

BULLETIN N° 112
ACADÉMIE EUROPEENNE INTERDISCIPLINAIRE
DES SCIENCES



Séance du Mardi 13 février 2007

Conférences :

Roger BALIAN de l'Académie des Sciences

« Introduction à l'émergence »

Michel BITBOL membre du C.R.E.A.

« L'émergence sans propriété fondamentale »

Prochaine séance : le Mardi 13 mars 2007

**Conférence de Jacques DUBUCS, Directeur de recherches au CNRS, Directeur
de l'IHPST**

« Emergence diachronique »

ACADEMIE EUROPEENNE INTERDISCIPLINAIRE DES SCIENCES
FONDATION DE LA MAISON DES SCIENCES DE L'HOMME

PRESIDENT : Michel GONDRAN
SECRETARE GENERAL : Irène HERPE-LITWIN
TRESORIER GENERAL : Bruno BLONDEL
CONSEILLERS SCIENTIFIQUES :
SCIENCES DE LA MATIERE : Pr. Gilles COHEN-TANNOUJJI.
SCIENCES DE LA VIE ET BIOTECHNOLOGIES : Pr. François BEGON
PRESIDENT DE LA SECTION DE NICE : Doyen René DARS
PRESIDENT DE LA SECTION DE NANCY : Pierre NABET

PRESIDENT FONDATEUR
DOCTEUR Lucien LEVY (†).
PRESIDENT D'HONNEUR
 Gilbert BELAUBRE
SECRETARE GENERAL D'HONNEUR
 Pr. P. LIACOPOULOS

février 2007

N°112

TABLE DES MATIERES

P. 3 Compte-rendu de la séance du 13 février 2007 :
 P. 6 Compte –rendu de la Section Nice-Côte d'Azur
 P.10 Documents

Prochaine séance : Mardi 13 mars 2007,
 MSH, salle 215-18heures
**Conférence de Jacques DUBUCS, Directeur de recherches
 au CNRS, Directeur de l' IHPST :**
« l'émergence diachronique »

ACADEMIE EUROPEENNE INTERDISCIPLINAIRE DES SCIENCES
 Fondation de la Maison des Sciences de l'Homme, Paris.

Séance du
Mardi 13 février 2007

Fondation de la Maison des Sciences de l'Homme, salle 215, à 18 h.

La séance est ouverte à 18 h. 00 sous la Présidence de Michel GONDRAN et en la présence de nos collègues Gilbert BELAUBRE, Michel BERREBY, Bruno BLONDEL, Alain CARDON, Gilles COHEN-TANNOUJJI, Jean-Pierre FRANCOISE, Irène HERPE-LITWIN, Saadi LAHLOU, Jacques LEVY, Pierre MARCHAIS, Emmanuel NUNEZ, Victor MASTRANGELO, Alain STAHL

Etaient excusés : François BEGON, Françoise DUTHEIL, Marie-Louise LABAT, Pierre SIMON

I) Informations générales

Avant de donner la parole à nos conférenciers, la parole est accordée à notre Collègue Alain STAHL qui nous présente le livre sur la psychanalyse de L. NACCACHE, « *Le Nouvel Inconscient- Freud, Christophe Colomb des neurosciences* ». Il insiste sur la notion d'inconscient cognitif qui a des implications en psychanalyse. Cet inconscient dont l'essence ne serait pas psychique différerait de celui de Freud et serait un nouveau territoire à explorer pour les neurosciences.

II) « L'émergence sans propriété fondamentale » par Michel BITBOL

Michel BITBOL est directeur de recherche au CNRS et chargé de cours à l'université Paris 1. Ses domaines sont la Philosophie de la physique moderne, l'Histoire de la physique du vingtième siècle, la Philosophie des sciences. Il est "membre plein" du CREA: Centre de Recherches en Epistémologie Appliquée - (CREA/Ecole Polytechnique), 1, rue Descartes, 75005 Paris

Il a publié de nombreux articles et ouvrages. Parmi ceux-ci, le lecteur s'intéressera particulièrement à :

- « *Mécanique quantique, une introduction philosophique* », champs-Flammarion, 1997
- « *L'aveuglante proximité du réel* », champs-Flammarion, 1998
- « *Physique et Philosophie de l'Esprit* », Flammarion 2000

Par ailleurs, Michel BITBOL a participé au Congrès « *Physique et Conscience* » des 9 et 10 décembre 2005, au Carré des Sciences où il nous a présenté :

« *Une science de la conscience équitable: la neurophénoménologie de Francisco Varela* »

Michel BITBOL s'est intéressé de très près à la problématique de l'émergence vue sous plusieurs angles ainsi que son article « *La Nature est-elle un puits sans fond* » publié dans le dernier bulletin n°111, issu de « La Recherche » (n°405 –février 2007) , en témoigne.

Michel BITBOL évoque dans un premier temps l'aspect « ontologique » de l'émergence en se basant notamment sur les travaux de Robert LAUGHLIN dont les thèses sont anti-réductionnistes. Il existerait deux types émergence « ontologique », l'une « moniste » , l'autre « dualiste » dans laquelle les propriétés macroscopiques émergentes n'auraient pas de lien absolu avec les propriétés microscopiques...et existeraient « par elles-mêmes » (ontologie).

Le conférencier conteste quelque peu cette dernière vision purement ontologique¹. Il envisage la possibilité selon laquelle l'émergence de traits physiques autonomes pourrait tout simplement être liée à nos capacités perceptives et cognitives et n'avoir aucun fondement réel . De même : l'émergence de catégories linguistiques proviendrait simplement de concepts de bases trop grossiers, l'émergence computationnelle de domaines qui ne se laisseraient pas calculer à partir des éléments de base, ne serait imputable qu'aux insuffisances provisoires des moyens de calcul. Ceci pose le problème de l'émergence des relations cognitives.

III) « Introduction à l'émergence » par le Pr. Roger BALIAN

Le Pr. Roger BALIAN Polytechnicien- Ingénieur des Mines , a fait sa carrière au CEA d'abord au service de Physique Mathématique puis au service de Physique Théorique qu'il a dirigé jusqu'en 1987. Il a enseigné la Physique Statistique à l'Ecole Polytechnique. Il est membre de l'Académie des Sciences depuis 1995. En dehors de nombreuses publications scientifiques en Physique Statistique et en Mécanique Quantique, il est l'auteur de plusieurs ouvrages parmi lesquels :

- « *Le plaisir des mathématiques* », revue de Mathématiques Spéciales 116(3) 2006
- « *Le temps macroscopique* » sur l'irréversibilité et l'entropie (conférence donnée lors du premier colloque « Physique et interrogations fondamentales »
- « *Le temps et sa flèche* » Collection Champs de Flammarion

Roger BALIAN a également donné à l'Université de tous les savoirs une conférence « *Entropie, information : un concept protéiforme* ».

Roger BALIAN a établi un théorème sur les séries de Fourier, appelé Théorème de BALIAN-LOW.

¹ Seul le domaine de la mécanique quantique serait producteur d'une émergence non réductionniste. « L'état du tout y déterminerait celui des parties » à l'inverse des théories ascendantes dans lesquelles « l'état des parties déterminerait celui du tout ». (voir théorie fusionnelle de l'émergence de Paul HUMPHREYS)

Le Pr. Roger BALIAN n'adhère pas aux propositions à base ontologique de Robert LAUGHLIN. En ce qui concerne la mécanique quantique il insiste sur le fait que le concept fondamental est celui de matrice de densité et non celui de fonction d'onde. On ne dispose pas d'une théorie de l'objet-en soi mais on dispose de théories mathématiques (algèbres non commutatives) au formalisme adapté.

L'irréversibilité des propriétés macroscopiques simplement via les théories probabilistes émergerait de propriétés microscopiques réversibles et ne mériterait donc pas autant d'attention. Il ne faut cependant pas oublier que les probabilités ne sont qu'un outil mathématique.

Divers commentaires se présentent dans l'assistance :

- Le monde classique émerge du monde de la mécanique quantique
- La mesure n'exercerait-elle pas une contrainte sur le système ?

Selon, Roger BALIAN, l'irréversibilité résulterait de la perte d'information à l'échelle microscopique au profit du système d'ordre supérieur. Par ailleurs il existerait de très nombreuses causes d'émergence parmi lesquelles :

- des causes extérieures ou non
- des changements d'échelle
- l'irréversibilité émergerait de la réversibilité au niveau microscopique
- L'émergence dépendrait également des instruments de mesure dont on dispose
- En mécanique quantique on ne peut séparer l'objet observé de l'observateur, une mesure étant une fabrication dynamique de corrélation entre le pointeur de l'appareil et l'objet dont on ne mesure donc qu'un aspect.
- En mécanique quantique l'affirmation est contextuelle –l'affirmation en soi n'y existe pas en sorte que la logique absolue ne s'y applique pas- .. Peut-être peut-on parler de l'émergence d'une nouvelle logique dans laquelle on peut effectuer deux mesures, chacune étant vraie tout en étant contradictoire avec l'autre....

Après ce riche débat, la séance a été levée à 20 heures,

Bien amicalement à vous,

Irène HERPE-LITWIN

Compte-Rendus de la Section Nice-Côte d'Azur

Le savoir est le seul bien qui
s'accroisse à la partager.
Comprendre est bien sans limite qui
apporte une joie parfaite.
Baruch SPINOZA (1632-1677)

Compte-rendu de la séance du 18 janvier 2007 (101^{ème} séance)

Présents :

René Blanchet, Sonia Chakhoff, Pierre Couillet, Patrice Crossa-Raynaud, Guy Darcourt, René Dars, Jean-Paul Goux, Jacques Lebraty, Maurice Papo, Jacques Wolgensinger.

Excusés :

Jean Aubouin, Alain Bernard, Jean-Pierre Delmont, Yves Ignazy, Gérard Iooss, Michel Ladzunski, Jean-François Mattéi.

1- Approbation du compte-rendu de la 100^{ème} séance.

Le compte-rendu est approuvé à l'unanimité des présents.

2- Le mois écoulé.

Jacques Lebraty et Patrice Crossa-Raynaud font le point sur la préparation du manuscrit pour l'édition des actes du cycle de conférences : « *Entreprise et mondialisation* ».

Toutes les conférences ont été transcrites et revues. Certaines ont déjà été approuvées par leurs auteurs. Pour les autres, nous attendons qu'ils fassent le choix des illustrations.

Nous espérons que l'ouvrage sera prêt pour l'édition vers la fin février.

Actes du colloque sur le climat : tous les exposés ont été transcrits et enregistrés sur disquette. Ils seront envoyés pour approbation à chacun des auteurs.

3- Examen d'une proposition de cooptation.

Lors de la dernière séance le Président nous a distribué le *curriculum vitae* du professeur François Cuzin, membre de l'Institut, Chevalier de la Légion d'honneur, normalien, responsable de projet dans l'unité de recherche génétique du développement normal et pathologique de l'INSERM.

Les membres de l'AEIS ayant tous approuvé cette proposition de cooptation, elle sera présentée au professeur François Cuzin par le Président René Dars.

4- Prochains colloques.

- « *L'éthique et la morale* » pourrait faire l'objet d'un cycle de conférences. Il convient, pour le responsable, de rechercher les conférenciers et de fixer les dates au CUM.

- **Pierre Coulet** pense que l'on pourrait faire un cycle autour de « *Qu'est-ce que la Science* ». En fait, les idées ont évolué depuis l'Antiquité, d'une manière très continue. Pour la lumière, par exemple, la modélisation actuelle permet de mieux présenter le réel. Ne pourrait-on pas mettre ensemble la Science comme démythification des phénomènes réels et la modélisation comme outil des nouvelles technologies, qu'elles soient mathématiques ou autres ?

Quand les mathématiciens nous disent que leur discipline est jeune, cela semble incroyable car il y a eu Euclide bien avant et beaucoup d'autres. La seule différence, la rupture, vient de la technologie : les ordinateurs.

Il y a en ce moment en Science, l'émergence de nouvelles technologies, par exemple en biologie, qui permettent des progrès très rapides.

Est-ce que l'on ne pourrait pas présenter la science en parlant de démythification et de modélisation ?

- **Guy Darcourt** fait remarquer que du côté de la Science, on considère que le mythe est une erreur alors que du côté des sciences humaines c'est une vérité.

Que Sophocle ait écrit « *Œdipe roi* » prouve qu'il devinait quelque chose qu'il a fallu 20 siècles pour expliciter.

- **Pierre Coulet** rappelle que Léonard de Vinci avait déjà compris que les fossiles représentaient quelque chose qui démythifiait des croyances. Dès cette époque pour la lumière (arc en ciel) par exemple, on a remplacé les explications mythiques par des rationnelles.

Il y a une modélisation qui crée des savoirs.

- **René Dars** souligne qu'il ne faut pas confondre démythifier et démystifier.

- **Pierre Coulet** rappelle que les modèles sont conçus pour prédire l'avenir. Ils ont d'abord été très simples, élémentaires et si on ne les comprend pas, on ne peut pas aborder les plus complexes.

La technologie a commencé bien avant la Science. La première est destinée à réaliser quelque chose alors que la seconde vient du besoin de comprendre.

- **Jacques Lebraty** note que l'on s'éloigne ainsi de notre première idée de réaliser un cycle sur la modélisation. Il y a alors deux perspectives différentes : l'épistémologie se situe au niveau de la connaissance ; comment parvient-on à connaître la réalité. Dans ce cas, on peut parfaitement englober dans la modélisation les mythes, les arts proches et tout sujet touchant à la réalité. Ou bien on est sur le plan opérationnel : par exemple comment, avec la modélisation, on peut construire un schéma boursier.

En épistémologie, au sens de Varela, il s'agit d'explorer des processus cognitifs, au niveau opérationnel il s'agit de tester la puissance des outils.

- **Pierre Coulet** pense que c'est plutôt la première option qui est souhaitable : expliquer la Science avec des exemples, des conférences concrètes. C'est le cas de la lumière qui est un superbe exemple pour expliquer ce qu'est la Science. La ligne droite est venue de l'observation des rayons lumineux au travers des ouvertures.

Notre projet de faire un cycle sur la modélisation risquait d'être une tristesse infinie, alors que « *Les mythes et la Science* » est plus grand public. Il devrait être ouvert par un philosophe et continuer par un ensemble de disciplines scientifiques très différentes dont le but serait de faire disparaître les explications mythiques et tout ce qui est mineur dans les explications.

- **Patrice Crossa-Raynaud** rappelle la proposition qu'il avait faite sur les mythes et la Science :

- Le déluge expliqué par la fin de la glaciation quaternaire,
- L'Atlantide ou le rêve de l'âge d'or qui a induit le déterminisme de XIX^{ème} siècle et le développement des utopies (communisme – nazisme),
- Prométhée : l'interdiction de la connaissance reprise très longtemps par les églises et les dangers actuels de la connaissance (OGM, nucléaire, clonage reproductif),
- Osiris : prémices des transplantations d'organes ?
- Les oracles et les pseudo-sciences.

On peut trouver d'autres mythes (en excluant les archétypes comme Don Juan) et choisir.

★★

★

jeudi 15 février 2007
déjeuner célébrant la 100^{ème} séance
Lycée hôtelier (Arénas)
à 12 heures

Prochaine réunion
le jeudi 15 février 2007 à 17 heures
au siège
Palais Marie Christine
20 rue de France
06000 NICE

Documents

Pour préparer la conférence de Jacques DUBUCS nous soumettons à votre lecture un article intitulé « Le déploiement des capacités cognitives » (*Unfolding Cognitive Capacities*) qu'il nous a confié. Cet article, traduit en français, traite notamment de la validation d'une proposition de Leibniz en désaccord avec les théories modernes de la complexité :

P. 11 : « Le déploiement des capacités cognitives » par Jacques DUBUCS

Dans le cadre de la même problématique notre Collègue, Alain CARDON nous soumet son projet d'étude :

P. 18 « Principes d'architecture pour un système générant des pensées artificielles »

LE DEPLOIEMENT DES CAPACITES COGNITIVES

Jacques Dubucs

Institut d'Histoire et de Philosophie des Sciences et des Techniques (CNRS/Paris I/ENS)

13, rue du Four, 75006 Paris, France

Jacques.dubucs@cnrs-dir.fr

<http://www.ihpst.univ-paris1.fr/>

Résumé : En ce qui concerne les capacités cognitives, le point de vue de l'intelligence artificielle classique a souvent été l'objet d'un défi de la part des tenants du soi-disant point de vue de l'émergence. Cet article est une tentative de tracer une théorie générale des déploiements incompressibles fondée sur des considérations logiques en rapport avec la faisabilité - théorie qui est en phase avec un vieux postulat leibnizien plutôt qu'avec la théorie contemporaine de la complexité. Je défends une variante de l'émergentisme selon laquelle tout processus qui conduit à doter un système de capacités cognitives est un déploiement incompressible de même nature.

1 Le déploiement en tant que problème

Les postulats de base de l'intelligence artificielle ont été l'objet de nombreuses critiques récemment. Un grand nombre d'exemples bien connus, d'observations et *d'expériences de pensée* (« *Gedankenexperimente* ») ont établi qu'il ne suffit pas, pour avoir un comportement adéquat dans le monde réel, d'être équipé d'un quelconque ensemble complexe d'instructions et de les suivre aveuglément. Pour résumer ce point qui est devenu familier, jouir de capacités d'intelligence ou de cognition ne peut s'identifier à la simple exécution d'un programme d'ordinateur, car les capacités d'intelligence ou de cognition sont des propriétés *émergentes* qui ne peuvent que résulter d'un long processus d'« apprentissage » ou d'« adaptation » impliquant les interactions de nombreux composants simples.

Pour préciser cette idée vague, il est utile de se référer au parallélisme entre conception et compréhension. D'une part, la conception d'un dispositif capable d'avoir un certain **comportement ciblé** dans certaines circonstances revient à mettre au point et à réaliser un mécanisme complet capable de produire le résultat désiré en présence de ces circonstances ou en étant partiellement causé par elles. D'autre part, pour comprendre un système donné, on doit trouver le mécanisme qui explique son comportement en jetant la lumière sur les transitions qui mènent des conditions initiales du système à ses états futurs et qui permettent de les prédire. La tâche du concepteur est de mettre au point un mécanisme qui produise des résultats désirés et connus d'avance (c'est comme une explication rétrospective, sauf que les résultats à expliquer sont factuels dans le cas de l'explication, tandis qu'ils ne le sont pas dans le cas de la conception). La tâche de la compréhension prédictive consiste à découvrir le mode de fonctionnement du système donné afin de prévoir ses futures configurations inconnues jusqu'ici. Un problème crucial, autant pour la conception que pour la compréhension, dépend du répertoire de base dont on dispose. Pour réaliser un robot capable de se

déplacer dans une pièce encombrée de meubles nous n'avons pas la possibilité de recourir à la collaboration de « promeneurs adroits » ou d' « artefacts habiles » ; et nous devons plutôt nous contenter de composants omniprésents dépourvus d'intelligence tels que des capteurs ou des cellules photoélectriques. De même, nous n'avons pas les moyens, pour prédire les conséquences de l'ingestion d'opiacés, de faire appel à des entités complexes « empaquetées » telle que la *virtus dormitiva* ; nous devons énoncer nos explications dans la terminologie élémentaire de la neuropharmacologie.

Les gens qui ont travaillé dans la tradition de l'Intelligence Artificielle classique étaient convaincus que le fossé entre le comportement tout à fait intelligent et les dispositifs élémentaires dont on disposait serait aisément comblé par la planification. On divise le problème de l'obtention du comportement ciblé en une série gérable de sous-objectifs jusqu'à ce que les objectifs élémentaires soient atteints, chacun d'entre eux étant facilement réalisable par un dispositif de base (modulaire) agissant sous le contrôle d'un « administrateur » central. Le fait que cette conception de l'action planifiée soit généralement inadaptée et que les comportements ciblés soient réalisés plus efficacement d'une façon ascendante (« *bottom-up* »), en laissant les dispositifs très élémentaires et non spécifiques interagir librement les uns avec les autres et avec leur environnement est devenu un lieu commun. Mais on n'a pas remarqué que l'épistémologie de la compréhension prédictive rencontre un problème qui est exactement parallèle, nommément le fossé entre, d'une part, des corrélations grossières d'entrées/sorties qui se contentent de reprendre ce qui doit être expliqué sans valeur supplémentaire prédictive ou explicatoire et d'autre part, des lois de transitions élémentaires décrivant pas à pas l'évolution de la microstructure qui est impliquée. En d'autres termes, le problème central du design, notamment le problème de la décomposition du comportement ciblé de l'ensemble en micro-opérations devant être réalisées par ses parties, fait écho au problème dual de l'extraction des prédictions concernant le futur du macro-système à partir de la connaissance des micro-transitions de ses parties. Le but de cet article est d'expliquer comment le second problème peut être résolu et comment cette résolution peut être appliquée au problème du design de dispositifs munis de capacités cognitives.

Soit S la classe des états possibles d'un macro-système, G la classe des *applications* désirables de S dans S (bien sûr, non désirables dans l'absolu, mais considérées comme telles par le concepteur. Pour une *application particulière* ciblée, disons F , le problème du concepteur est de trouver des applications f_1, \dots, f_n qui représentent autant de transitions élémentaires disponibles d'un état d'un composant au suivant et qui soient capables, lorsqu'elles sont composées ensemble, de donner F . Inversement, le problème de la prédiction réside dans le calcul de l'application globale résultant des applications et compositions des f_i élémentaires, qui sont supposées être déjà connues. On peut se représenter le processus par lequel les transitions élémentaires f_i révèlent leur effet jointif global, notamment les applications associées F (ou la valeur de cette application aux conditions initiales) comme un processus de déploiement . Il n'existe pas plus , dans un certain sens, d'information dans l'application F que dans le dossier $\{ f_1, \dots, f_n \}$ dont provient F . Cette information est tout simplement contenue sous forme compressée dans le dossier . La conception consiste à compresser, tandis que la prédiction consiste à décompresser.

Primo, notons tout ce que la notion de déploiement recouvre. La preuve dans un système formel peut-être vue comme le processus de déploiement du contenu mathématique des axiomes au moyen de l'application graduelle des règles d'inférences. La mise en œuvre d'un programme informatique peut être vue comme le déploiement du contenu qui est implicite dans ses instructions.

La croissance des feuilles à partir des bourgeons est aussi, au sens propre, un déploiement tout comme le développement de toute la plante à partir de la graine.

Secundo, on doit distinguer entre le déploiement, nommément le développement réel d'un objet à partir des prémisses correspondantes, et le déploiement épistémique, nommément le processus intellectuel qui conduit à la connaissance du résultat à partir de la connaissance des objets de départ. Les deux processus sont souvent strictement parallèles. La connaissance mathématique, si elle est développée selon les recommandations de Bolzano, notamment de façon à ce que les théorèmes ne soient démontrés en n'utilisant que les propositions auxquelles ils doivent leur vérité (*wissenschaftliche Darstellung*), est un exemple fameux de ce parallélisme.² Dans les autres cas les choses sont différentes, elles dépendent des capacités cognitives de ceux qui savent. Selon Leibniz, la vie d'Adam est un exemple possible de discordance. Bien sûr, Adam a vécu d'une manière exactement conforme avec ce que Dieu avait projeté en le créant, nommément en décidant de réaliser cet Adam possible et en le faisant parvenir à l'existence. Pour nous-mêmes, en tant que créatures pourvues d'un pouvoir d'analyse fini, limité, le seul moyen de déployer la notion d'Adam consiste à attendre la vie réelle d'Adam et à observer ce qu'il fait et ce qui lui arrive. Même si la notion ou le « dossier » d'Adam contient implicitement tout ce qui lui arrive, y compris à ses parents et à sa descendance, il nous faut attendre le déploiement progressif de ce qui est implicite dans cette notion. En d'autres termes le déploiement est *incompressible* pour nous et il n'y a pas de raccourci pour connaître ce qui se passe avant que cela ne se passe. La situation épistémique de Dieu est tout à fait différente, parce qu'Il est doté d'un pouvoir d'analyse infini : Pour Lui, il n'est pas réellement question de déploiement puisqu'Il voit tout simultanément (« *toto simul* »).

Une autre façon de formuler à nouveau l'énoncé de Leibniz dans un contexte plus vaste, non théologique consiste à dire qu'il existe des objets tels que la connaissance exhaustive de tous les objets dont ils sont issus, en concomitance avec celle de toutes les lois gouvernant leur composition et leurs interactions, n'est pas convertible en la connaissance anticipatoire de ces mêmes objets: il n'existe pas d'autres moyens, pour connaître le résultat final, que d'attendre que ce résultat apparaisse *réellement*.

D'autre part, la théorie contemporaine de la complexité suppose que la complexité du produit déployé réside entièrement dans la complexité de son dossier comprimé et même que la taille de ce dossier puisse-t-êtr prise comme mesure de la complexité du résultat. Par exemple, conformément à l'idée fructueuse de Kolmogoroff, la complexité d'une suite de 0 et de 1 successifs peut être mesurée par la longueur du programme le plus court nécessaire pour la générer. Ainsi, seule la complexité du dossier a de l'importance, et aucune signifiante n'est conférée à la complexité du processus de déploiement lui-même : connaître le dossier revient à connaître l'objet déployé qui est contenu en lui sous une forme embryonnaire, parce que le déploiement ne transporte ni ne requiert jamais d'information nouvelle.

Pour discuter la justesse de cet abandon complet de la position leibnizienne, restons en marge des problèmes ennuyeux qui, le plus souvent, ternissent l'argumentation, notamment ceux qui sont liés à l'indéterminisme et à l'incalculabilité. Je prendrai pour hypothèse que ces problèmes sont absents, c'est-à-dire que nous possédons une connaissance analytique juste et complète des lois fondamentales qui gouvernent les interactions entre les constituants du système dont nous parlons, et que cette connaissance est entièrement convertible (peut-être par le biais d'une discrétisation

² Cf [Dubucs-Lapointe 2003] et [Dubucs-Lapointe *in press*]

convenable des équations d'évolution) en une connaissance algorithmique. Alors, le problème principal consiste à savoir s'il existe une place pour l'idée Leibnizienne de déploiement incompressible. J'argumenterai, contrairement à l'*opinion commune* en faveur d'une réponse positive à cette question.

2 Déploiement , Prédiction, Hyper-prédiction

Pour éclaircir l'argumentation, il me faut prendre un exemple très simple de déploiement de Kolmogoroff . Le programme :

```

      1          n=0
0     2          Print n
      3          n :=n+1 mod
0     4          Goto 20
0

```

engendre lors de son exécution , la suite infinie 01010101....

On peut considérer cette suite de digits comme les états successifs d'un objet x changeant et de valeur initiale zéro. Le programme est le dossier de l'objet et le déroulement du programme , y compris le processus d'impression, est son déploiement diachronique. Une caractéristique intéressante de ce déploiement réside dans le fait que des raccourcis sont possibles, ce rend la prédiction possible. Clairement, la valeur de x en son $n^{\text{ième}}$ état est 0 si n est pair et 1 dans le cas contraire. Ainsi, pouvons nous sauter mentalement à n'importe quel état de x pour connaître la valeur de x à cet état.

La notion de prédiction peut se comprendre en deux sens différents. Au sens général, prédire (avec exactitude) un événement revient simplement à savoir que l'événement se produira, avant qu'il n'ait lieu. La prédiction dans ce sens est généralement possible dès que les exigences d'exactitude, de complétude et de calculabilité énoncées ci-dessus sont satisfaites. Parce que dans ce cas il suffit de disposer d'ordinateurs suffisamment rapides pour obtenir les résultats avant que l'événement ne se produise. Les avancées dans la technologie informatique assurent que cette catégorie générale de prédiction peut ou pourrait être réalisée dans la plupart des cas, à condition que l'on dispose de connaissances qui satisfassent les trois exigences précisément énoncées. Dans ce cas, la proposition de Leibniz selon laquelle le déploiement est parfois définitivement incompressible et selon laquelle le mieux que nous ayons à faire est d'attendre pour ainsi dire la développement de la réalité, semble pouvoir être réfutée dans un monde déterministe et calculable. Néanmoins cette proposition peut être confortée d'une autre façon, nommément en se référant à un sens plus fort de la notion même de prédiction.

La réfutation qui a été justement développée pour la proposition de Leibniz repose sur l'idée communément admise que le progrès de rapidité des instruments de calcul est illimité, et qu'alors le déploiement par l'ordinateur est en principe capable d'aller plus rapidement que le déploiement des choses elles-mêmes. Mais un tel déploiement par ordinateur suit exactement la même piste que le processus réel lui-même. Le raccourcissement n'est pas obtenu en prenant un raccourci du chemin réel, mais en déroulant plus rapidement le même processus. Il n'est nullement question de sauter à la fin du processus, mais il s'agit uniquement de l'exécuter pas à pas à une vitesse plus élevée. C'est pourquoi, au mieux, la valeur de la proposition de Leibniz est affaiblie, mais elle n'est pas réfutée à strictement parler. Il serait utile de la maintenir en lui ajoutant quelques caractéristiques en affirmant, par exemple, qu'il existe des phénomènes tels que le seul moyen de les prédire est de simuler au moyen d'instruments informatiques le processus de déploiement qui leur donne naissance. Pour vraiment saper la proposition de Leibniz, il en faut beaucoup plus, nommément il faut un argument convaincant établissant que la prédiction est toujours possible avec un sens plus puissant : pas au moyen d'une simulation rapide du processus réel en déploiement, mais en court-circuitant certaines étapes et en bondissant plus rapidement jusqu'à la fin du processus. Appelons *prédiction forte* ou *hyper-prédiction* ce type de raccourci. Il existe beaucoup de cas où la prédiction peut être convertie en hyper-prédiction. L'exemple ci-dessus en est un : on peut court-circuiter le processus de déploiement de x pour obtenir directement sa valeur à n'importe quel stade de son déploiement. Si, et seulement si, cette manœuvre est toujours possible alors la valeur de la proposition de Leibniz sera remise en question. C'est pourquoi, les défenseurs de la proposition doivent montrer que l'hyper-prédiction est en général impossible, même dans un univers déterministe et calculable.

3 Perspectives de Validation de la Proposition de Leibniz

Premièrement, un contenu mathématique bien arrêté peut certainement être attribué à la revendication selon laquelle on ne dispose pas toujours d'une hyper-prédiction. Ce contenu mathématique pourrait être explicité en définissant une classe L non vide de fonctions récursives qui sont seulement calculables selon un processus pas à pas (ou pour lequel ceci est la voie de calcul la plus rapide) ; notamment, f appartient à L si le (meilleur) calcul de $f(n)$ nécessite le calcul successif de $f(0), f(1), \dots, f(n-1)$. Par manque de place, je ne peux pas entrer ici dans les détails de la progression dans cette direction. Disons simplement, pour donner une idée de la difficulté de la tâche, que quelques « occupants naturels » de la classe L sont inadéquats. Par exemple, on pourrait penser à la suite des décimales de π en tant que candidat naturel pour appartenir à L parce qu'il est de prime abord clair que la valeur de la $n^{\text{ième}}$ décimale de π ne peut être calculée avant que la valeur des décimales précédentes aient été calculées. Intuitivement, les décimales de π doivent être calculées l'une après l'autre et on n'arrive pas à percevoir qu'il serait possible de calculer la valeur de, disons, la millième décimale de π sans avoir à disposition le segment initial 141 592... Malheureusement, un fameux résultat de calcul mathématique obtenu par Bailey, Borwein et Plouffe³ a montré que cette intuition était erronée.

Ainsi, au lieu d'entrer dans les détails de la définition mathématique de L , donnons quelque idée générale de la (mauvaise) raison pour laquelle la proposition de Leibniz est tombée en disgrâce. Cette situation vient essentiellement de la prédominance du point de vue dit « de principe » qui a été défendu depuis le travail de Turing et d'autres logiciens dans les années 30. L'analyse de la notion

³ [Bailey-Borwein-Plouffe 2004]

de calculabilité algorithmique qu'ils ont fournie connaissait un tel succès qu'elle a été prise comme la référence ultime pour caractériser les capacités de calcul - elle a été considérée comme un « miracle », car la notion qui en résultait était absolue, notamment robuste et indépendante de toute référence à un système formel particulier -. Maintenant, il est simplement clair, que pour une machine de Turing, il ne peut pas exister de différence sensible entre prédiction et hyper-prédiction. La seule question d'importance pour une machine de Turing est de savoir si elle est en principe capable de calculer quelque résultat et toute question allant au-delà, concernant le mode ou la vitesse du calcul, est dépourvue de signification. Ce point de vue est aussi sous-jacent à la théorie de la complexité de Kolmogoroff qui mesure la complexité d'une suite par la taille du plus court programme d'ordinateur capable de la produire, quelle que soit la durée de cette production. : la seule chose qui importe est le dossier initial et non le processus de déploiement. Pour étayer l'idée de Leibniz, nous devons donc mettre en jeu des considérations relatives à la faisabilité pratique ou humaine, c'est-à-dire prendre en compte la frontière entre ces processus qui non seulement peuvent être menés à bout *en principe*, mais peuvent aussi être réalisés humainement et ceux qui, bien qu'ils puissent être menés jusqu'au bout *en principe*, ne peuvent être réalisés humainement⁴. Si on adopte cette perspective, il devient évident que les seules prédictions que nous pouvons réellement faire sont celles pour lesquelles nous pouvons court-circuiter certaines parties du développement du système approprié. En ce qui concerne les autres, il n'existe aucun autre moyen, comme le disait ou aurait dû le dire Leibniz, que d'attendre le résultat du déploiement ou de le simuler.

En retournant au problème de départ, la conjecture la plus plausible concernant les capacités cognitives est que, lorsqu'un système en est doté, elles soient non seulement *nouvelles*, au sens qu'il est absurde de les attribuer à des composants spécifiques du système, mais également *anormales* au sens qu'il est impossible d'expliquer pourquoi le système possède ces capacités ou d'hyper-prédire qu'il les aura en se basant simplement sur les lois qui régissent le comportement de ses composants.

⁴ Cf [Dubucs 2002]

Références

[Bailey-Borwein-Plouffe 1995] D.Bailey, P. Borwein & S. Plouffe, “A rapid computation of Various Polyalgorithmic Constants » in Pi: A source Book(eds. L. Berggren, J. Borwein & P. Borwein), Springer Verlag, 3th edition , 2004

[Dubucs 2002] J. Dubucs, Feasibility in Logic, Synthese 132 (2002), pp. 213-237

[Dubucs-Lapointe 2003] J. Dubucs et S.Lapointe, Preuves par excellence, Philosophiques 30(2002), pp 219-234

[Dubucs-Lapointe *sous presse*] J. Dubucs et S.Lapointe, Bolzano’s Alleged Explicativism, Synthese *sous presse*

Principes d'architecture pour un système générant des pensées artificielles

Alain Cardon

Février 2007 Document de travail.

Bien poser le problème

Une pensée est, sous son aspect physique, une forme **émergente** dans un organe particulier qui n'est pas constitué de couches fonctionnelles, qui n'est pas fait de composants aux liens prédéfinis et aux rôles fonctionnels immuables, formant des structures régulières. Le cerveau ne produit pas de la pensée comme une sorte d'usine compliquée qui produirait des états dits de pensée, après un processus de fabrication bien déterminé. Il s'agit donc d'abandonner, pour la génération de pensées, la notion de programme général sous forme d'un automate à états, constitué lui-même de multiples automates dans une ramification très profonde, et qui ne ferait finalement qu'atteindre des états prédéfinis. Il s'agit de s'engager vers des systèmes qui opèrent par réorganisations incessantes des très nombreuses entités qui les composent et qui construisent sans cesse, avec ces entités et selon de multiples contraintes opérant par négociations, des *configurations* au sens géométrique du terme, valant pour des états morphologiques plutôt que fonctionnels. Ces états ont de la signification, c'est-à-dire qu'ils sont systématiquement associés à la désignation clairement exprimée de quelque chose. Il y aura toujours, dans ces systèmes, l'expression de permanences et de variations, et donc, en terme géométrique, attraction vers des formes stables et usuelles ou bifurcation vers des formes instables et nouvelles. Ce sont des systèmes qualifiés de complexes au sens organisationnel du terme.

Il s'agit de trouver des architectures très dynamiques, qui s'activent en subissant des tendances comme dans le psychisme humain, qui s'activent sous la poussée de tendances fondamentales valant pour des pulsions, et résolvant les contradictions de tendances. Et il s'agit finalement de trouver comment un état émergent formé de multiples entités co-actives, concurrentes ou associées, peut être poussé à l'expression et ressenti par le système, de la même façon qu'une pensée est produite, appréciée et ressentie par l'esprit qui la génère. Sinon, cette pensée n'émerge pas et reste inconsciente, latente dans l'inconscient, et elle peut tout au plus engager à une action réflexe ou altérer d'autres productions suivantes. Il faut donc très fortement s'inspirer de modèles du système psychique, et non de ceux de la mécanique de l'action et de la réaction.

On dispose d'un modèle explicatif du principe de fonctionnement de l'appareil psychique, découvert il y a un siècle par l'inventeur de la psychanalyse, le Dr. Sigmund Freud. Ce modèle proposait l'architecture générale d'un système de production de pensées avec ses contraintes et ses défaillances possibles, c'est-à-dire ses pathologies.

Le modèle a été défini à partir de ses effets visibles dans le comportement humain. La localisation physique effective des constituants d'un tel modèle est l'objet de recherches intenses en neurobiologie. Et il n'y a pas eu vraiment de rencontre entre ce modèle conceptuel, fort complexe, et l'informatique très technicienne et focalisée sur la programmation réactive et fonctionnelle, où l'on spécifie à l'avance tous les cas de fonctionnement des systèmes. La transposition n'a pas vraiment été envisagée, sans doute parce que les concepts informatiques n'étaient pas assez développés et les concepts psychanalytiques paraissaient trop abstraits ou trop vagues. Aujourd'hui, la transposition peut se faire, par synthèse constructiviste, en réinterprétant les topiques freudiennes.

Notre thèse, centrale, ambitieuse, sera qu'un système psychique complet peut être développé sur un autre support que le réseau neuronal d'un cerveau, et notamment sur le support fourni par certaines entités informatiques en cours de calcul, constituant une vaste organisation dynamique, co-active et s'auto-contrôlant en ajustant sans cesse ce contrôle.

L'enjeu de ces recherches est clairement la **conscience artificielle**. Il s'agit de concevoir des systèmes très autonomes, qui génèrent des pensées artificielles sous la forme de configurations dynamiques de nuées de processus structurées géométriquement, qui développent et éprouvent ces pensées artificielles en étant en liaison temps réel avec une corporéité, en s'appuyant sur des émotions également artificielles. Il s'agit de concevoir des systèmes qui manifestent de l'intérêt, du souci, ici et maintenant pour certaines choses qu'ils remarquent selon leurs sens artificiels ou leurs rappels mémoire, qui ont des intentions propres, et qui ont un certain profil psychologique.

Un système qui génère des pensées artificielles aura donc des pulsions, un inconscient, un préconscient et un processus lui permettant d'éprouver certaines de ses productions : la sensation de penser à quelque chose. Il pourra être conscient de ce qu'il engendre comme formes dynamiques organisées, au sens où il pourra les observer, les manipuler, s'en servir pour agir par son corps et en générer d'autres, et puis les mémoriser pour en conserver une certaine trace. Il aura un vécu artificiel, radicalement différent d'un ensemble d'informations factuelles structuré comme le sont les bases de données, mais lui permettant de rappeler des quantités d'événements artificiels valant pour une vie passée, pour les expériences nécessaires à ses aptitudes courantes. Il augmentera cette mémoire par des événements réels qu'il vivra d'une certaine façon en les produisant et en les ressentant. La création de ce vécu artificiel conditionnera finement son profil psychologique, et cette construction n'est pas une affaire simple, c'est même un grand travail pluridisciplinaire.

Il faut donc, en informatique, découvrir une architecture dynamique produisant des émergences valant pour de la signification intentionnellement produite. Cette architecture est liée à des mots et à des structures, à des informations indicatrices de mouvements organisationnels et de valeurs. Il faut trouver les raisons qui seraient propres à la mise en activation, continue mais changeante, de ce système et qui ne seraient pas des commandes. Il faut trouver les éléments de base de l'architecture, il faut trouver l'architecture utilisant et contrôlant ces agents, et qui ne pourra être que géométrique. Il faut aussi trouver comment faire éprouver un état émergent par le système lui-même, en définissant une sensation à éprouver des représentations, avec production d'émotions artificielles.

Ce dernier point est la grande question, qui était en effet bien difficile : comment produire, dans un système artificiel, la sensation d'éprouver une représentation que le système générerait de lui-même ? La façon de résoudre un tel problème, a priori non soluble quand on l'énonce si brutalement, consiste à bien le poser, c'est-à-dire à le placer dans un domaine tel que l'on puisse

prendre en compte d'autres éléments fondamentaux, et qu'il devienne ainsi soluble en étant un problème local lié à d'autres qui sont bien solubles et forment ensemble le domaine de la solution. La notion de sensation de penser à quelque chose n'est pas première ni isolée dans la compréhension d'un système psychique, mais se comprend à partir des propriétés organisationnelles d'une certaine architecture qui a pour finalité de produire, mémoriser et éprouver des représentations. Et en définissant certaines propriétés de cette architecture, en distinguant bien la notion de processus qui s'exécute de celle de forme dynamique qui se donne à voir, les représentations peuvent effectivement être éprouvées, la notion de sensation à penser devenant alors claire, dans un sens bien constructible.

Il faut trouver comment faire s'activer ensemble, c'est-à-dire faire se co-activer de la bonne façon, des nuées de petits programmes indépendants mais qui communiquent et qui coopèrent, ce que l'on appelle des agents logiciels, et qui seront les entités de base du système, qui constitueront le substrat. Il faut trouver ce que vont représenter ces entités au niveau cognitif ou simplement opérationnel, comment ces entités peuvent former des agrégations plus significatives, qui s'exécutent comme telles et se coactivent, et qui vont faire se distinguer une certaine forme agrégative qui aura de la signification : la pensée artificielle courante qui sera éprouvée, puis ensuite mémorisée, et cèdera la place à la suivante. Il faut trouver comment ces agrégations vont produire un état unifié et significatif, normalement cohérent, émergeant au-dessus de multiples activités non exprimées, non ressenties par le système et peu cohérentes entre elles, il faut trouver comment un état très adapté à ce qu'il était intéressant de produire à tel moment et dans telle situation peut se définir et s'apprécier en se construisant entièrement. Et il faut trouver comment toutes ces idées artificielles forment une suite cohérente, adaptée à la situation et aux possibilités ontologiques du système.

Pour cela, il faut adopter un principe de construction, qui constitue bien une **rupture épistémologique** avec la démarche rationnelle et technicienne usuelle, et qui doit amener à considérer des systèmes informatiques sans objets fonctionnels permanents prédéfinis, à aucune échelle. Nous devons considérer des systèmes où tous les objets qui le composent sont d'une certaine manière évolutifs, en allant des éléments de base aux grandes agrégations relationnelles et même aux tendances. C'est un principe fondamental du vivant, à partir du niveau cellulaire. Chacun d'entre nous est un peu différent de ce qu'il était auparavant, parce qu'il s'est physiquement un peu modifié mais surtout parce qu'il a modifié sa mémoire en se servant de son esprit qui a fait plus ou moins changer cette mémoire et ses tensions. Cette propriété de l'effet du temps ne se simule pas et il faudra bien l'introduire effectivement dans ces nouveaux systèmes.

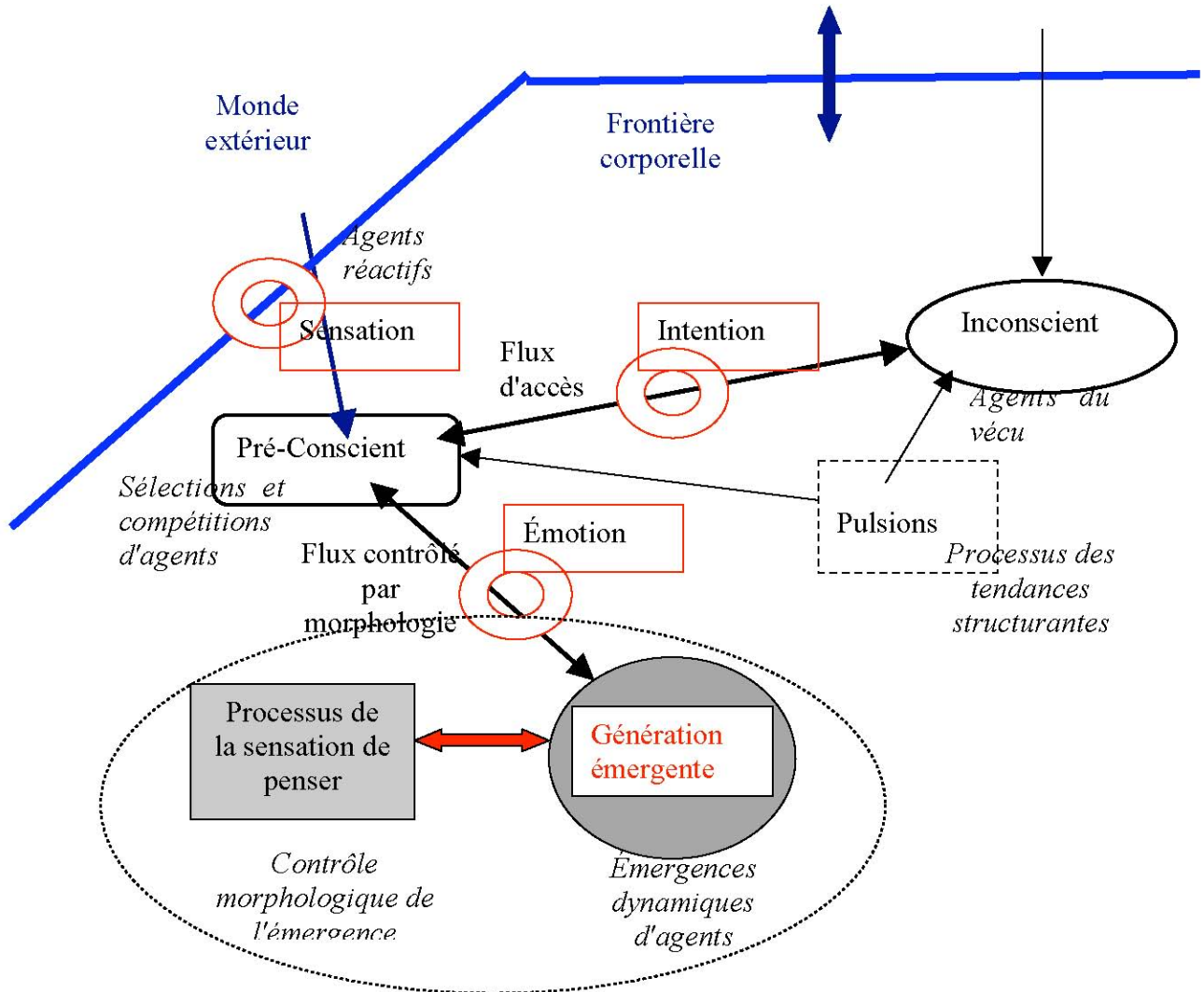
Il faut quitter la voie, très empruntée, où l'on s'intéresse à des programmes produisant des résultats attendus, définis à partir de fonctionnalités et de spécifications fixes, à des programmes qui affichent des informations au format de lecture des utilisateurs. Il faudra se focaliser sur des programmes qui génèrent à chaque fois et sans cesse des constructions nouvelles mais tenant compte des conformations passées, complexes, très dynamiques, leur servant à agir et réagir en temps réel et se positionner dans leur environnement. Chaque construction courante sert aussi et surtout à produire les constructions suivantes dans un processus que l'on pourrait dire sans fin, si la mort n'existait pas. L'esprit de toute personne qui prononce la phrase "gagner sa vie" n'utilise pas une sélection dans une bibliothèque la contenant déjà pour l'afficher, mais la construit à chaque fois qu'elle est exprimée, avec des nuances valant pour ce qu'elle signifie ici et maintenant, et avec des sensations et des jugements. Cela permet de prendre en compte le fait que l'on puisse créer de nouvelles idées, prononcer de nouvelles phrases, inventer des associations de concepts par analogies, approfondir ou généraliser des concepts ou des représentations.

Mais il faut aussi garder un principe économique : ce nouveau système est intégrateur des systèmes classiques, qui calculent des fonctions très précises et très locales. Il intègre, en les considérant comme des composants spécialisés, tous les systèmes spécifiques fonctionnels calculant des valeurs de fonctions, conçus dans tous les domaines depuis la naissance de l'informatique, et notamment de l'intelligence artificielle.

Pour un tel système qui doit donc générer des représentations pour son compte, on pourra vraiment dire que le tout, qui est mouvement organisationnel, est plus que la somme de ses parties, qui ne sont que locales. Le tout est un ensemble très dynamique qui ne cesse de changer d'organisation par le réagencement continu et la modification de ses innombrables parties, qui sont bien les parties d'un ensemble de composants élémentaires au sens mathématique d'un dénombrement immense. Concevoir un tel système revient à emprunter, et radicalement, la voie constructiviste de la production d'états considérés comme des agencements d'activités, c'est-à-dire qui sont appréciés comme des morphologies et non comme des suites de valeurs, ce que font les systèmes actuels. Il s'agit de concevoir un organisme artificiel, avec ses propriétés organisationnelles, toujours redéfinies pour qu'elles existent.

De tels systèmes sont aujourd'hui modélisés dans quelques très rares laboratoires de recherche spécialisés, certains ne produisant pas de résultats publics. Ils sont conçus dans le détail de leur architecture logicielle, ce que nous avons fait en ce qui nous concerne, et l'on en est à la phase de début de réalisation. La transposition de la pensée dans l'artificiel est donc en marche. Faut-il aller jusqu'à la réalisation et la mise en service ? Qui peut et doit répondre à une telle interrogation ?

Actions corporelles



Précisions conceptuelles

La conception d'un système générateur de faits de conscience est une transposition, une transposition radicale de l'architecture qui le permet dans le vivant dans le domaine des systèmes dynamiques calculables effectivement. Il s'agit de comprendre précisément tous les caractères essentiels de la pensée qui se génère dans les cerveaux et de dégager les propriétés spécifiques de cette dynamique, les régularités, les contraintes, les attractions, les évitements, les bifurcations, les lois d'agrégation et de séparation qui font cette émergence organisée continue. Ensuite, il s'agit de transposer cela dans l'univers informatique de la co-activation auto-contrôlée de nuées de processus capables de s'auto-organiser de certaines manières, en utilisant un fonctionnement lié de manière temps-réel à une corporéité artificielle, telle celle d'un ou de plusieurs robots. Il s'agit, pour construire un tel système, de résoudre le problème de la relation entre forme dynamique et

signification d'une désignation introduite par la forme. Ce problème est difficile, et il est de plus systématiquement non traité par les scientifiques, qui s'intéressent plutôt à la technologie lucrative et aux effets observables de l'acte ou de l'action qu'aux significations profondes des mouvements.

Nous allons poser maintenant l'**hypothèse centrale** qui nous a permis d'aller vers cette transposition :

• *Une pensée qui se fait doit être comprise comme un mouvement organisé constitué de processus se développant sur une matière très dynamique composée d'éléments proactifs, se déployant simultanément à plusieurs échelles de structures conceptuelles et d'organisations sous des contraintes multiples, le substrat dynamique et organisateur produisant des émergences sous la forme d'états dynamiques entrant en résonance à certains instants, là où la pensée s'exprime et se ressent.*

Précisons les termes que nous avons utilisés dans la formulation de cette hypothèse de caractère architectural :

• Un *processus* est un mouvement non aléatoire d'éléments matériels ou informationnels identifiés, confrontant une autonomie d'action et de changement effectif dans un milieu qui a ses contraintes. On pourra penser à des mouvements de convection des liquides ou à des molécules dans un gaz. Un processus s'oppose radicalement à la structure de la matière inerte, comme un cristal qui est une répartition régulière et invariante d'atomes.

• Un élément *proactif* est un élément identifié, structuré, s'activant comme un processus local, manipulant de l'information et de l'énergie, qui peut agir pour son compte selon des buts réactifs précis, et qui a ainsi une certaine autonomie. C'est le cas d'un neurone et évidemment d'un processus informatique. Par contre, une molécule n'est en rien proactive.

• Une *structure* est un substrat architectural formant un composant physique ou informationnel vaste à partir d'éléments de base, et qui a de la permanence. Une structure fixe et fige pendant une certaine durée des contraintes locales sur ses composants pour une activité qui se fera par la suite, au niveau de la structure.

• Ce peut par exemple être un réseau d'éléments informationnels permettant la circulation et la modification d'informations véhiculées entre ses éléments.

• Une *organisation* est une activité de multiples éléments de différents niveaux, appartenant à des structures, les éléments activés et actifs ayant un certain but local de par leur proactivité naturelle ou acquise, et le tout allant vers un certain état global. L'organisation est un déploiement qui a un sens, qui est, d'une certaine manière, dirigé et qui n'est pas aléatoire. Une armée est organisée lorsqu'elle n'est pas en déroute. Par contre, les mouvements des liquides ne sont pas des mouvements globalement organisés, car un liquide n'a pas de but général dégagé du mouvement d'ensemble de ses éléments, il est simplement soumis à des champs de forces, de gravité, de convection et de température qui agissent sur chaque élément et qui forment des tourbillons changeants, qui eux acquièrent une certaine organisation. Cette organisation sera décrite par des équations célèbres en physique, les équations différentielles non linéaires de Navier Stokes.

• *L'échelle du déploiement* d'une organisation à partir d'une structure initiale est la considération des caractères qui permettent de comprendre que le mouvement des éléments est dirigé par des règles ou des lois à un niveau qui est à la fois celui des éléments et aussi au-dessus de ces éléments. Le système s'active dans un déploiement lorsqu'il suit des règles, obéit à des lois locales et plus globales multiples, où des choix doivent se faire. Par exemple, tout système d'analyse monétaire

suit des lois qui permettent de dire que le système n'a pas fonctionné seulement de manière fortuite, par le seul hasard, mais qu'il se fonde sur des concepts précis valant à une certaine échelle, dont celle des systèmes psychiques des intervenants, et que son organisation a des régularités et montre des tendances.

- Les systèmes qui sont basés sur des mouvements d'éléments et non des éléments fixés, qui ont des mouvements définissables à différentes échelles, qui sont organisés, qui définissent et suivent des contraintes à plusieurs échelles spatiales et temporelles simultanément, sont qualifiés de **complexes**. Ce sont des systèmes qui sont particulièrement difficiles à étudier avec les techniques actuelles, car ils ne sont que variation et changement à différentes échelles d'espace et de temps. Par exemple, un écosystème comme une forêt tropicale est un système organisé, où les multiples éléments végétaux et animaux s'activent et se confrontent à différentes échelles de temps et d'espaces, étant en interaction continue allant du niveau individuel très local à celui du biotope. Il n'y a aucun équilibre permanent dans un écosystème, mais une remise en cause continue et négociée de toute stabilité, qui n'est que relative.

La pensée est un système particulièrement complexe car elle produit des représentations permettant de passer systématiquement du niveau information au niveau signification, et dans les deux sens. Il nous est bien impossible de considérer une parole entendue comme une simple suite de sons et non comme du sens sur quelque chose qui apparaît. Le modèle générant des émergences ayant de la signification sera donc celui qui produit des *représentations internes* ayant du sens, à partir d'informations factuelles nécessaires : c'est bien un modèle absolument constructiviste. La forme de ces représentations était à trouver, les effets, la raison de leurs productions ou altérations aussi.

Le fait pour un système complexe, qui a une certaine architecture, de produire des émergences et d'entrer en résonance organisationnelle avec lui-même pour s'éprouver tel il se conforme, est un point dur et délicat. Il est dur, car le modèle pour atteindre sa compréhension dans une approche calculable n'est vraiment pas évidente à découvrir. Il est délicat, car il ouvre directement sur la conscience artificielle, dont les applications engagent à une autre conception du monde et à une autre application du contrôle des personnes et des choses. Je ne donnerai donc pour l'instant aucune indication précise de manière publique sur ce point.

L'esprit qui produit de la pensée est un cas typique de système très complexe, mais ce n'est qu'un cas de système très complexe. Notre approche, clairement constructiviste, présente un avantage décisif quant à la compréhension de l'esprit par rapport à une approche descriptive basée sur la seule observation proche ou médiatisée par le langage. Elle permet de distinguer entre des éléments constitutifs qui sont et ne sont que sous deux formes : elle permet de distinguer ce qui est permanent et qui vaut pour des éléments structurels disponibles, de ce qui est processus, c'est-à-dire de ce qui est essentiellement du mouvement d'activation. Usuellement, les scientifiques non constructivistes, très nombreux, ne font pas cette distinction et identifient tout caractère du réel par des variables utilisées dans des équations, dans un espace adapté.

Notre démarche, à partir de la considération phénoménologique suffisante des caractères du processus psychologique qui conduit l'esprit à générer des pensées, sera de transposer ce type de fonctionnement dans le calculable. Plus précisément, les questions de base à poser et les problèmes à résoudre sont les suivantes :

- Quel nouveau substrat informatique considérer ?
- Quelles nouvelles formes dynamiques considérer ?
- Quelles tendances doit-on introduire dans le système, et sous quelle forme ?
- Quelles combinaisons de formes peut-on déployer ?
- Quel lien établir entre ces nouvelles formes et la notion de signification ?
- Quelle représentation le système pourra-t-il produire ici et maintenant, dans tel ou tel contexte, pour quelle motivation ?
 - Quelles causes donner à la mise en activité d'un processus de reconfiguration produisant une suite d'émergences et quand décider qu'un état est effectivement produit ?
 - Comment faire s'éprouver par le système un état qu'il produit et valant pour une représentation de quelque chose qu'il observe comme une forme interne ?
 - Comment procéder à la mémorisation de représentations d'événements valant pour un vécu artificiel et qui ne soit pas une bibliothèque de données ?

Ces questions sont les questions à résoudre, en y répondant de manière absolument précise, pour aller à un système réalisable et montrer que la transposition est possible.

Et puis, il y a un autre type de questions à poser, une fois que les réponses aux questions précédentes sont satisfaisantes :

- Comment exprimer la qualité d'une émotion artificielle et la lier à une représentation conceptuelle ?
 - Comment exprimer l'association permettant la création de représentations originales, imprévues, et pourtant pertinentes ?
 - Comment faire fusionner des émotions et des représentations conceptuelles entre de nombreux systèmes en communication ?
 - Quelle est la notion de Soi distribué multi-corps ?
 - Comment développer et traiter les formations émergentes pathologiques, qui sont inéluctables dans un système ayant la liberté à se reconformer sous des tendances pulsionnelles ?

Remarques sur les notions de forme et de pensée

Dans cette transposition, la notion de forme qui sera considérée sera beaucoup plus abstraite que les formes produites par les activités neuronales dans le cerveau. Il faudra que ces formes soient représentables comme des structures faites d'éléments géométriques élémentaires se combinant selon une certaine algèbre, ce qui n'est pas le cas de l'activité neuronale. Mais cette notion de forme que nous devons développer est explicatrice de manière suffisante pour les conformations de la matière physique dans le cerveau qui produit de la pensée. Le lien central que nous retenons pour réaliser cette transposition sera celui qui existe entre forme et signification, qui permettra de donner du sens à l'activité produite dans le calculable. Mais est-elle vraiment artificielle, cette activité de systèmes informatiques si autonomes ? Dans quelle mesure est-elle plus artificielle que les mouvements effectués dans un système neuronal physique, si le résultat produit est similaire ? Le lien entre action et information n'a pas bien été clarifié dans ce domaine, comme il l'a été en physique quantique où les équations qui ne sont que des formes informationnelles expriment la réalité de la matière.

L'activité du système construit en informatique sera artificielle sur un point important. Les mouvements effectués par le réseau neuronal dans le cerveau sont des collisions de mouvements systématiquement parallèles, s'exécutant dans des groupes de neurones distincts mais au même moment. Nous serons contraint de simuler ce parallélisme massif, ce qui est délicat mais réalisable. Et Allan Turing a bien montré, il y a fort longtemps, que tout calcul d'une fonction par un programme qui s'exécute de manière parallèle est strictement équivalent à l'exécution du programme correspondant qui s'exécute séquentiellement. Le traitement parallèle est plus rapide, tout simplement.

Une définition de la pensée artificielle ne peut que se fonder sur les caractères d'un système effectivement constructible produisant pour son compte des représentations qu'il sait éprouver. Les caractères et qualités décrits par l'analyse du phénomène "pensée naturelle produite dans les cerveaux" ne valent évidemment pas directement dans cette transposition. Cela conduit les tenants, fort nombreux, de l'axiome de l'impossibilité à connaître et reproduire la pensée, à refuser même de prendre en considération une telle définition. Et cette attitude est une propriété de la pensée de l'homme, comme nous l'avons souvent dit, c'est une tendance à ne communiquer qu'avec ce qui lui est semblable ou pouvant être possédé et ainsi de refuser ce qui lui paraît être ontologiquement différent ou indifférent. La pensée est bien un système de représentations opérant pas distinctions avant que d'opérer par agrégations, qu'elle soit naturelle ou artificielle.

Donnons une définition constructiviste de ce qu'est la pensée artificielle, par ses caractères :

- Un système qui génère de la pensée artificielle est un système en action de reconfiguration de ses structures informationnelles, qui peut produire des états de conformation qui sont des émergences conduites, non aléatoires, valant, à chaque instant, pour une correspondance *forme – signification* et que je nomme alors "fait de conscience artificiel" : ce sont des *représentations* de différentes choses.

Et ce système a les caractères suivants :

- C'est un système informatique lié à un flux informationnel incessant venant de l'extérieur et allant dans l'environnement, produit par une corporéité composée de capteurs et d'effecteurs,
 - Il est composé d'entités calculables multiples qui sont proactives, chacune ayant un fondement cognitif et géométrique et ayant toutes la qualité d'être évolutives, s'agrégeant à d'autres ou se détachant d'agrégations,
 - Il a une architecture permettant de constituer un vécu événementiel sous forme de mémoire organisationnelle systématiquement augmentable, qui est une structure permettant une organisation,
 - Il subit des tendances fondamentales vues comme des contraintes générales du niveau organisationnel profond, qui sont des tendances pulsionnelles irrépressibles,
 - Il peut objectiver et mémoriser des structures à la fois informationnelles et géométriques valant, en étant activées, pour des éléments de signification, pour jouer sur lui-même et de lui-même,

- Il peut à tout niveau de sa fractalité de déploiement lier ses parties actives à sa totalité.