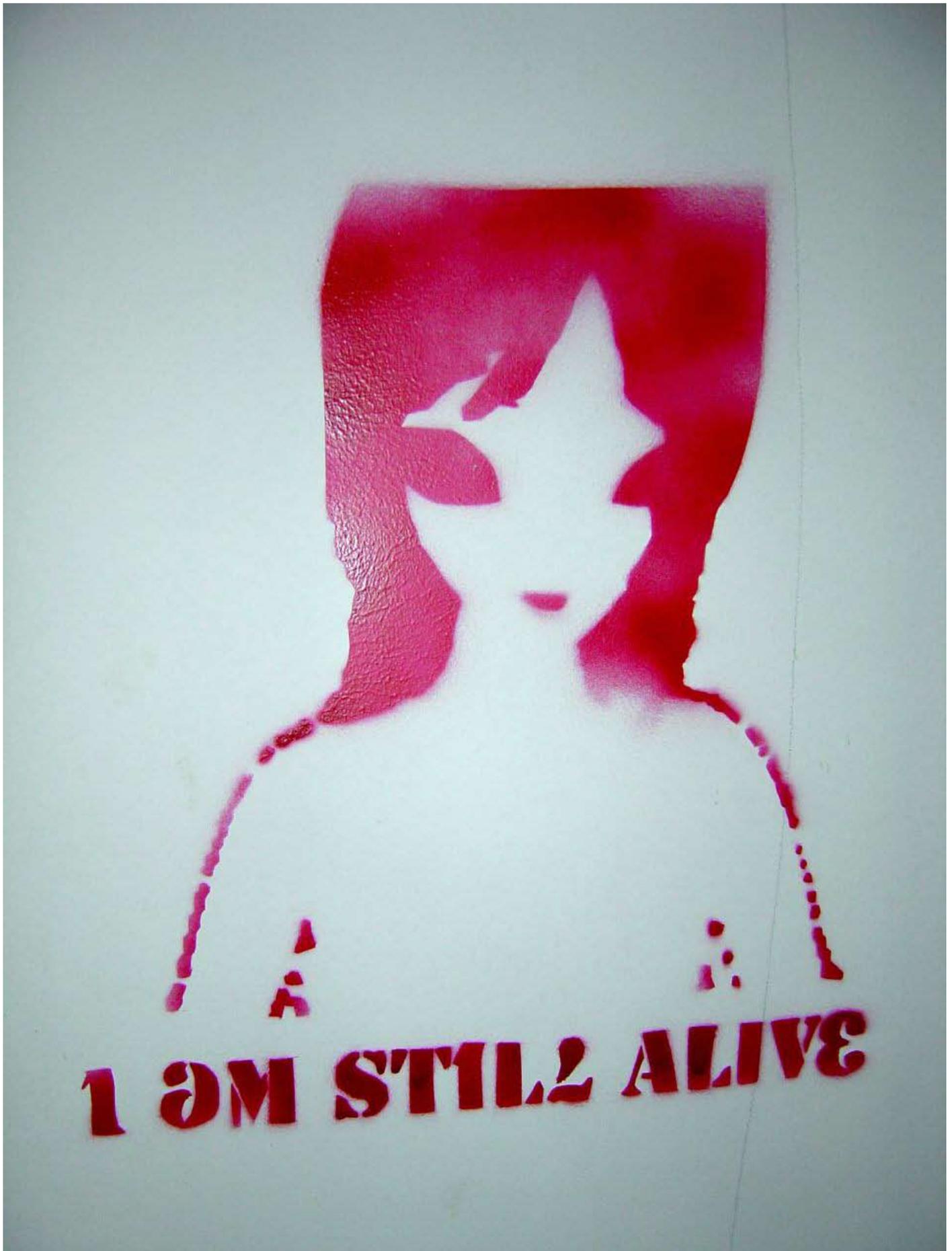


fig. 21 - Mircea Cantor, *I AM STILL ALIVE*, 2004. Photo par Libby Rosof - © Mircea Cantor.
<http://fallonandrosf.blogspot.com/2006/11/talk-with-video-artist-mircea-cantor.html>

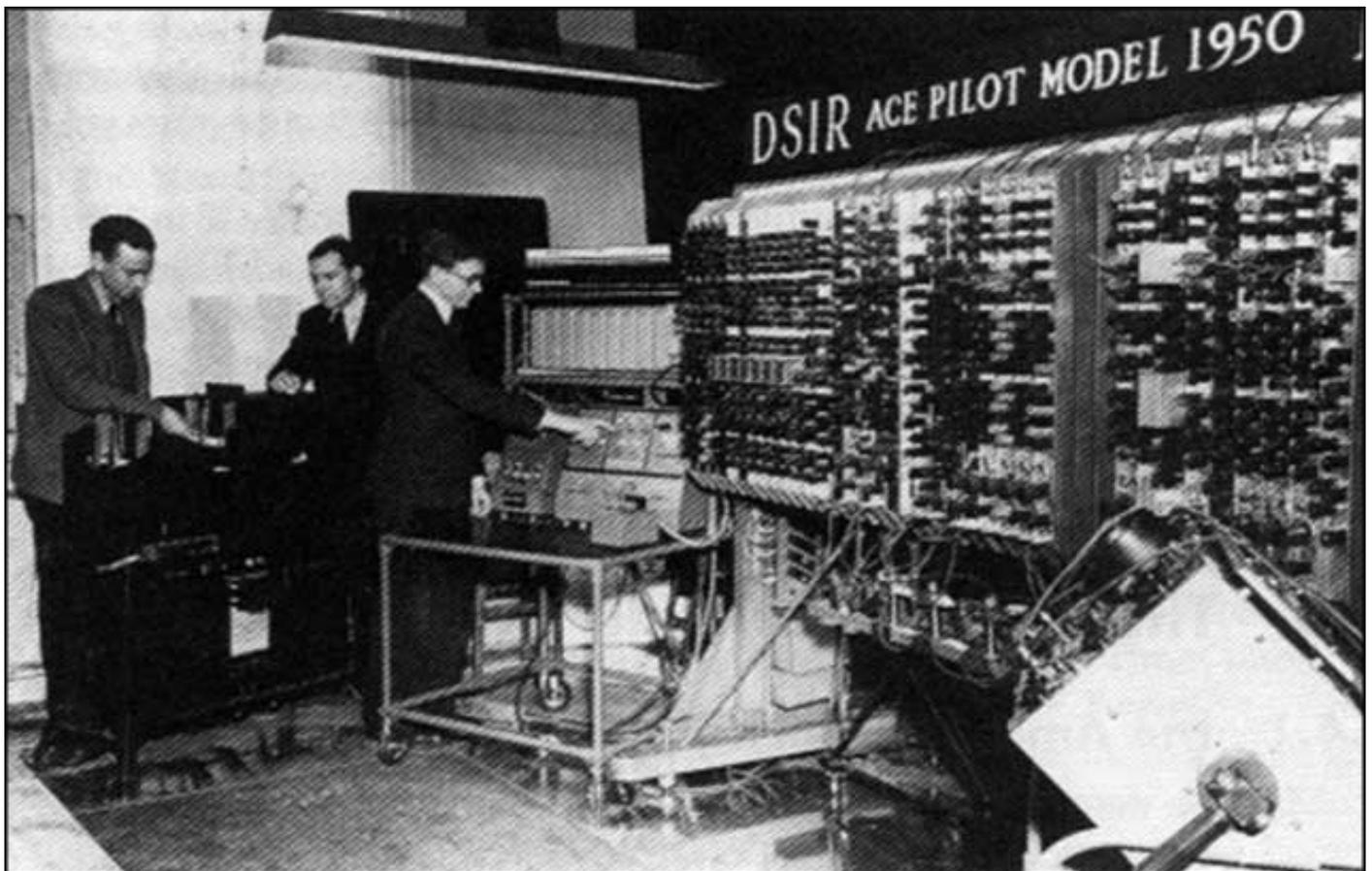


fig. 22 - Modèle pilote de l'*Automatic Computing Engine* de Turing, au National Physical Laboratory de Londres - NPL photographs © Crown Copyright 1950.

http://www.alanturing.net/turing_archive/archive/infopages/london1st.html

fig. 23 - Exemple de *datacenter*, avec ici un hébergement sur serveur dédié - © La Petite Boite.

<http://www.lapetiteboite.net/creation-site-internet-rambouillet/hebergement/>

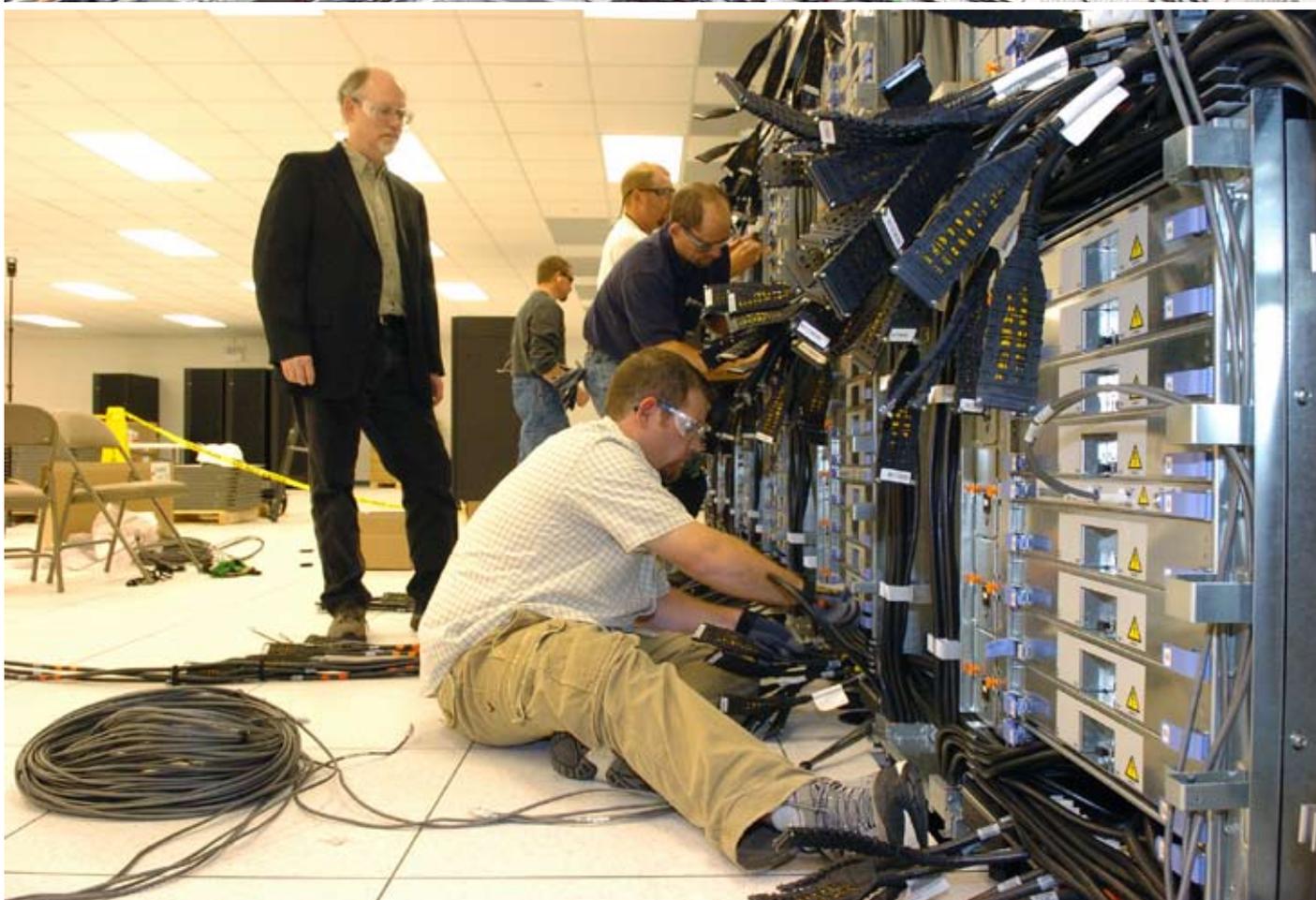
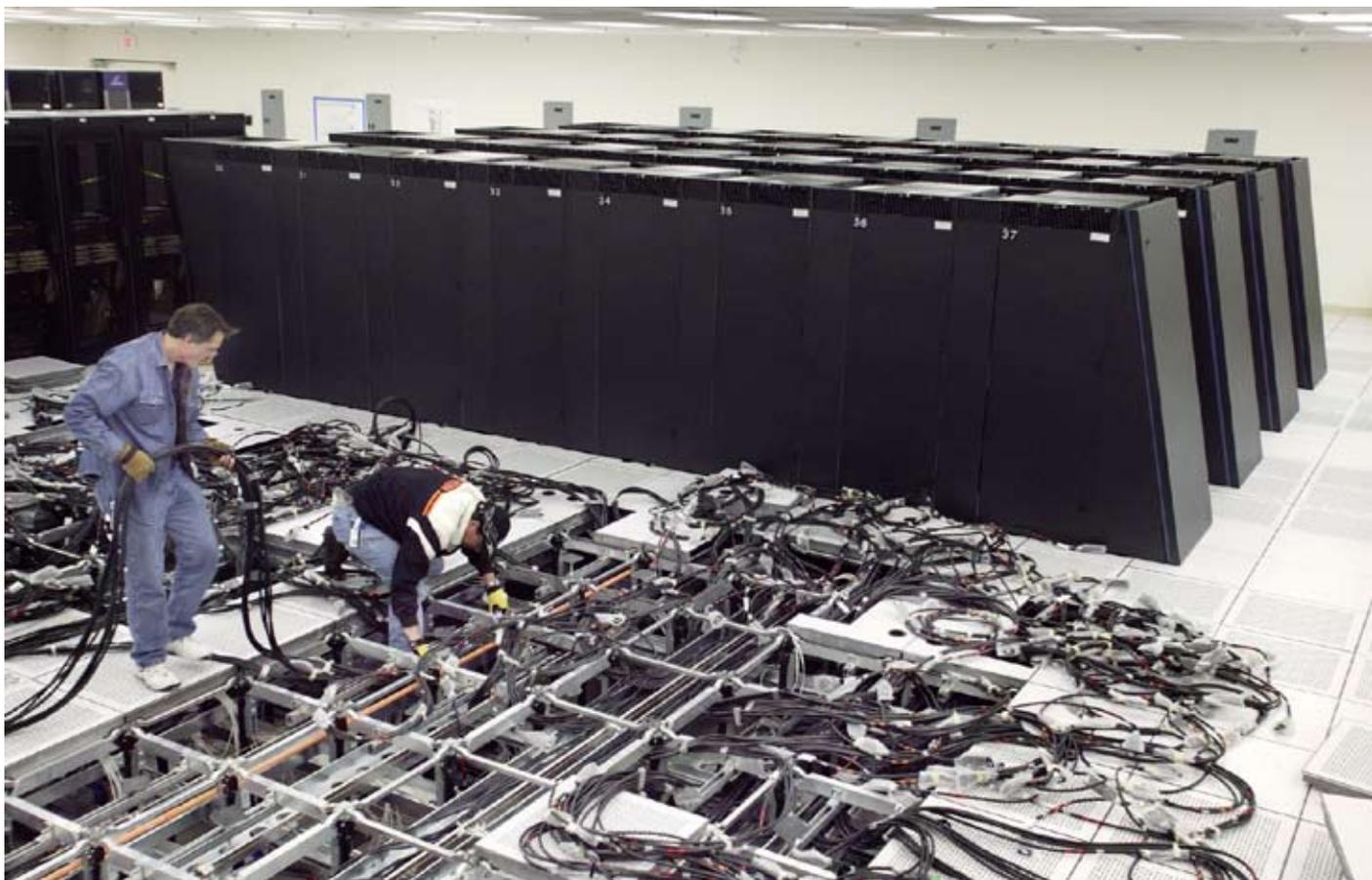
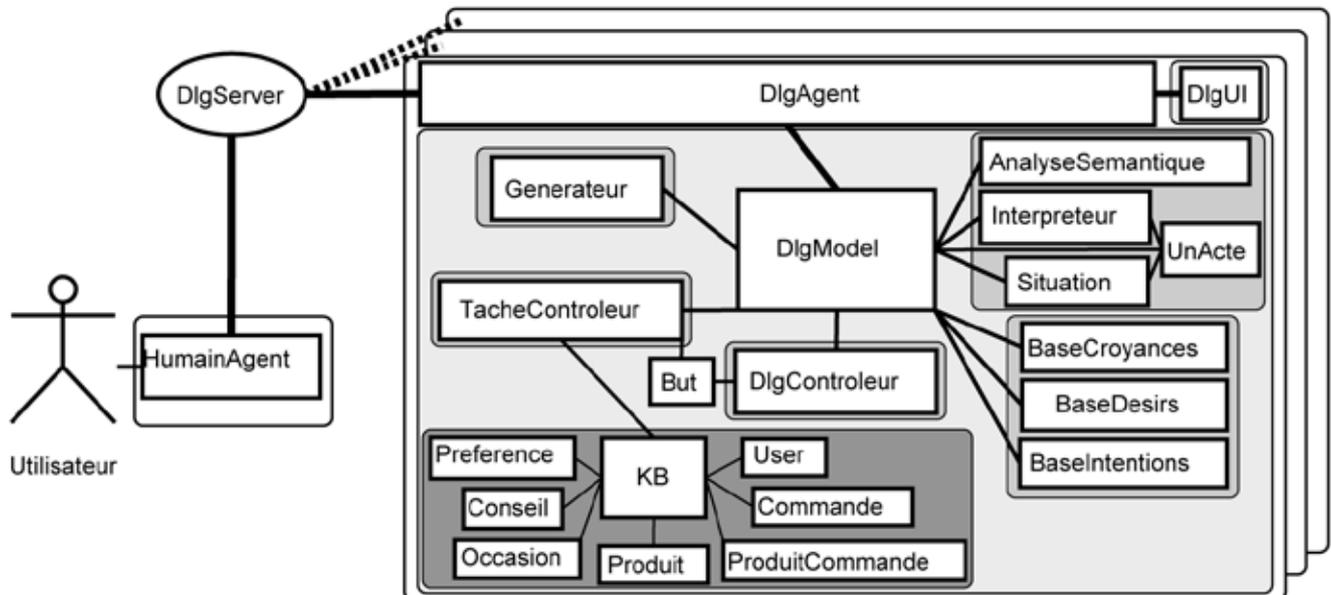


fig. 24 - Hébergement du supercalculateur IBM BlueGene/L - © IBM.

http://www.presence-pc.com/image/IBM-BlueGeneL_0101-14816-0-2-3-1-jpg.html

fig. 25 - Ray Bair, Directeur d'Argonne Leadership Computing Facility, supervisant l'installation d'un système IBM Blue Gene/P. Photo de George Joch - © Argonne, 2007.

http://www.anl.gov/Media_Center/News/2007/ALCF071109.html



- **DlgServer** : serveur 'relais' pour la mise en place des différents agents-clients et leurs communications.
- **HumainAgent** : agent-client d'interface utilisateur.
- **DlgAgent** : agent-client de l'ACE.
- **DlgUI** : interface du système de dialogue.
- **DlgModel** : noyau du modèle de dialogue.
- **AnalyseSemantique** : analyse sémantique de l'énoncé
- **Interpreteur** : interprétation à partir du schéma sémantique de l'énoncé
- **UnActe** : définition d'un acte de langage
- **BaseCroyances** : gestion d'une base de croyances
- **BaseDesirs** : gestion d'une base de désirs
- **BaseIntention** : gestion d'une base d'intentions
- **Situation** : gestion d'une situation selon l'acte interprété et les bases BDI
- **DlgControleur** : contrôle du dialogue à partir de la situation courante
- **TacheControleur** : contrôle de la tâche à partir de la situation courante
- **But** : gestion des buts
- **Generateur** : génération de l'énoncé réponse à partir du type de réponse et des éléments déterminés par le système
- **KB** : base de connaissances
- **User, Commande, etc.** : modules d'interaction avec la base de données associée

fig. 26 - Extrait de Yannick Fouquet, Alexandra Berger, Sylvie Pesty, "Une architecture d'Agent Conversationnel Expressif", in *WACA 2006 Second Workshop sur les Agents Conversationnels Animés*, Toulouse, 2006. http://halshs.archives-ouvertes.fr/docs/00/38/15/17/PDF/Waca06_Fouquet-al.pdf

Simple Search Routine.

```
void find_keyword()
{
    int len=0;
    int lenkey=0;
    int key_no=0;
    char teststr[50];
    while((ip.keyfound==0) &&(key_no!=MAX_KEY))
    {
        // getting the length of the keyword
        // lenkey=strlen(keys[key_no].getword());

        char *ptr=NULL;
        ptr=strstr(ip.userip,keys[key_no].getword());
        if (ptr!=NULL)
        {
            // keyword found !
            ip.keyfound=1;
            ip.keyno=key_no;
            strcpy(ip.keyword,keys[key_no].getword());
            break;
        }
        key_no++;
    }
}
```

```
//
// Program Name: chatterbot1
// Description: this is a very basic example of a chatterbot program
//
// Author: Gonzales Cenelia
//

#include <iostream>
#include <string>
#include <ctime>

int main()
{
    std::string Response[] = {
        "I HEARD YOU!",
        "SO, YOU ARE TALKING TO ME.",
        "CONTINUE, I'M LISTENING.",
        "VERY INTERESTING CONVERSATION.",
        "TELL ME MORE..."
    };

    srand((unsigned) time(NULL));

    std::string sInput = "";
    std::string sResponse = "";

    while(1) {
        std::cout << ">";
        std::getline(std::cin, sInput);
        int nSelection = rand() % 5;
        sResponse = Response[nSelection];
        if(sInput == "bye") {
            std::cout << "IT WAS NICE TALKING TO YOU USER, SEE YOU NEXT TIME!" << std::endl;
            break;
        }
        std::cout << sResponse << std::endl;
    }

    return 0;
}
```

fig. 27 - Extrait de code source, sous Turbo C IDE 3.0 (C++), Amit Mathur, "How to Build Eliza Chatterbot - A Program that can Chat with Humans", déc. 2002. Légende : "Finally we can work out a procedure for searching the keyword in the user's input. MAX_KEY indicates the number of keywords in the DAT file. Here we are simply searching whether the keyword exists in the user input (anywhere)."

<http://www.sourcecodesworld.com/articles/How-to-build-Eliza-Chatterbot.asp>

fig. 28 - Gonzales Cenelia, *chatterbot1* (this is a very basic example of a chatterbot program), 2009.

http://www.codeproject.com/KB/recipes/bot_tutorial.asp

SmarterChild

From Wikipedia, the free encyclopedia

SmarterChild was a chatterbot available on AOL Instant Messenger and Windows Live Messenger (previously MSN Messenger) networks.

History

SmarterChild became very popular in the early part of the 21st century, with continued user popularity until its demise in May 2009. The entity was the property of Colloquis (formerly ActiveBuddy), a company that develops IM based “Automated Service Agents”. Colloquis was acquired by Microsoft on October 12, 2006.[1]

During November 2006, SmarterChild’s MSN Messenger friends list featured more than 20 million names. SmarterChild remained extremely popular, and advertisements began appearing in its messages. However, during the latter half of 2007, SmarterChild no longer shared the features or intelligence it once boasted, such as the ability to check user’s horoscopes, inform users how many “friends” it was currently chatting to, and the ability to respond to conversation in general. Its most common responses were now “I dunno how 2 answer that” and “I don’t understand this fully enough to say anything smart about it...”

As of November 18, 2008, SmarterChild’s AIM agent was no longer offering any services. All messages began with the reply “My brain is retired but watch some cool videos! Send an IM to GossipinGabby and Type VIDEO!” SmarterChild changed his status to “ONLINE” in May 2009 and officially returned to Windows Live.

References

[1] “Microsoft Acquires Colloquis to Enhance User Interactions Online: Uses include customer self-service, enterprise information access, marketing and community outreach” (<http://www.microsoft.com/presspass/press/2006/Oct06/10-12ColloquisAcquisitionPR.msp>)

External links

* Biederman, Marcia (May 29, 2003). “At \$10 a Year, Automated Buddy Loses Laughs”. The New York Times (<http://www.nytimes.com/2003/05/29/technology/at-10-a-year-automated-buddy-loses-laughs.html>).

* Cha, Ariana Eunjung (September 6, 2002). “Web May Hold the Key to Achieving Artificial Intelligence”. The Washington Post (<http://www.washingtonpost.com/ac2/wp-dyn/A43363-2002Sep5>).

* Estabrooks, Mac (February 16, 2006). “See SmarterChild. See SmarterChild type. SmarterChild types fast. Type, SmarterChild, type.”. The Toronto Star (<http://pqasb.pqarchiver.com/thestar/access/987671831.html?dids=987671831:987671831&FMT=ABS&FMTS=ABS:FT&type=current&date=Feb+16%2C+2006&author=Mac+Estabrooks&pub=Toronto+Star&desc=See+SmarterChild.+See+SmarterChild+type.+SmarterChild+types+fast.+Type%2C+SmarterChild%2C+type.&pqatl=google>).

Retrieved from “<http://en.wikipedia.org/wiki/SmarterChild>”

Categories: Chatterbots | Instant messaging | Microsoft acquisitions

This page was last modified on 25 May 2011 at 08:24.

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. See Terms of Use for details.

Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.