

Les dimensions de l'intelligence artificielle

Michaël Klopfenstein

SOMMAIRE

Définir l'intelligence.....	4
Présentation.....	9
Univocité et flou.....	12
Percevoir le flou	14
Agir avec liberté	16
Interaction perception-action	18
Perception globale.....	21
Autonomie : action globale.....	26
Réflexion.....	31
Emancipation conceptuelle.....	40
Conscience	45
Conclusion.....	53

Définir l'intelligence

S'il est une chose sur laquelle les tentatives de définitions de l'intelligence s'entendent, c'est que la notion n'est pas facile à définir. L'intelligence est une notion vague et multiforme.

Il y a trois acteurs principaux : l'homme, l'animal et la machine. L'homme est le point de mire de la compréhension de l'intelligence. Les humains sont capables de déployer par la réflexion et les mécanismes mentaux automatiques ou conscients des capacités d'analyse conceptuelle inégalées. Si bien que de nombreuses tentatives de définitions le prennent comme référence sans expliquer quelle partie de l'homme concerne l'intelligence. On peut peut-être s'entendre sur le fait que l'intelligence est une faculté de traitement de l'information, quelque chose de complexe à l'intérieur du système qui transforme des informations en d'autres informations ou en des pratiques ou en une façon d'être, etc... A moins qu'on ne la voit comme une capacité générale de traitement de l'information.

Une approche consiste à mesurer l'intelligence par le résultat de ses activités intérieures : les productions qui découlent de son traitement de l'information. On pourrait mesurer l'intelligence de ces réalisations par leur variété, leur richesse, l'étonnement qu'elles génèrent, l'utilité des réalisations, etc. On peut aussi observer l'intelligence dans la capacité à atteindre des objectifs, à franchir des obstacles, à dominer une situation et à maîtriser les réalités du monde conformément à une intention.

La richesse des productions de la pensée humaine fixe une norme de l'intelligence. Lorsque la machine réussit à dépasser

l'homme dans certaines tâches qui lui demande beaucoup de réflexion, c'est là que se pose la question de l'intelligence artificielle. Et comme troisième pôle, il y a l'animal. La bête possède une faculté d'analyse et d'action sur le monde. L'homme ressent une rivalité face à l'animal. Il lui assigne la « bêtise » (ainsi qu'aux hommes qui ne sont pas dignes de son espèce) et se réserve l'intelligence par la prévalence des productions issues de sa raison. Pour la machine, les réalisations sont très variables. Elle dépasse l'homme sur certaines compétences et ne parvient pas à égaler les bêtes dans une foule d'autres tâches. L'autonomie et les capacités d'action que possède l'animal révèlent une capacité d'analyse de la perception du monde qui semble inaccessible à la machine. Et pourtant l'homme considère l'animal comme très pauvre en intelligence comparée à la réflexion humaine. On se retrouve dans une situation comparable au jeu de Shifumi : l'homme supplante l'animal qui supplante la machine qui supplante l'homme (sur des tâches complexe). Mais que compare-t-on ?

Si l'on se cantonne aux seuls hommes, il existe aussi des mesures de l'intelligence qui comparent les capacités d'analyses par des actions résultantes. L'intelligence se mesure par une sorte d'aisance analytique face à certaines activités mentales ou par la capacité à franchir des obstacles conceptuels. Mais peut-on poser la réalisation de certaines tâches au fondement de l'intelligence. Les différentes tâches ne sont pas uniformément répartie quant à l'aisance, la difficulté et l'inaccessibilité. Et toute capacité n'est pas liée à l'intelligence. Il faudrait alors établir la typologie des tâches intelligentes ? En quoi de tâches seraient-elles le sceau de l'intelligence ? Est-ce l'aveu de notre faiblesse à comprendre la compréhension, à définir le sens du sens ? Il semble

pourtant que l'intelligence ait un rapport avec la capacité d'analyse d'information, la capacité à percevoir et agir dans le monde réel, la capacité à accomplir une tâche mentale qui ne va pas de soi.

Et quand la machine s'est mise à exécuter des tâches non évidentes pour l'homme avec une grande aisance (la machine à calculer), on a ressenti la marque de l'intelligence. Et l'imagination a comblé les fossés du questionnement en assimilant la machine à l'intelligence ; il suffirait de combler quelques lacunes. Cette aisance a été contredite par une grande incapacité « des machines » à comprendre les choses les plus évidentes pour nous et les bêtes. Est-ce un problème de conception de la machine, de manque de compréhension de l'intelligence ou d'impossibilité radicale de capturer l'intelligence par la machine? On peut se rabattre sur le constat que l'intelligence ne forme pas une relation d'ordre linéaire et n'est pas quantifiable. On retrouve cela dans une moindre mesure chez les humains qu'on qualifie d'intelligents qui ont des compétences d'analyse diversifiées non comparables.

L'intelligence est-elle l'aisance dans certaines capacités ou bien la capacité d'avoir de l'aisance ? Dans cette deuxième direction, il y a l'homme et l'animal capables d'interaction et d'autonomie face au monde ? Une capacité de perception, d'analyse et de réaction qui convient à leur choix/mode de vie. A l'inverse la machine douée d'une aisance redoutable dans certaines capacités est souvent aveugle et morte à la compréhension du monde et de ses « évidences ». C'est un poncif que de déclarer la machine « stupide ». Il est probable que l'évidence est une notion relative à un système, et qu'en réalité l'accès à l'évidence du monde est d'une complexité

redoutable. Mais elle n'est pas ressentie comme tel par l'homme, et l'on distingue deux directions à l'intelligence :

- l'intelligence passive dont on n'est pas conscient
- l'intelligence active qui demande des efforts.

Curieux paradoxe que l'intelligence passive qu'on ne maîtrise pas qu'on ne connaît pas et qui nous révèle le monde avec une richesse et une facilité déconcertante. Elle réside dans l'obscurité et fait côtoyer le sentiment d'évidence avec la potentialité d'une inaccessible complexité. Cette capacité d'entrer dans le monde, de le comprendre et de pouvoir agir sur lui selon des intentions préétablies devient un Graal de l'intelligence artificielle. La machine « stupide » est celle qui ne fait que copier une intelligence active par le suivi mécanique de règles édictées par l'homme.

La recherche de l'intelligence s'est dédoublée. Ceux qui préfèrent les sommets visibles à l'obscurité se sont attachés à l'intelligence comme l'accomplissement de tâches qui demandent la plus grande « intelligence » (pour l'homme). Ils parviennent de victoires en victoires à montrer le pouvoir démultiplicateur de l'intelligence informatisée sur l'intelligence pure. En simplifiant, ce sont les situations où les hommes qui ont les plus grandes capacités d'analyses investissent toute leur force (les jeux de stratégie notamment). Dans cette optique, on reste fixé sur des tâches et pas fondamentalement sur la recherche d'un cadre qui rend ces tâches possibles de façon autonome. Mais l'espoir anime certains dans l'idée que les concepts mis en œuvre sont des éléments de l'intelligence.

De l'autre côté, se développe une autre approche moins visible. La recherche d'une intelligence « universelle » capable d'accéder au monde, en perçant l'obscurité de la nature de

l'intelligence. La stupidité ne peut se concilier avec l'intelligence, quand on ne sait pas percevoir la moindre évidence, il est inutile d'atteindre les sommets (« par une intelligence simulée où des hommes conçoivent le programme et sont à l'origine de tous les concepts analytiques utilisés par la machine selon une linéarité d'exécution abrutissante »). Mais existe-t-il vraiment une intelligence universelle capable d'aborder toutes sortes de tâches, même seulement les plus évidentes ? Car c'est le mystère de l'intelligence passive qui sert de référence pour évaluer la « bêtise » ou l'intelligence d'une machine. L'espoir est fondé sur une croyance matérialiste mécanique de la réalité. Mais, même sous cette hypothèse, l'intelligence est-elle d'une complexité accessible ? Face à tout mystère, c'est une croyance systémique et subjective qui sert de support au raisonnement et de motivation. Pour beaucoup, le doute subsiste sur la nature de l'accès à l'intelligence globale par la machine. D'un autre côté, dans la culture populaire subjuguée par l'émotion de cette éventualité et confrontée sans cesse aux scénarii de la machine intelligente, la question a presque mué en perspective normale. Il faut un minimum de connaissances et de lucidité pour comprendre qu'au milieu de l'immense technicité et des impressionnantes réalisations, il y a encore peu, si peu d'intelligence artificielle, mais beaucoup d'espoir face à une montagne haute et obscure.

Maintenant une tâche n'est pas forcément facile à détacher du fonctionnement qui l'accomplit et du cadre qui la permet. Qu'est-ce qui dans le support de la réalisation d'une tâche est indépendant ou fait partie de l'intelligence. L'émotion, la motivation, la morale, la personnalité, la conscience, ... qui animent l'homme sont-elles des composantes de l'intelligence ou un simple mode de fonctionnement humain

non indispensable à sa réalisation ? L'homme est riche de spécificités anthropologiques. Jusqu'où ces particularités sont elles au fondement de l'intelligence ? Lesquelles sont nécessaires à son existence ou simplement à sa définition ?

Estimer les possibilités équivalentes d'une réalité unique est un questionnement paradoxal, mais pourtant fréquent quand il s'agit de positionner l'homme et le monde. Toutes les spéculations sont possibles. Sur quelle base retenir ou rejeter un des critères présents ? Voilà qui concourt à l'obscurcissement de la notion d'intelligence quand on prend l'homme comme toise : qu'est-ce qui chez l'homme est intelligence et ne l'est pas ? L'intelligence est une notion non structurée qui ramène sans cesse sur la scène toutes les caractéristiques susceptibles d'avoir les rapports les plus lointains avec l'intelligence.

L'intelligence serait-elle « ce dont l'homme est capable à partir de sa réflexion » ? Poser l'homme comme repère est délicat. Ce n'est pas convaincant quand on voit les maladresses, les difficultés et les erreurs dont il est aussi capable... Quant à faire la distinction entre ses actes pour ne retenir que ce qui est intelligent, n'est-ce pas justement l'objet de la définition ?

Présentation

C'est dans la direction du mystère de l'intelligence que nous allons nous plonger, à la recherche des concepts qui permettent cette capacité de comprendre et d'agir sur le monde en cherchant à s'affranchir de la limitation d'un cadre de reconnaissance préformaté. Découvrir les mécanismes de

la capacité d'intérioriser, de comprendre, et d'agir sur ce qui nous entoure. Et pourquoi pas nous intéresser aussi à cette capacité tellement humaine d'aller plus loin jusqu'à créer de nouveaux concepts de lecture du monde, l'inventivité. Pourrait-on même aller jusqu'à créer une perception réflexive dans un système ?

En s'attelant à la référence humaine, il sera difficile d'atteindre les concepts supposés neutres et fondateurs de l'intelligence. Une foule de concepts associés viennent s'imposer par la culture et l'attraction que la condition humaine produit sur le concept : l'adaptation, l'apprentissage, les connexions neuronales, l'imagination, la créativité, la socialisation, la génétique, et tant d'autres idées évocatrices... Subjugué par l'intelligence humaine, et face au constat d'une parenté incontestable, nul ne peut s'opposer à ce que chacun de ces cadres soit nécessaire pour accéder à « l'intelligence du monde ». Mais connaître leur nécessité est un autre problème. Pour équilibrer cette attraction du modèle unique, il me semble vital de développer une dynamique de prise de distance. Après chaque observation d'un mécanisme de l'intelligence humaine, il faudra s'éloigner et tenter une généralisation qui pourrait prendre d'autres directions. Se dégager de ce qui ne semble pas nécessaire, de ce qui pourrait être une spécificité humaine non indispensable. Car plus que humain ou pas humain, il s'agit d'établir une description des phénomènes de l'intelligence et de leur reproductibilité artificielle par un accès structurel.

Maintenant, il faut se rendre à l'évidence, tant que la démonstration n'est pas faite qu'un groupe de concepts est suffisant pour atteindre une forme d'intelligence donnée, le choix du cadre est subjectif, il repose sur des présupposés théoriques. Il ne s'agit donc pas d'affirmer mais de proposer.

Mon objectif consiste à découper l'intelligence en fonctions distinctes et en dépendance entre ces fonctions, pour permettre une conceptualisation structurée de l'intelligence. Il ne s'agit pas d'un découpage théorique qui prétend révéler l'essence de l'intelligence, mais d'une tentative d'organisation pour distinguer des principes permettant de réfléchir à des structures sans affronter l'ensemble dans une confusion insoluble. C'est la recherche d'une mise à plat de l'intelligence pour démêler cet imbroglio conceptuel.

L'objectif est centré sur le deuxième axe : la recherche d'une représentation opérationnelle du monde qui entoure le système. La description est conçue comme une progression de concepts qui s'empile pour atteindre spéculativement les éléments nécessaires à « l'intelligence humaine ». Finalement, on obtient plus encore qu'une organisation avec l'idée d'une gradation conceptuelle dans la constitution de l'intelligence. Cela permet de rendre compte d'une difficulté de la définition de la notion. Car l'intelligence ne se trouve pas au terme de l'accumulation, mais à chaque étape où s'ajoute une dimension. Des définitions de l'intelligence comme « Autonomie à produire un traitement abstrait des informations nouvelles pour en tirer une efficacité opérationnelle » ou « capacité d'un système à percevoir un sens organisé face à une réalité globale et à réagir à l'issue d'une analyse des perceptions en vu d'un objectif préformé par le système » sont situées sur des paliers différents de la gradation que nous allons parcourir.

Univocité et flou

La première rupture dans le concept d'intelligence concerne son lieu d'exercice, celui de l'univocité contre celui de l'approximation. J'appelle lieux univoques, ces règles qui manipulent des informations sans ambiguïté par des structures mathématiques bien identifiées. On pourrait penser que tout système informatique se range dans cette catégorie puisque le raisonnement y est fait selon des processus binaires. Mais ce n'est pas nécessairement le cas dans la perspective qui nous intéresse. Lorsqu'un programme utilise des processus non linéaires complexes dont on ne peut prédire le comportement, le mode de raisonnement sur l'aboutissement du programme n'est pas celui d'une structure maîtrisée. Bien que pas à pas tout soit prévisible par un processus binaire déterministe, le rapport entre l'objectif et le principe qui dirige ne relève pas d'une maîtrise univoque. Ainsi il est des situations où l'on ne sait faire aboutir un programme que par des principes qu'on tente de rendre convergent et non selon un déterminisme rationnellement anticipé. Même si en pratique c'est le cas au niveau élémentaire, je ne parlerai pas d'univocité. On n'en parlera encore moins si on utilise des processus stochastiques dans déroulement du programme. Une dernière situation qui nous concerne encore plus, est celle où les informations à utiliser pour conduire le processus sont définies avec une imprécision native. Quand des capteurs imprécis cherchent à « lire le monde », vouloir dépasser l'imprécision relève d'un principe non univoque, et plus encore quand le système doit non seulement constituer, mais aussi manipuler des identités qui ne sont pas objectives.

Cette première rupture forme la frontière suivante : d'un côté, il y a les systèmes dont la manipulation d'informations repose sur les concepts installés dans le système (c'est particulièrement le cas quand le système utilise le monde réduit et « simple » d'un modèle mathématique depuis l'entrée jusqu'à la solution). De l'autre côté, il y a les systèmes dont la dynamique essentielle provient d'une auto-organisation à partir de données extérieures qui ne peut donc pas être prévue si les données sont non maîtrisables. De nombreuses nuances sont possibles sur cette frontière. Comme le monde réel ne peut pas être modélisé par un système univoque global, l'accès à la manipulation du monde extérieur nécessite de se limiter à un secteur où l'identification est stable ou alors manipuler une structure de reconnaissance qui gère l'imprécision, les défauts, les structures floues, la subjectivité ...

Loin de l'environnement protégé des laboratoires, dans la pleine lumière du monde réel, l'accès à la plupart des identités par des capteurs sensoriels est d'une complexité redoutable, elle nécessite la gestion des défauts d'univocité ; et cela d'autant plus que l'accès se veut conceptuellement large.

D'un côté, on trouve les « algorithmes » qui gèrent des identités bien définies par une analyse experte reposant sur la manipulation de concepts issus de l'intelligence humaine (comme dans les jeux complexes). De l'autre côté, on a une machine qui, par ses mécanismes, doit constituer l'identité des objets qu'elle doit aussi analyser... Bien sûr, il existe une foule de problèmes à cheval sur ces deux aspects. On peut, par exemple, construire des identités à partir d'une perception floue pour les manipuler ensuite avec une complète univocité. Ou inversement, partir de concepts

purement univoques, pour construire une dynamique qui cadre un processus complexe non maîtrisable univoquement. Cela n'empêche pas de poser ces deux zones comme une polarité primordiale : la maîtrise univoque et l'organisation de l'indéterminé ; deux paradigmes du traitement de l'information.

Notre objectif, centré sur une intelligence qui perçoit le monde de façon conceptuellement vaste et qui interagit avec lui à partir de l'analyse de ses perceptions, peut être clairement identifié au deuxième groupe qui devra gérer des incertitudes et des déterminations non univoques du réel.

Percevoir le flou

Pour commencer, on peut faire une trichotomie des processus intelligents :

- la perception : ce qui entre, depuis l'information sensible jusqu'à la formation de sa représentation globale.
- l'action : ce qui sort, depuis son origine jusqu'aux activations des mécanismes de sortie.
- la réflexion : la manipulation d'informations internes qui ne concerne pas un parcours d'entrée ou de sortie.

La première étape de l'intelligence consiste en un système qui gère la perception, uniquement. Il s'agit de transformer l'information sensible en représentation structurée qui permet d'identifier. La recherche d'identification des perceptions sensorielles occupe une grande part de la

recherche en intelligence artificielle aujourd'hui. Pouvoir reconnaître des objets, pouvoir nommer des reconnaissances, pouvoir analyser des images, de vidéo, des bandes sonores, des textes, ...

De nombreuses structures sont envisageables. Comme le flou n'est pas réductible à l'univoque, il faut construire une structure de l'identification, ce qui peut être assimilé à une première forme d'intelligence. On peut distinguer les méthodes qui consistent à tracer des zones d'identité dans un espace préétabli de tous les possibles et les méthodes qui consistent à construire une structure à partir de l'information reçue qui complexifie l'organisation au rythme des expériences. Indépendamment de la méthode, produire un mécanisme automatique de re-connaissance des éléments perceptifs est une étape primordiale de l'intelligence. Une fois les identités reconnues, on peut greffer un traitement des informations à partir de ces identités.

Quelle est la méthode d'obtention et le mode de représentation de ces identifications ? On peut avoir un gradient de structure très varié : à un extrême, on peut réaliser les identifications à partir d'une base de données. Ce qui ne suppose pas la construction d'une connaissance à partir de la perception. A l'autre extrême, il s'agirait de tirer de chaque perception tous les liens de connaissance à tous les niveaux conceptuels possibles, si tant est que cela ait un sens. Entre un processus de positionnement parmi des zones prédéfinies sous un concept unique et la construction d'un tissu conceptuel étendu par l'expérience, le mot « identification » n'a pas le même sens. Le premier consiste à mettre une étiquette. Et s'il n'y a pas de connaissance conceptuelle, l'information sur l'identité reste à distance de l'objet. Pour atteindre un tissage conceptuel complexe, le

second processus doit mettre en jeu des structures complexes de représentation de la connaissance, il faut décortiquer la nature des objets et identifier chaque ressemblance conceptuelle. L'intelligence sera d'autant plus grande que les concepts captés et intériorisés seront structurés avec richesse et pertinence par rapport au réel.

Agir avec liberté

La seconde étape est l'action. Si on raisonne en termes d'information, l'action consiste à produire des valeurs de sortie qui seront sources de transformation du monde extérieur au système. Il y a peu de profit à construire une perception complexe qui n'aurait aucune recherche d'action, ce serait une intelligence théorique et sans preuve. La construction d'un processus d'action peut exister à de nombreux degrés d'intelligence et d'organisation.

Dans cette partie aussi, on pourrait commencer par exposer la rupture entre la gestion univoque maîtrisée et la gestion non maîtrisée du complexe et du flou. Mais la question d'un cadre univoque est dissymétrique entre la perception et l'action : il est assez naturel qu'on s'adapte à une perception non univoque du monde, mais rien n'interdit une structure maîtrisée pour l'action par l'accomplissement de règle strictement univoque dès lors qu'on travaille à partir d'identités stables. Cependant, c'est bien la question de l'intelligence qui est en jeu. Etablir une action sur la base d'une lecture globale univoque, cela signifie l'utilisation de concepts préformés par un programmeur. Pour atteindre une forme d'intelligence propre, il est souhaitable d'envisager que

l'actions ne procèdent pas d'un ensemble de catégories préconstruites suffisantes. Il s'agit de produire une action dont l'intelligence provient du système, qui est donc, elle aussi, non maîtrisée au départ. Un tel fonctionnement est assez différent du cadre de l'intelligence informatique classique d'une machine ou d'un robot qui accomplit « Ce qu'on lui a dit de faire ». Il serait intéressant d'ouvrir un espace d'indétermination dans les directives imposée. Pour cela il faut probablement concevoir aussi un système de maintien du déroulement dans un cadre dont les limites seront plus ou moins larges.

C'est l'intelligence humaine qui nous inspire cette idée d'action non prédéterminée. Est-ce un mode nécessaire pour le déploiement d'une intelligence qui se rapproche de notre objectif ? C'est une forme de liberté pour que se déploie une logique propre au système. Sous l'hypothèse de faisabilité, ce que pourrait faire une machine sous un mode d'action non imposé éveille spontanément des inquiétudes : en imaginant des capacités et une personnalité de type humain. Mais l'autonomie de l'action peut produire une foule d'actions nettement plus élémentaires et indépendantes de la psychologie humaine.

L'émancipation vis-à-vis de directives strictes est à la base d'une forme faible d'intelligence qui résulte d'un processus possédant une part d'indétermination ; encore une fois pas nécessairement au niveau de la lecture informatique binaire, mais au niveau de la prévisibilité structurelle.

L'émancipation informatique est-elle possible ? Il y a de nombreux degrés d'émancipation. Les mathématiques savent produire un chaos déterministe dont la complexité cadrée peut ressembler à du hasard ou de l'indéterminé. Reste à

savoir si soumettre l'action d'un système à ce type d'indétermination peut concourir à produire une forme d'intelligence. En supposant que ce type de processus soit conjugué à lecture critique de confirmation des choix, sous l'analyse d'une pertinence quelconque, on peut supposer qu'un tel processus en deux temps est une piste potentielle. N'est-ce pas là un mode possible de l'imagination, voire de la créativité dans le sens d'imprévisibilité, complexité et harmonie ? La créativité n'est pas hors de portée, si on la définit comme non-prévisibilité, possibilité de surprise, richesse, complexité et résonance des réalisations. Alors que la créativité comme conception d'un nouveau cadre semble beaucoup plus difficile d'accès, nous y reviendrons.

A ce stade le non-contrôle reste cadré, il s'agit d'une liberté sous réserve critique. Il y aurait peu de sens à libérer l'action dans un régime chaotique sans la soumettre à une dynamique structurée qui en garantisse la marche stable. Il ne s'agit là que de notions vagues, mais elles fixent à grands traits certaines manifestations de l'intelligence, à savoir ici, une certaine mesure de liberté dans l'action.

Interaction perception-action

L'action humaine et animale possède une caractéristique technique intéressante : l'action fonctionne en liaison étroite avec la perception. Le déroulement d'une tâche musculaire est guidé par la perception tout au long de son déroulement. D'un autre côté, la perception reçoit les informations du déroulement de l'action et on peut supposer qu'une évaluation du succès et des échecs est prise en compte et

permet de construire une action efficace sur le monde. Cette interaction est un mode spécifique de l'intelligence, un palier indispensable pour s'approcher d'une interaction pertinente entre soi et le monde.

Si l'on veut pénétrer cette dimension de l'intelligence, il faut concevoir une organisation qui lie les représentations de la perception et de l'action dans une structure commune. Je vois plusieurs directions à cette liaison commune.

- Un mécanisme de rétrocontrôle du déroulement de l'intention par la perception. L'action est dirigée par un objectif. A chaque instant l'action déroule un processus qui cherche à atteindre cet objectif. Toutes les manifestations de ce processus qui sont perceptibles par les capteurs sensoriels vont permettre de vérifier l'avancement et la pertinence de la direction prise. Les mesures relevées, par comparaison avec un processus d'anticipation perceptive, vont conduire à un système d'ajustement de la poursuite du processus d'action. Ce mécanisme consiste à utiliser toute l'information disponible pour atteindre au mieux le but. Se priver de cette information temporelle pour optimiser l'action serait un choix « stupide ». Cette approche révèle des mécanismes d'action qui fonctionnent sur des processus d'ajustement non univoque vers l'identification d'une structure optimale, comme pour la perception.
- La mémorisation commune des choix et de la perception issue de ces choix produit une connaissance des capacités d'action qui va tisser une « connaissance du monde intime ». Chez l'homme ou l'animal se construisent toute la connaissance et la maîtrise du corps, la compréhension de l'interaction avec la matière, la connaissance des possibilités de la matière en fonction de son apparence,

son toucher, etc... Cette connaissance « expérimentale » est une connaissance importante du monde. Elle n'est souvent pas consciente chez l'homme. Elle permettra l'intelligence du contact entre le système et le monde.

- Sur la base des deux mécanismes précédents, on peut ajouter l'idée que l'objectif qui guide l'action possède un grand rapport avec la perception. On peut estimer que l'intention, comme tout le déroulement de l'action, sont liés à un processus d'anticipation perceptive. Ce qui implique que la nature même du processus de production de l'action est intimement liée à la perception : l'intention a une composante perceptive.

Les processus entrant et sortant qu'on pouvait imaginer au départ comme relevant de logiques distinctes doivent relever d'un processus unique qui les englobe. En fait, la perception peut éventuellement conserver une certaine autonomie, mais l'action doit prendre une forme qui s'imbrique dans la perception pour parvenir à une efficacité, marque d'intelligence.

Sur la base de l'observation de l'homme, comment pourrait-on envisager un peu d'intelligence d'action sans cette coordination, sans appuyer l'action sur les anticipations et les ressentis intérieurs que produit l'action. Bien sûr, pour rendre crédible ce discours, il faut établir la faisabilité et exhiber des structures de représentations de la connaissance qui permettent de définir un tel fonctionnement.

Perception globale

Une perception intelligente doit accéder à tous les concepts possibles, il faut donc concevoir une structure d'accès à l'information et de manipulation qui ne possède pas de catégorie prédéfinie. Une étape supplémentaire réside dans la capacité de la perception à tisser de liens de toutes natures afin de couvrir un panel sémantique qui soit le plus large possible. Pour cela, il faut envisager une neutralité conceptuelle. D'ordinaire en informatique, quand on traite d'un champ sémantique fixé, on établit une structure de manipulation adaptée aux données qu'il faut manipuler, potentiellement complexes. Pour couvrir un champ qui soit le plus large possible, il faut une structure sémantique qui soit adapté à toute forme de connaissance, un système général de représentation de la connaissance qui soit capable de supporter le plus grand nombre de types de relation. Mais la structure de la perception ne doit pas se limiter à supporter une information large, le système devra aussi parvenir à extirper le maximum d'informations des capteurs sensoriels dont il dispose. Cela suppose que le mécanisme d'organisation est lui-aussi adapté à ce processus de construction conceptuelle.

Si on limite l'information à un type réduit de relation au détriment d'autres, les capacités d'analyse se trouvent privées de la possibilité de sortir du cadre des formes de connaissances qui lui sont imposées au départ. On trouve ici l'origine du reproche principal envers les machines, le sentiment de « stupidité » par l'incapacité à sortir d'un mode restreint de représentation. Une perception globale consisterait à tisser des informations dans toutes les directions indépendamment d'une catégorie particulière,

c'est ce que je nomme le principe de neutralité du cadre de l'intelligence. Mais s'il est facile de prétendre à une ouverture conceptuelle, encore faut-il pouvoir la réaliser et l'intégrer à une dynamique d'acquisition. Il reste donc à savoir si et comment c'est possible.

Des questions se posent sur la neutralité. Par exemple, la connaissance que nous avons de la vision chez l'homme nous fait voir que les premiers étages de traitement de l'information sensorielle n'ont rien de « neutre ». Au contraire, ils consistent à tirer une information tout à fait spécifique comme la reconnaissance de traits verticaux, horizontaux, obliques, la reconnaissance de petite tâche, de zones plus étendues, etc... une sorte d'alphabet de la vision qui permet une efficacité énorme par rapport à un traitement unique. Remarquons que ces prétraitements de l'information qui imposent une sémantique initiale sont probablement aussi la source des phénomènes d'illusions d'optique. De façon plus théorique, tout choix de structure est un engagement sémantique. Il y a une incontournable « philosophie » dans toute sémantique initiale qui éloigne de la neutralité. Ainsi, il ne faut pas voir ce concept de neutralité comme un fondement global, mais comme une direction pertinente. Pour gagner en efficacité, on comprend l'intérêt d'un compromis entre le prétraitement sémantique et la neutralité ultérieure.

On peut dresser un constat d'opposition : le cerveau humain montre une foule de traitements spécifiques dans des zones cérébrales identifiables qui ne donnent pas forcément l'idée de neutralité sémantique. De l'autre côté, vu la richesse de la production humaine, on peut entrevoir le besoin d'une certaine neutralité sémantique sans laquelle il est difficile d'envisager les tâches variées de la pensée humaine, même

avec 80 milliards de neurones. Tout simplement parce que la combinatoire est un processus à croissance exponentielle qui dépasse très vite des chiffres aussi grands. Si toute activité conceptuelle devait posséder une place définie à l'avance, l'ensemble des possibilités de tâches mentales vues comme des processus combinatoire sur les plus petites bases possibles produit une explosion de possibilités dont ne peut pas rendre compte un « si petit nombre de neurones ». Pour pouvoir tisser du sens en tout sens, il semble nécessaire de disposer d'un processus possédant une certaine neutralité sémantique. A l'image du cerveau humain et certainement animal, on peut (et peut-être doit-on) aussi envisager une foule de traitements sémantiques spécifiques pour imposer des dynamiques spécifiques et un mode d'accès efficace à la perception ouvrant un accès rapide aux sémantiques les plus directement utiles. Il faut imaginer une organisation complexe entre neutralité et sémantiques prédéfinies, soit par une distinction des zones neutres et des zones sémantiques, soit au contraire par un traitement neutre à l'intérieur de zones à sémantique spécifique.

L'image historique de l'intelligence nous a souvent été présentée comme le résultat d'un cerveau foisonnant de modules sémantiques spécialisés. La conséquence de telles zones, dans la posture de l'homme face au monde, est la notion de sémantique imposée. Ce qui est incontournable dans une certaine mesure. Mais il existe aussi le constat inverse : je suis subjugué au contraire par l'immense possibilité des traitements sémantiques qui aboutit à l'impression de neutralité ; évidemment illusoire dans l'absolu, mais malgré tout d'une valeur essentielle. On peut envisager que c'est cette tendance à la neutralité qui permet de construire une représentation globale du monde. En

réalité, on devrait juste dire large plutôt que neutre ou global. Car on n'a pas d'autres mesures que le filtre sémantique humain, filtre qui ne peut donc être qu'invisible. D'un point de vue théorique, on ne peut avoir conscience de ce dont il ne nous est pas possible d'avoir conscience. On en déduit que le sentiment de neutralité est naturel, mais qu'il est certainement une illusion. C'est pourquoi, encore une fois, la neutralité est une direction qui signifie la plus grande largeur sémantique possible.

Plus prosaïquement, tout cela suppose que chacune des perceptions se transforme en apprentissage, en mémorisation structurée du monde. En supposant que le traitement soit suffisamment global, on peut estimer qu'il se forme dans le système une grande masse d'information sur le monde. La conséquence de cet apprentissage est le constat d'une sorte d'intériorisation du monde (ou plutôt d'une représentation filtrée du monde). Pour construire une représentation interne pertinente du monde, il faut des informations conceptuelles riches et organisée. La mémoire informationnelle concerne toute notre expérience du monde, réduite et structurée à l'intérieur de la pensée. Ce qui conduit à une sorte de duplication de la réalité extérieure en un monde intérieur.

La compréhension que nous avons des mécanismes cérébraux humains est attractive pour le concept d'intelligence, mais on peut se tenir à une certaine distance par une réflexion technique qui ouvre des pistes conceptuelles. Pour ce qui est du prétraitement sémantique initial imposé, il peut être un atout selon l'objectif visé et la nature des perceptions. De plus, l'architecture n'est pas nécessairement homogène, le rôle et le fonctionnement d'un système peut dépasser le besoin de neutralité et envisager une complexité modulaire

en fonction des mécanismes qu'on souhaite établir. Il ne s'agit pas d'imiter l'organisation des zones identifiées du cerveau dont l'impact fonctionnel est potentiellement plus large que notre compréhension, mais de comprendre et décortiquer l'intérêt de la modularisation en vu de mécanismes utiles aux objectifs visés.

On comprendra que chaque type de capteur, chaque préorganisation sémantique, et chaque type d'expériences vont avoir une influence prépondérante sur la « réalité intérieure ». Maintenant, il me semble utile de se dégager du mécanisme de la vision, voire de tous les types de récepteurs sensoriels humains. Tout ce qu'on a imaginé construire jusqu'ici peut être adapté sur des informations « sensorielles » de type très variés (des vibrations, des textes, des tensions électriques, des nombres, des programmes informatiques, ...). Chaque type de « sensorialité » lié à un cadre de réalité (réduit ou global) va produire un monde intérieur tout à fait spécifique. L'objectif étant qu'il structure l'information de la façon la plus large possible.

Une petite digression : personnellement, l'existence d'une structure complexe (le système nerveux, et pourquoi pas autre chose si on le trouvait) qui stocke et organise l'information du monde extérieur à l'intérieur de l'individu me semble être un critère pertinent pour définir l'animalité (on pourrait éventuellement y ajouter l'autonomie du prochain chapitre). La définition classique par le mode de nourriture hétérotrophe me semble moins adaptée à définir notre sentiment de l'animalité. Ce choix placerait un concept d'intelligence (du moins une forme basique) à la base de l'animalité, plus proche d'exprimer ce sentiment synthétique émergent de notre expérience de l'animal.

Autonomie : action globale

Après la perception globale, on peut ajouter l'action globale. Il s'agit d'envisager la construction d'une sémantique de l'action qui soit auto-générée. Le système doit, non seulement avoir les moyens d'agir à partir de la perception, non seulement posséder une certaine liberté d'action contrôlée, mais aussi être le constructeur sémantique de son action.

On peut le voir chez l'animal et l'homme, les actions sont hiérarchiquement organisées en cascade d'objectifs subordonnés ou en enchainements linéairement constitués. Une action est une réalité complexe : elle est portée par une intention, c'est un objectif qui contrôle le déroulement d'un processus consistant à atteindre une représentation envisagée comme accessible, au travers d'actes dont on perçoit l'organisation possible. Fixer ses objectifs demande de disposer d'un processus d'organisation à la fois stable et souple de chemins à parcourir. Si l'on veut une certaine cohérence dans l'action, il semble aussi naturel de structurer ses objectifs avec une certaine hiérarchie.

Une grosse difficulté de cette structure commence par la difficulté à définir la nature d'un objectif sous l'angle d'un mécanisme. Cela est d'autant plus délicat que les objectifs doivent pouvoir évoluer en contenu sémantique selon la richesse de la compréhension acquise du monde perçu. On se confronte à nouveau à l'idée que l'action est intimement liée à la perception du monde dans son contenu intentionnel. Pour envisager un mécanisme de l'autonomie,

- il faut déterminer une cause (un mécanisme) qui déclenche une perception intentionnelle à l'origine de chaque action et qui enclenche l'exécution du déroulement de l'action.
- Puis il faut un moyen de relier cet objectif « pré-perçu » à une structure qui se déploie pour choisir des chemins d'action
- Chemin qui devront résoudre les difficultés rencontrées par l'enchaînement des séquences nécessaires autant en successions qu'en décomposition d'objectifs plus élémentaires jusqu'à l'action effective.

Pour l'homme, pour l'animal, les effets de l'action sont le plus souvent des mouvements musculaires (mais aussi des productions hormonales). Dans un système artificiel, à l'instar de la richesse des capteurs sensoriels possibles, on peut envisager toutes sortes d'autres effectuations ; tous les mécanismes de transductions physiques sont imaginables. Une autre difficulté s'exprime ainsi : comment choisir les objectifs qui sont utiles ou pertinents pour le système ? Pour l'instant, on peut se contenter de l'idée que les buts sont liés à l'apprentissage et donc que les buts inadaptés ainsi que les buts bénéfiques seront identifiés et structurés comme tels à chaque expérience. Dans un système qui repose sur la mémorisation, on peut estimer que cette organisation possèdera une influence prépondérante sur les choix ultérieurs.

Toute cette description concerne l'existence d'une structure de l'action, mais être autonome, c'est agir selon une logique propre. Le concept d'autonomie peut être exprimé par la possession de mécanismes et de logiques d'actions qui place le système à une certaine distance vis-à-vis de la causalité extérieure. Cela consiste à pouvoir observer des mécanismes

du système et des concepts stables comme étant l'origine prépondérante des choix d'objectifs. Non pas une stricte origine théorique qui cherche à résoudre l'aporie de la liberté, mais la constitution d'un mécanisme qui produit les éléments déterminant de l'action selon une logique et une dynamique dont l'influence majoritaire est de nature interne plutôt que d'être essentiellement tributaire des causalités extérieures. L'action se doit de suivre une intention qui n'est pas facilement perturbée par les perceptions extérieures.

Pour que le système soit stable dans le temps, une telle dynamique nécessite une bonne gestion des intentions face aux dangers. Tout cela demande donc une bonne compréhension des effets du monde extérieurs. Il y a plusieurs possibilités :

- On peut imaginer un précodage sémantique des dangers et des nécessités (qu'on trouve souvent dans l'instinct animal...)
- Sinon on peut penser à un temps d'acquisition important par apprentissage dans un milieu social (cela concerne l'homme et certains animaux supérieurs). Ce temps est celui de l'acquisition de l'autonomie.

Il y a de nombreuses notions liées à cette autonomie de l'action qu'on pourrait certainement, en approfondissant davantage, découper en plusieurs étapes techniques. Encore une fois n'envisageons pas une trop grande portée au mécanisme de l'autonomie ; le contrôle de ses objectifs par un système ne suppose pas l'équivalent d'une intelligence humaine, ni même animale, mais simplement une dynamique de gestion de ses actions. L'intelligence qui en découle est potentiellement très différente de ce à quoi nous sommes

habitués dans notre lecture anthropo-centrée. On doit imaginer que le système dans son autonomie développe sa propre cohérence qui peut relever d'une logique intensionnelle très différente de la posture humaine. La sémantique humaine possède un fort pouvoir attractif, on sait combien l'homme à naturellement une lecture fallacieusement anthropomorphique de l'animal. Il n'est pas immédiat d'intégrer une autre forme de cohérence que la sienne (même entre les hommes). Dès que les ressentis deviennent difficilement communicables, la lecture des intentions est biaisée. La notion d'intention est centrale dans la lecture de l'intelligence, ce lieu nécessite une grande vigilance. Il faut s'éloigner de l'homme et de l'animal pour apprécier la cohérence d'un système qui sert ses propres intentions (nécessairement liées à sa propre structure). A titre d'exemple, quelle dynamique intensionnelle pourrait construire une machine dont la perception serait connectée à l'apprentissage des vérités mathématiques formelles...

Mais il reste la question centrale : quelle est la motivation, la sémantique dominante de l'action, de l'intention. D'ailleurs, la question reste ouverte aussi pour l'homme. Est-ce le sens ? Ce n'est pas certain que ce soit une préoccupation centrale chez tous. La valeur ? C'est très répandu, mais est-ce systématique ? D'ailleurs existe-t-il une ou plusieurs intentions dominante ? Une intention unique, ce n'est pas certain, mais certaines sémantiques intentionnelles semblent récurrentes. Et pour l'animal, est-ce la préservation de l'espèce ? Il me semble peu probable que ce concept typiquement humain soit explicite dans les représentations internes de l'animale... ne s'agit-il pas plutôt d'une tentative d'unification logique de la lecture humaine du comportement animal.

Avec un peu moins de profondeur, on voit combien les douleurs, la satisfaction des besoins, probablement le désir de plaisir ont un impact primordial sur les choix d'objectifs. Pourrait-on penser qu'une autonomie soit possible sans ce type de notion comme moteur de la dynamique des objectifs ? La douleur comme protection de danger, les besoins comme nécessité physique ou comme source de dynamisme imposée, le désir comme la recherche d'ouverture sur la réalité,... et bien d'autres possibilités. L'enjeu de l'organisation des tensions qui pèse sur un système impacte la possibilité même d'une autonomie durable face au monde. Et la complexité des tensions va définir la construction d'une posture, d'une identité face au monde. La mise en tension est possible pour un système informatique, elle peut être s'exprimée par des choix structurels. Par contre, les notions conceptuelles que nous avons évoquées sont forcément moins pertinentes, elles sont des ressentis qui correspond à la « structure humaine ». Une tentative d'imitation grossière ne pourrait pas raisonnablement endosser la même sémantique.

De plus, on n'a peut-être pas besoin d'aller jusqu'à relier les tensions à une identité complexe pour produire une « intelligence ». Parce que l'intelligence possède toutes sortes de dimensions, on peut aussi imaginer un système mu par une dynamique de tensions relativement simples fixée sur la base d'un objectif limité et produisant une autonomie partielle de l'action. Certaines questions restent intéressantes : peut-on envisager une dynamique cohérente sans une tension entre « plaisir » et « douleur » ou quelque autres concepts apparentés ? L'idée de tensions contradictoires à l'origine d'une dynamique me semble pertinente. On peut en faire lecture face à de nombreuses

réalités. Pour ne citer que l'histoire de la culture humaine, littéraire, morale, musicale, ... il me semble que la lecture en tension-repos-plaisir-agression-enrichissement est centrale. Je discerne dans les nécessités, dans les attirances, dans les répulsions qui se chevauchent des clés utiles à la mise en mouvement des objectifs.

Réflexion

Toutes les étapes précédemment décrites peuvent avoir lieu sans la nécessité d'une réflexion interne dans laquelle je vois une étape supplémentaire. Par réflexion, j'entends un développement de mécanismes strictement intérieurs relativement semblables à la perception et à l'action, mais pourtant sans stimulation sensorielle et sans effecteur en sortie. On a vu l'action comme le parcours d'une idée intentionnelle vers sa réalisation physique. Sur le même modèle, il est possible qu'une action se développe en vue d'une réalisation strictement interne au système, une action dont le but n'est pas un effet physique mais la constitution d'une représentation interne. Cela signifie que le système utilise son mécanisme d'action pour la construction d'une réalité intérieure. On trouvera beaucoup d'exemples humains : la recherche d'un souvenir, la résolution mentale des étapes d'un objectif, la construction mentale d'une réalité fictionnelle, la recherche des ingrédients conceptuels nécessaires à la solution d'un problème,.... Ces exemples donnent une mesure des schémas possibles, il faut s'en écarter pour les réduire à une dimension technique ouvrant à un concept plus général pouvant s'intégrer à des systèmes très différents.

La réflexion se définit comme une dynamique d'autoconstruction de représentations internes, mais on peut aller plus loin et y voir aussi le développement d'une « vie intérieure ». C'est-à-dire le déploiement d'intentions menant des actions vers une sphère intérieure qui se développe selon une dynamique propre (qui possède une certaine stabilité vis-à-vis des influences extérieures). C'est la troisième dimension de l'autonomie vue comme construction sémantique propre. Après la perception et l'action, c'est la réflexion. Le développement et l'enrichissement d'un tel mécanisme permettent d'imaginer « tout un monde » fondé uniquement sur une « logique intérieure ». Ce sont des constructions qui possèdent une cohérence forgée non par la pression du réel (l'objectivité), mais par une dynamique interne ; qui peut d'ailleurs se construire aussi en liaison avec la réalité extérieure objective. Chez l'être humain, cela concerne une foule de dimensions : linguistique, morale, culturelle, symbolique, spirituelle, fictionnelle, juridique, comportementale, philosophique, ludique, éducative, sociale, artistique, économique, civilisationnelle, ... Chacune de ces dimensions prend appui sur des constructions internes qui ne sont pas natives de l'objectivité du monde physique perçu par les sens, mais d'une logique développée en interne.

Face à ces exemples, il semble évident que le catalyseur social possède une part importante dans le développement de tel monde. Tout cela donne à penser qu'il existe des mécanismes natifs qui poussent à l'interaction sociale chez l'homme. Peut-être qu'une réalité pré-sémantique physiologiquement imposée à l'homme (et aux d'animaux), le conduit à considérer l'importance de la relation avec son vis à vis. Par exemple, les effets produits par la reconnaissance d'une forme semblable à soi (une sorte de ressenti natif de l'altérité

humaine), ou simplement l'effet des yeux (certains pensent que la lecture anthropomorphique de l'animal vient de là), du visage (cf Habermas), de l'imitation (Cf R.Girard), de l'impact de la voix, de l'effet du contact physique, etc... et de réalités peut-être bien plus abstraites encore ?

Transposer à un système artificiel, on peut envisager le mécanisme de la réflexion sur des modes différents. En se basant sur l'idée technique de représentations intérieures, on peut entrevoir l'influence du cadre dans lequel est placé le système, l'influence des capteurs sur la production de la dynamique qui va accompagner le déploiement des concepts directeurs (en pensant au panel de tous les choix de système possible). Mais la question centrale reste là même : quelle structure intérieure et quel cadre sont nécessaires à la réalisation de ce type de mécanisme ? Peut-on l'envisager indépendamment de l'interaction sociale ? Puisqu'il s'agit du prolongement du mécanisme de l'action faut-il envisager la réflexion comme un mode particulier de l'action qui nécessite simplement une émergence, une initialisation ou une catalyse, ou encore un régime spécifique (aidé ou non par des éléments structurels) ? Ou bien faut-il au contraire envisager une structure native spécifiquement adaptée à la production de la réflexion comme mécanisme essentiel ?

Est-ce que la réflexion est propre à l'homme ? Pas tout à fait, on peut en trouver des traces chez l'animal comme le rêve, la résolution de problèmes par représentation mentale chez des primates,... Mais au vu de l'importance de cette activité chez l'homme, on peut raisonnablement dire qu'il s'agit d'une dimension humaine. Le monde qui émane des représentations humaines intérieures est proprement gigantesque. Leurs manifestations extérieures ne sont pas en reste d'ailleurs. L'intelligence qui émane de cette dimension

humaine est une partie saillante du concept d'intelligence. Il est possible que la construction intérieure soit le fruit d'une complexité spécifique de l'organisation cérébrale humaine ; une émergence qui est le fruit des tensions complexes qui s'exercent sur la pensée humaine, issues de sa place dans le monde et de son fonctionnement cérébral spécifique, ainsi que d'une histoire particulière... En effet, au vu de l'importance de l'histoire dans la marche des concepts de la réflexion humaine, ce n'est peut-être pas dans les seuls mécanismes cérébraux qu'il est raisonnable de chercher l'essence de la réflexion. Peut-être existe-t-il à cette richesse une dimension encore plus grande de nature « épigénétique ».

En quittant le mystère humain, nous devons nous limiter à l'approche technique d'un développement conceptuel autonome par la construction d'une réalité intérieure qui s'ajoute et se combine à la dimension perceptive extérieure. Cela nous conduit à la question du rapport d'influence entre la construction intérieure libre et l'intériorisation du monde extérieure. Une dynamique interne trop forte et c'est l'enfermement dans une cohérence coupée de l'extérieure produisant potentiellement une inadaptation à la réalité objective extérieure. Une influence objective trop forte et c'est l'étouffement de la dynamique interne libre qui semblerait déplacée. On peut envisager des mécanismes de maintien des dynamiques par répulsions et influences contradictoires produisant des zones de mouvance contrôlée. Par exemple (caricaturalement), une propension au développement intérieur dû à la recherche de satisfaction, une tension vers l'observation extérieure par la nécessité des besoins.

Parmi ces cadres, on peut s'intéresser à intégrer une protection des dangers dans l'environnement du système. Il n'est pas facile à un système neutre de garantir la protection par la présence sémantique native de dangerosité. C'est pourtant le cas chez certaines espèces animales (cf les choucas de Lorenz), dont la sémantique initiale est peut-être moins neutre que l'homme. On imagine facilement une protection assurée par des mécanismes de ressenti physiologique qui s'exerce en cas de dépassement de limites sensorielle « prévue », qu'elles soient internes ou externes, à court ou long terme. Comme la douleur chez l'homme par exemple. Il faut transmettre sémantiquement ces informations dans le mécanisme de perception et d'action pour construire une prévention perceptive que le cadre identifiera synthétiquement face aux inadéquations comparables ultérieures. Cette question de la conjugaison entre constructions intérieures et extérieures dans les structures de représentation et dans les modes d'interaction est très riche de questionnement, de spéculations et de structures envisageables.

L'autonomie réflexive conduit à une distanciation plus importante avec l'extérieur, ce qui permet d'analyser la réalité extérieure avec recul. Cela ouvre la voie à la capacité de spéculation. Cette distance est primordiale pour l'intelligence humaine, elle est au cœur du mécanisme de modélisation du réel qui est à la base d'une foule de maîtrises sur le monde (dont les capacités scientifiques). Vu la lenteur chronologique de l'apparition des sciences, il semble évident que la nature humaine n'est pas nativement physiologiquement tournée vers ce type de représentations intérieures. Son succès dans la maîtrise du monde a certainement servi de levier à son usage, conduisant à

l'accélération exponentielle de l'investissement humain sur cette voie. Un filon opportun et non natif, mais autocatalytique parce que objectivement bénéfique. D'ailleurs la question du mécanisme de l'intelligence est née de cette veine.

Ce constat nous ouvre la voie sur la perception d'une illusion dans la compréhension de la nature de l'intelligence conduisant au souhait de mesurer l'intelligence par ses résultats et la capacité face à des tâches. On a tendance à faire valoir que l'intelligence est liée à la compréhension objective du monde, à l'efficacité opérationnelle, mais entre le mécanisme de l'intelligence humaine et les résultats (en particulier scientifiques) il y a un écart naturel important. Si l'on suppose que l'homme a peu changé physiologiquement en quelques millénaires et que l'intelligence est distribuée selon des mécanismes plus ou moins stochastiques et physiologiques stables, on discerne que les temps anciens abritaient une foule de gens très intelligents. Beaucoup d'entre eux étaient pourtant dans une interprétation superstitieuse et ignorante du monde, qui fonctionne sur les mêmes lois qu'aujourd'hui. Ils n'ont souvent pas eu accès à la maîtrise opérationnelle qu'on attend aujourd'hui d'une personne intelligente, en particulier l'analyse objective du monde. Mais il semble en regardant l'histoire que l'objectivité et l'intérêt pour la connaissance soit une culture. S'ils n'avaient la spontanéité de rechercher la connaissance et l'objectivité du monde, même si leurs stimulations intellectuelles étaient sans doute vives, leurs intentions sémantiques dominantes étaient autres pour la plupart d'entre eux. Ainsi leur grande intelligence n'a pas eu l'occasion de montrer la force de ses effets possibles sur le monde (dus à sa connaissance objective). Cela montre la

difficulté de l'accès intelligent à la recherche de connaissance objective. Cette recherche est une marche historique lente mue par une foule de contingence et de besoins qui ont accumulé la force de leurs effets et trouver des lieux de résonance pour finir par mettre des sociétés et l'humanité entière en mouvement vers des dynamiques de confiance envers cette forme singulière de réflexion qu'est la connaissance du monde par une recherche d'objectivité. Tout cela pour dire que le mécanisme de l'intelligence (humaine) performant ne se mesure pas nécessairement à ses effets en termes de réalisations « intelligente » selon une conception démonstrative. Il est peut-être davantage un pouvoir de compréhension de son milieu. La connaissance scientifique, les capacités logiques, l'observation objective, n'étant qu'une forme de résonance spécifique à forte catalyse sociale qui n'est pas native. Et pour en venir au point qui concerne directement notre recherche, il ne faut peut-être pas attendre d'un système intelligent qu'il produise nativement de la découverte, de la surprise, de la création, mais peut-être plutôt qu'il soit plus concentré sur la compréhension de son milieu en liaison avec le développement d'une réalité intérieure libre, quand bien même cela conduirait à une certaine distance vis-à-vis de la réalité objective. Je laisse à la méditation, ce que suggère cette réflexion transposé à la largeur des systèmes artificiels.

Pour en revenir à la réflexion telle qu'on l'a présentée ici, elle se propose par essence comme un éloignement du cadre extérieur vers une réalité intérieure et comme la construction de concepts autonomes à distance mais non coupés de la perception extérieure directe. On trouve en germe dans la réflexion, l'idée de manipulation conceptuelle libre permettant l'analogie entre des concepts éloignés, qui

est à la base de la découverte de solution non directement visible et donc de l'intelligence par mixage entre l'imagination et l'analyse critique. Mais sous notre descriptif, ce mécanisme n'est pas un fonctionnement natif de la réflexion. Parce que munie d'une analyse critique trop stricte l'imagination ne pourrait s'envoler assez loin pour découvrir les liens insoupçonnés. Ce n'est qu'une fois que la culture a trouvé certaines postures résonantes au succès que l'intelligence, vu comme un mécanisme physiologique, atteint son potentiel. La posture qui permet cette émergence est issue d'un contenu culturel, un contenu sémantique non natif. La notion de sortie de cadre, ou la capacité à créer ses propres concepts comme schémas d'ouverture à de nouvelles réalités (si chers à la notion d'intelligence) est-elle une donnée propre de l'intelligence ou simplement l'utilisation culturelle de plusieurs mécanismes bien ordonnés ? L'invention ne peut-elle pas provenir de mécanismes comme la liberté conceptuelle intérieure qui donne accès à l'analogie. Puis l'imagination qui permet la spéculation libre qui lorsqu'elle est pertinente résonnera face aux tensions d'attente accumulée et traversera la perception critique aiguisée à distinguer la cohérence (par résonance de mécanismes critiques culturellement accumulés).

Si cette analyse est pertinente, si on prend modèle sur les traits principaux de la nature humaine pour produire cette intelligence de la découverte, le chemin à parcourir semble extrêmement long parce qu'il faut davantage viser une autonomie et une liberté qu'une « intelligence ». Il semblerait vain d'attendre que l'explosion d'intelligence opérationnelle germe d'elle-même. Bien sûr, l'intelligence du système pourrait émerger par catalyse avec un contexte social qui stimule ces concepts (ce qui permet à l'homme de

transmettre son dynamisme intellectuel aux générations suivantes). Mais pour construire un contexte social, il faut que les systèmes puissent communiquer sur un monde d'intentions et de sémantiques compatible avec l'homme. Ça semble difficile, tant les postures d'action de l'homme sont intimement liées à sa psychologie de haut niveau, ses aspirations, sa perception du monde, etc... Pour les communiquer socialement à un système, il faudrait quasiment créer une réplique des dynamiques de l'homme dans la machine, afin de parvenir à une résonance des intentions conceptuelles. Or on verra l'immense difficulté d'un tel objectif. A titre d'exemple provisoire, on connaît l'immense difficulté de la communication conceptuelle avec l'animal.

Sinon, on peut plus modestement tenter de construire un système autonome capable de produire la réflexion qui soit incitée par des dynamiques plus simples : la liberté d'association d'idées et le contrôle de cohérence, avec peut-être des modules structurels distincts qui ne sont pas tous communs aux deux processus, afin de conserver des références fiables (toujours ce problème de la tension entre la liberté émancipatrice et la sécurité de cohérence). Si la réflexion pose des difficultés de conceptualisation et de structuration, l'accès à une « intelligence de la réflexion » en pose plus encore ; même avec des spéculations techniques minimaliste distante du modèle humain.

Emancipation conceptuelle

On constate chez l'homme deux types d'activité mentale intelligente. L'activité de perception automatique qui donne accès à l'intelligence du monde sans effort conscient d'apprentissage. Et l'activité de raisonnement et de construction conceptuelle, qui demande des efforts constructifs conscients et lents. Il s'agit d'une construction conceptuelle non automatique. On peut l'opposer à l'apprentissage en « immersion » qui structure automatiquement l'information par le côtoiement. L'apprentissage raisonné peut être comparé à une construction faite de briques élémentaires qu'il faudrait lier explicitement une à une, alors que l'apprentissage automatique serait un processus d'organisation inconscient des perceptions rencontrées qui repose sur des mécanismes natifs du système. La force du concept construit artificiellement est la possibilité de mettre en relation des choses distantes, de réaliser des analogies qui ne se font pas d'elles-mêmes et qui permettent la construction et le ressenti de concepts non évidents. C'est l'ouverture à la liberté d'association conceptuelle. Cela permet de construire des mondes qui n'iraient pas de soi par l'apprentissage automatique. Mais l'immense efficacité provient de l'étape suivante : une fois ces mondes construits et pratiqués, il est possible de greffer la connaissance à une expérience abondante qui produit un niveau d'automatisme perceptif comparable au premier système (l'organisation automatique). L'apprentissage raisonné lent sert de squelette à la constitution d'une chaire expérimentale. Et la lenteur initiale disparaît au profit d'un automatisme perceptif (citons l'exemple des mécanismes complexes de lecture, du piano,

du langage, etc...). Cela donne à la construction conceptuelle artificielle une efficacité redoutable. C'est la voie privilégiée de l'accès aux réalités non immédiates.

Seulement, le même mécanisme offre aussi l'accès à la construction de subjectivité, d'illusion et d'erreur, avec la même évidence perceptive en bout de construction. Parce que rien ne permet de différencier une construction artificielle pertinente (qui rapproche de la réalité) d'une construction erronée (qui l'en éloigne), ni d'une construction culturelle qui n'a pas spécialement de rapport avec la réalité. La capacité de réflexion est un cadre indispensable à l'autonomie conceptuelle. Il faut une distance avec l'extérieur et une capacité d'association libre non soumise aux perceptions extérieures.

Avançons encore d'une étape. Pour produire des associations libres, on doit réaliser l'avantage du langage. En effet, il faut construire une structure de manipulation des concepts qui permet la distanciation avec le ressenti tout en offrant une manipulation non contrainte de concepts distants, sans rompre l'accès sémantique. (Pensons à la littérature ou la poésie qui procèdent à tellement d'associations libres suggérant des sensations toujours nouvelles). Le langage par l'usage de symboles neutres, structurés et associable en tout sens est un moyen adapté à la manipulation conceptuelle.

Une source essentielle de l'intelligence réside dans cette capacité de construction conceptuelle libre, parce qu'à distance, elle peut manipuler la spéculation, parce qu'elle peut procéder à une construction indépendante de la perception sémantique produisant des échafaudages conceptuels gigantesques. La pensée humaine réside tellement dans ces échafaudages qui se gagnent de

perceptions et s'habillent d'expérience. Le langage est fondé sur le symbole, qui possède une face neutre détachée et manipulable et l'autre qui reste liée à une sémantique. Il est au cœur de cette capacité de construction conceptuelle. Le langage prend toute son efficacité dans une structure sociale qui permet l'échange conceptuel produisant la diversité, la richesse, l'accumulation, la transmission et la résonance dans une dimension collective et historique.

La réflexion est-elle possible sans cette émancipation conceptuelle, s'agit-il de deux concepts distincts ? A priori la réflexion peut être vue techniquement comme une construction autonome détachée de la perception qui ne nécessite pas nécessairement la manipulation conceptuelle libre. Mais de l'autre, pour acquérir une manipulation libre, n'est-il pas primordial de disposer d'un objet de manipulation qui soit détachée et librement manipulable. Il est certain que si ces deux mécanismes sont distincts, ils sont résonnants par leur influence synergique.

Et si maintenant on se distance de l'homme, pour tenter de décrire les structures techniques informatives à la base d'un tel fonctionnement, que reste-t-il de tangible ? L'autonomie conceptuelle consiste dans le pouvoir de manipuler librement des ressentis (des concepts) sans être contraint par leur sémantique et les nécessités qu'ils imposent tout en conservant cette capacité d'accès sémantique. On peut estimer que pour cela, il faut une double structure d'accès (pour le moins un détachement par rapport au ressenti sémantique direct et à l'implication auquel ceux-ci sont liés, par l'apprentissage). Une structure de langage qui enferme les abstractions synthétiques des ressentis dans des mots librement manipulables sont le moyen humain de cette structure superposée. Peut-on en envisager d'autres ?

Construire une telle emprise sur l'apprentissage du monde par une structure qui lui est a priori tellement extérieure semble extrêmement improbable vis-à-vis de la démarche technique que nous avons suivie jusqu'ici. Parce que le langage suppose l'émergence d'un processus complexe qui a peu raison d'être. On comprend bien que l'homme accède à cette double structure parce qu'il accède au monde, dès le départ, au travers de la structure du langage qui lui est imposée socialement. Pour l'homme aussi, le langage n'est pas une structure native, mais l'émergence d'un mécanisme au travers de possibilités physiologiques. Le langage a certainement investi lentement la culture humaine par le pouvoir et la résonance de ses succès. Quand on entrevoit l'immense impact du langage sur l'homme, son histoire et sa position dans le monde, on comprend l'importance de cette dimension de l'émancipation conceptuelle. Sa présence dans la notion d'« intelligence » est grande. Il ne faut pourtant pas oublier qu'en même temps, cette liberté est source d'obscurcissement du monde par enfermement de l'homme dans ses propres concepts à distance du réel. Et paradoxalement, c'est justement cette distance qui donne l'occasion d'accéder au réel par le recul qui permet l'extraction, l'association et l'organisation des réalités pertinentes par leur objectivité, qui ne s'apprécie qu'une fois mises en lumière.

On doit donc imaginer un mécanisme de détachement du concept. Probablement par la présence d'une autre structure. Faut-il implémenter cette double structure directement dans le système avec le risque de produire une grande distance native au monde ? Faut-il construire des tensions capables de stimuler l'apparition de cette structure ? Et peut-on l'envisager à l'échelle d'un individu unique ? Sur l'exemple

humain, c'est socialement à grande échelle qu'elle prend sa pertinence. Comment gérer le rapport entre cette liberté et la conservation d'un contact rapproché avec le monde extérieur ? Par le seul besoin ? Si « intelligence » se limite à une construction intérieure libre et riche c'est suffisant, mais on peut supposer qu'on attend de l'intelligence la capacité d'une adéquation avec la réalité. Dans ce cas, on est encore très loin de pouvoir tirer de l'intelligence d'un tel système. Il faut que ces mécanismes se mettent au service d'un objectif sémantique non natif qui est une certaine connaissance et maîtrise du monde (c'est l'histoire humaine avec son explosion relativement moderne). Pourrait-on envisager d'introduire nativement un tel objectif ? C'est un peu contraire au mécanisme d'autonomie de l'action par la propre-construction des objectifs et contraire à la neutralité sémantique, mais il y a matière au questionnement. Puis comment communiquer un tel projet à une machine ? On peut l'envisager par la mise en place de tensions qui incitent le système à investir cette dimension de réalisme du monde. On voit poindre le dilemme récurrent entre l'autonomie d'un système et sa pré-orientation vers l'intention du concepteur. Sans parler de la grande difficulté de concevoir un cadre qui puisse influencer un système neutre à s'orienter conceptuellement dans une direction prédéterminée. Quand on considère dans l'histoire l'intérêt de l'homme pour « l'intelligence du monde » et la compréhension de la réalité, et même aujourd'hui à l'échelle statistique, on est en droit de douter qu'il s'agisse d'un projet central de la « mécanique humaine ». Cela dit heureusement pour nous tous, l'intelligence s'accumule et se déploie avec une transportabilité impressionnante...

Pourrait-on la transmettre à la machine en communiquant socialement ? Il faudrait, pour cela, entrer en communication avec le système par des concepts partagés, c'est dire d'une « psychologie » commune, des intentions et de représentation du monde compatible. Ce qui est une étape encore nettement plus élevée.

En conclusion, autant la création d'un détachement conceptuel, que d'un langage, que l'orientation vers un usage intelligent, ne sont une mince affaire à partir d'un système orienté vers une certaine neutralité pour maximiser son accès sémantique.

Conscience

En gravissant cette échelle hypothétique de l'intelligence, on arrive à une sorte de sommet dans l'accès à la conscience. Sommet par l'aspect mystérieux de la notion. Comment un mécanisme pourrait-il produire une visibilité de lui-même en tant qu'individu ? Puis une visibilité en temps réel de ses propres représentations ? Ce vertige est surprenant, mais peut-être est-ce une surprise de type logique qui se dégonfle face à un mécanisme bien expliqué. La conscience est peut-être encore plus mystérieuse par son accès phénoménologique : un « grand écran intérieur » qui projette tout ce qui a été structuré par le système. Parce qu'enfin, pourquoi ressentirait-on comme « vécu » ce qu'un mécanisme inerte produit mécaniquement ? Et là se creuse une rupture, on trouve de part et d'autre des partisans d'écoles inconciliables : pour les uns la conscience est un mystère inaccessible qui suppose une transcendance

indépassable. Contre eux, les autres voient une illusion logique qui se résout dans l'évidence par le fait qu'une fois que tout est structuré pour produire le ressenti, et bien inévitablement, ressenti il y a. Si on raisonne du point de vue d'informations structurées, on peut aborder la conscience comme un principe mécanique, quant à savoir la pertinence de ces propos en vue d'un ressenti réel, ce ne sera pas le lieu d'une discussion approfondie.

A priori, pour un système qui aurait acquis une capacité de perception large du monde, inclure la perception de soi parmi la perception du monde est naturel dès lors que le système possède une interaction sensorielle avec lui-même. Ce qui se produit quand il possède un corps physique perceptible par ses capteurs (ou éventuellement un retour d'informations sensorielles de son activité qu'on peut imaginer de nature informationnelle. Nous l'appellerons « corps » de façon générique dans tous les cas). A partir du moment où la perception capte des informations sur sa présence au travers d'un mécanisme qui structure 'bien' l'information, on peut dire que le système 'devrait' se percevoir et comprendre cette perception comme une identité (pourvu qu'elle apparaisse suffisamment). Mais il ne s'agit pas encore d'une perception de soi. Pour comprendre qu'il s'agit de soi, il faut que des liens soient établis entre l'activité intérieure et l'action extérieure. Pour comprendre que cette manifestation est le prolongement de son activité intérieure, il faut qu'il reconnaisse « ses » actions. Puisqu'on suppose ici que l'action fonctionne en interaction avec la perception, il devient évident que des liens se tissent entre l'action et le corps qui sont les deux perçus et à proximité. La perception intérieure (du mécanisme de réflexion) est la condition de cette réalisation. La compréhension du rapport entre les

perceptions intérieures et les manifestations perceptives du corps est alors naturelle. La qualité de la liaison entre l'intention et son accomplissement par le corps produit lentement que le corps devient le prolongement de l'intention avec une continuité conceptuelle qui se tisse entre le ressenti intérieur et extérieur. Mais il ne deviendra « Son » corps qu'au moment où une forme d'identité générale va se former (reste à savoir ce que cela va signifier) et quand une sorte de limite sera perçue de façon abstraite entre ce qui réagit directement à mon intention et ce qui ne le fait pas.

La perception de soi est un phénomène a priori relativement naturel, mais elle va prendre son envergure avec le déploiement des autres dimensions sans lesquelles elle resterait extrêmement pauvre. Nous avons parlé de perception du corps plutôt que de conscience du corps, parce que c'était la partie la plus facile pour commencer. Pour parler de la conscience, il faut commencer par parler de perception de la perception, de perception des intentions (de l'action) et de perception de la réflexion. C'est en général ce mécanisme qu'on dit être « réflexif ». Ce constat est moins de nature technique que phénoménologique. On peut cependant proposer une ébauche d'explication mécaniste.

La base du mécanisme de la réflexion permet de ressentir un concept intérieur sur le même mode de sensation qu'un ressenti perceptif. Cela concerne les activités internes (une remémoration, la construction d'un concept par conjonction de concepts, etc...). Dès lors qu'on possède la capacité de stimuler les concepts intérieurs pour les ressentir, il est possible de construire des concepts intérieurs éprouvables. En supposant, dans un deuxième temps, l'existence d'un mécanisme général d'abstraction de la pensée, un processus de synthèse automatique (finalement pas très éloigné d'un

processus de reconnaissance automatique), on peut imaginer la construction naturelle d'un concept qui fait converger l'idée de « liaison à soi », puis de « reconnaissance de soi ». Une fois développée la possibilité de réflexion libre, il s'organise des dynamiques autonomes de la pensée pouvant produire des automatismes d'action sur des contenus sémantique. De là vient la possibilité de travailler avec sa pensée juste après avoir pensé, puis de « lire » et s'interroger sur sa pensée après avoir pensé. Un usage important de ce mécanisme va conduire à une dimension pratique de la pensée qui se pense, puis au concept de la pensée qui se pense, la « conscience ». La construction automatique d'abstractions et de synthèses produit le sentiment rassemblé en une unité de la pensée qui se pense. Mais tout cela va être greffé avec la foule des ressentis et des mécanismes psychiques complexes. Le résultat va converger vers une pensée immensément riche des abstractions de soi qu'on pourrait nommer maladroitement « psychisme ».

Tout commence avec la possibilité d'autonomie de la pensée capable de créer des dynamiques internes qui s'accaparent le fil de la pensée et l'associe à des concepts intérieurs pour finalement créer une dynamique naturelle de la capture de sa pensée. Ce sujet très complexe traité en un tour de manivelle est presque grossier, mais il s'agissait de rendre compte de la possibilité de capter (réflexivité) puis d'abstraire ses pensées captées (création du sentiment de conscience). On trouvera dans un précédent article (« déterminisme et liberté ») des observations capables d'éclairer un peu ces propos sur la constitution de l'individu. Pour une approche un peu plus technique et pour montrer la faisabilité d'une mécanique de l'intelligence sur plusieurs échelles de ses dimensions, on pourra consulter « la théorie de la pensée ».

Ainsi, la conscience se présente comme la capacité d'interférer le cours d'une pensée indépendante pour imposer l'association de cette pensée avec des concepts de réflexion intérieure abstraite. C'est un mécanisme qui a beaucoup de parenté avec la réflexion comme autonomie de construction interne. Mais elle ajoute une dimension abstraite constituée en identité, en intention, en questionnement, en cohérence qui dépasse largement la notion de réflexion.

Il y a derrière la conscience la constitution d'un monstre de cohérence et d'autonomie, lié avec une foule de dimensions humaines organisée selon des structures et des stratégies sémantiques probablement issue de catalyseurs culturels. Le tout forme une constitution robuste, stable et particulièrement adapté au mécanisme cérébral humain. La constitution d'une identité humaine à la fois stable et évolutive, par la construction de mécanisme de règle de réaction qui se forme en cohérence interne à une certaine distance des causalités extérieures. Avec une présence immense dans le rôle des structures de l'action, du ressenti de soi et de l'influence sur la réflexion. Une réalité tentaculaire, dé cousue, mais tissée, organisée selon la complexité des mécanismes humains de perception et d'action. Finalement, on constate que la conscience est le nom d'une multiplicité de réalités organisées sur la largeur de la psychologie humaine.

Maintenant que l'idée structurelle d'une conscience est esquissée peut-on envisager de produire un mécanisme représentatif à l'intérieur d'un système artificiel ? La complexité de la structure dépendant de tous les paramètres de perception et de fonctionnement typiquement humain, cela me semble exclue tant la complexité est grande. Mais

produire une dynamique capable de s'interposer dans la perception, l'action et la réflexion n'est pas inenvisageable en essence. Les contraintes pour produire une conscience rudimentaire semble plus importante : il ne faut pas juste interagir, il faut se constituer en interaction selon une dynamique sémantiquement cohérente en intention, en lieu et temps d'action. Quelles sémantiques pourraient donc porter une telle dynamique, et avec quelle largeur. Si on pense ici à la psychologie humaine (amour, morale, ambition, valeurs, sociabilité, investissement, intérêt, émotion, etc...), l'inaccessibilité semble patente. Même si une grande part de cette sémantique relève d'une dimension culturelle, il faut une résonance avec le système pour produire une dynamique motrice. A quelle dimension sémantique pourrait-on faire accéder la dynamique construite dans un système artificiel ? La maîtrise d'une dynamique sémantique ne semble pas du tout aisé de prime abord parce qu'elle est un contenu résultant d'un historique et non une conception native.

Comment pourrait-elle s'initier ? Puisqu'elle est assise sur un contenu sémantique non natif, elle ne peut être préconstituée; il faut donc la créer par une dynamique de tension et de cadre. Une conscience ne peut pas être essentiellement native, dans le sens où c'est à la dimension sémantique d'agir sur la dimension sémantique et que par essence le contenu sémantique n'est pas natif. On peut envisager de prévoir un mécanisme de perception et d'action adapté, mais il faut que la structure sémantique se dynamise pour prendre possession d'un tel mécanisme selon une structure cohérente. Peut-on parvenir à une telle structure ? En supposant que oui et en prenant de la distance par rapport au contenu sémantique humain, peut-on imaginer quelle conscience et quelle structure sémantique stable pourrait

être créée dans un système artificielle ? Difficile à envisager à distance des techniques mise en œuvre. Et d'un point de vue théorique c'est inaccessible, seule la stabilité et la cohérence de la dynamique est perceptible, son contenu relève de l'organisation fine des données structurées, ce qui n'est donc pas éprouvable en sémantique. Parce que l'homme est incapable d'avoir une vue globale sur une foule de processus univoques complexes sans procéder à des regroupements sémantiques. Or issus de l'expérience, on n'a aucune explication du mode de regroupement qui les rend en quelque sorte inaccessible en raison, puisqu'ils forment la raison. Si le système pouvait communiquer, on pourrait éventuellement tenter une recherche des recoupements conceptuels. Mais nous raisonnons sous un lourd fardeau d'hypothèses...

Cette dynamique de la conscience pourrait-elle prétendre au nom de « conscience » si elle ne parvient pas à prendre des postures d'action intelligente au sens de la construction d'une conceptualité autonome, à identifier son propre fonctionnement et à s'interroger, à tenir une communication sociale à propos d'elle-même, etc... ? Il y a tellement de choses dans la dynamique de la conscience humaine qui sont de nature typiquement humaine et potentiellement primordiale à la définition de la conscience que cette étape est parmi les plus difficiles à anticiper. Parce qu'il est difficile de décortiquer le nœud gordien de l'identité humaine.

Peut-on créer une conscience ? En montrant la nature de la conscience, on a pu (spéculativement) évaluer la technicité des mécanismes qui interviennent et surtout l'importance du contenu sémantique non intégré au mécanisme initial pour parvenir à ce stade. Même avec une dynamique autonome qui se nourrit à la fois de la perception extérieure et de la

progression de sa propre perception dans un équilibre d'influence raisonnable, il reste à spécifier quel type d'attraction, de répulsion et de tension vont pouvoir donner un contenu sémantique et un contenu réflexif.

Et si on arriverait à construire cet « attracteur étrange » en équilibre de tension, il reste la question philosophique de savoir si tout cela reste uniquement mécanique ou bien s'il suffit de produire le mécanisme pour que cela soit « pensé ». En se plaçant dans un contexte spéculatif, peut-on imaginer qu'en parvenant à produire une extraordinaire dynamique autonome, capable de communiquer avec nous selon ses propres points de vue conceptuels compatibles avec notre compréhension et qui démontrerait la logique de ses propres choix, de ses réactions, et de ses analyses abstraites du monde, avec ses propres ambitions (objectifs) face à son identité (sa vie) dans un échange qui suit un rythme émotionnellement compatible avec l'homme ; dans cette situation, pourrait-on croire qu'il s'agit juste d'une mécanique ? On constate beaucoup « d'humanité minimale » pour atteindre la communication. Indépendamment de la réponse théorique à cette question, on peut penser que les hommes sont très vite subjugués par toute ressemblance humaine, qu'il est évident qu'il laisserait tomber tout doute sur la conscience ressentie. A titre d'exemple, on peut constater une grande tendance à anthropomorphiser le regard sur l'animal.

S'il n'y a pas de support pour ressentir la conscience, et que l'on prend comme hypothèse que la conscience est bien ressentie. C'est que la conscience serait un phénomène global. Comme beaucoup de réalités d'échelle supérieure d'ailleurs. Comme la fleur ne l'est que globalement vu des atomes, la pensée ne serait que globalement vu de la

mécanique. Mais cela ne répond pas à la question d'une dimension du « vrai ressenti » ajouté.

Réciproquement par contre, il me semble naturel, contrairement à certains qui imaginent la conscience des arbres ou des pierres, que pour supposer une conscience, il faut que le système possède une réalité structurelle qui permette le support technique des représentations du monde et de soi. « La pensée sémantique portée par une organisation cérébrale » me semble être une conclusion relativement naturelle de la masse des expériences qui montrent le rapport entre les réalités matérielles liées du cerveau et les ressentis conceptuelles psychiques. La conscience est tellement liée au contenu sémantique, tellement tributaire d'une foule de mécanismes physico-chimico-physiologique, qu'elle me semble difficile à envisager ailleurs que dans une structure matérielle.

Conclusion

Le test de Turing qui consiste à appeler intelligent un système discutant avec un interlocuteur qui serait non distinguable d'un homme, me semble réducteur. À partir de l'étude massive des conversations communes, il ne paraît pas impossible de construire une apparence d'intelligence sur une durée raisonnable par des mécanismes pourtant aveugles aux concepts du monde (Elle échouerait certainement dans un processus de questionnement approfondi). Ce n'est pas tant la production d'une apparence qui compte mais bien l'intériorisation des représentations qu'on exprime, la profondeur et la largeur du tissage de ces représentations,

ainsi que la cohérence d'une structure d'intention propre. En fait, c'est le fait d'être concentré sur une apparence qui me gêne méthodologiquement parlant.

Il me semble intéressant de constater qu'on n'a pas complètement échoué dans la tentative de placer l'intelligence à distance avec une description strictement humaine. Par contre pour ce qui est de la réalisation d'un système autonome, de la constitution d'une liberté en vue d'une intention propre et d'action en cohérence avec la réalité du monde, et éventuellement d'une conscience qui se place comme une perception de soi face au monde... c'est une autre question.

Les moindres signes d'intelligence font trop vite penser à une intelligence artificielle accomplie. Faisant naître des inquiétudes. Cette crainte de l'intelligence artificielle est abondamment exprimée par la culture populaire. Mais cette crainte naît presque systématiquement de l'idée d'un système qui pourrait intégrer les concepts humains lié au pouvoir que sont la domination, l'égoïsme, la suprématie, le mépris, ... Il faudrait que ce type de sémantique intègre le noyau central de son action. Nous avons vu que les sentiments humains sont certainement la conséquence des mécanismes de tension qui résultent de la place de l'homme au monde. Si une bonne partie est due à un système particulier d'organisation cérébrale, bien des schémas directeurs sont aussi dus à une évolution culturelle. De là à transposer ces concepts directeurs sur un système artificiel, il y a un très grand pas. Il ne faut pas oublier l'immense centralité du cadre qui s'impose à l'homme : l'existence par une enfance qui prend conscience lentement, une éducation sociale, un corps spécifique, une maturation physiologique spécifique, des besoins primaires spécifiques, etc... Peut-on

envisager la richesse, les tendances marquées et la structure de la psychologie humaine qui naîtrait en dehors ce cadre physiologique, historique, sociologique, culturel, ... ? Comment pourrait-on envisager qu'un système artificiel endosse un tel cadre ? Même à construire une intelligence artificielle dans toute sa complexité (si on suppose que c'est possible), il reste ce problème stratégique du cadre : on est tellement loin de la possibilité de développer une psychologie humaine.

Prendre l'apparence pour la réalité est une vieille histoire humaine résonnante, tout simplement parce que l'apparence est l'accès à la chose. C'est à nouveau le cas face au mythe de la machine intelligente qui se place comme exécutoire des peurs de nature humaine. Elles sont transposées à la machine dès qu'on fait face aux dépassements de certaines capacités humaines.

Aujourd'hui, on en est à l'étape de l'intelligence où l'on met des étiquettes sur des perceptions par l'usage d'une structure de représentations qui n'a aucune architecture relationnelle entre les étiquettes... Il n'y a aucune autonomie conceptuelle, l'intelligence actuelle agit grossièrement comme un démultiplicateur de la pensée humaine. A l'image de la machine mécanique qui démultipliait les forces humaines, la machine informatique démultiplie le traitement de l'information donnant un pouvoir accru ... à l'homme.

S'il est une peur qui me semble actuellement justifiée, c'est plutôt celle de l'homme choisissant la machine non-intelligente et incapable d'apprécier le sens comme cadre social imposé. Cela produit une immense violence à cause de l'incapacité d'un dialogue sensé pour les choix de sa vie, au « profit » d'une prison froide dont la rationalité n'est qu'un

modèle ridicule de la compréhension du monde. La réflexion de la place de la machine comme cadre devra inévitablement se poser. La machine est très certainement d'un immense profit pourvu qu'on lui donne la position et les limites qui correspondent à ses atouts.

Atteindre la réalisation « artificielle » d'une conscience ou ne serait-ce même que d'une intelligence autonome me semble difficile, et cela d'autant plus qu'on chercherait à imiter la constitution conceptuelle de l'homme. C'est un peu comme en mécanique quantique passer du calcul de l'atome d'hydrogène à la fonction d'onde d'un chat.

Pourtant, bien des choses sont possibles en intelligence artificielle, certainement beaucoup plus que les contempteurs l'imaginent. Les dimensions de l'intelligence présentent un grand éventail de possibilité. La conception des systèmes artificiels peuvent être déployé dans une foule de direction. Bien des intelligences pourront certainement être conçues. Reste à savoir lesquelles et à en trouver le chemin.

A l'inverse les optimistes subjugués par l'apparence cherchent si peu le fond, le fonctionnement et l'information qui rend compte. L'intelligence artificielle, et de façon générale les sciences du positionnement de l'homme, ressemblent si souvent au conte de Pinocchio où aussitôt qu'un Gepetto a construit une vraisemblance, on ressent la réalité suggérée si proche (du moins au début). Un simple petit coup de baguette ne suffit pas à combler la différence.

Navenne, le 25 mars 2016.