

Les problèmes de la conscience artificielle

Alain Cardon
AEIS, LITIS- INSA de Rouen

Plan

1. Un domaine avec des problématiques très éloignées.
2. Premier problème : trouver une modélisation constructiviste.
3. Deuxième problème : réaliser le système ou pas ?
4. Conclusion

1

Un domaine avec des problématiques très éloignées

Le problème général

- Comprendre vraiment le fait de penser à quelque chose, intentionnellement, comprendre la sensation de penser et comment peut se déployer cette pensée en appelant les suivantes.
- Le problème philosophique central pour l'homme : comprendre ce qu'il est et ce qu'il vaut.

Autres aspects du problème

- 1 Modéliser la génération de la pensée sous forme d'un **système** (*recherche*);
- 2 Concevoir et réaliser le système modélisé : concevoir l'architecture du système générant intentionnellement des faits de conscience éprouvés (*recherche-ingénierie*).
- 3 Appliquer le système là où la technologie rejoint les intérêts du marché (*technique*).

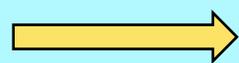
Des domaines éloignés

- Psychologie, psychiatrie, sociologie, linguistique :
 - Comprendre dans des domaines spécifiques en analysant les faits observés selon l'éclairage de chacun de ces domaines.
- Neurobiologie :
 - Comprendre en étudiant les éléments physiques observables selon les techniques disponibles : niveau du neurone et calculs considérables.
- Calculable (informatique, mathématiques) :
 - Faire des systèmes, selon des architectures précises, avec des approches de modélisation basées sur le traitement d'informations symboliques.

L'approche rationnelle

L'IA et les connaissances

- L'IA est centrée sur la rationalisation maximale de la notion de connaissances : atomes cognitifs, structures hiérarchiques, métriques d'écart, règles, inférences, logiques précises, significations, causes – conclusions.
- Aucune prise en compte des notions de subjectivité, de pulsion, de point de vue, de « climat sensible et émotionnel », de sens :



Limitations, mais des caractères nécessaires

Un point clé

- Un élément clé de cette recherche est de considérer que des traitements d'informations peuvent être **appréciés** par le système qui les produit, s'il a une architecture très adaptée pour réaliser cette appréciation (n'existe pas au niveau du modèle de Turing).
- Le domaine du traitement de l'information semble qualifié pour pouvoir traiter la conscience artificielle.

Ouverture pluridisciplinaire ?

- Comment constituer un domaine pour aller de la caractérisation d'un système générant intentionnellement des pensées adressant le sens des choses à une spécification informatique ?
- Comment constituer un **vécu** pour le système, ce qui fonde sa possibilité de penser et de ressentir à chaque fois quelque chose ayant du sens pour lui-même et culturellement ?

2

Les éléments architecturaux d'une approche calculable constructiviste

Éléments généraux de conception

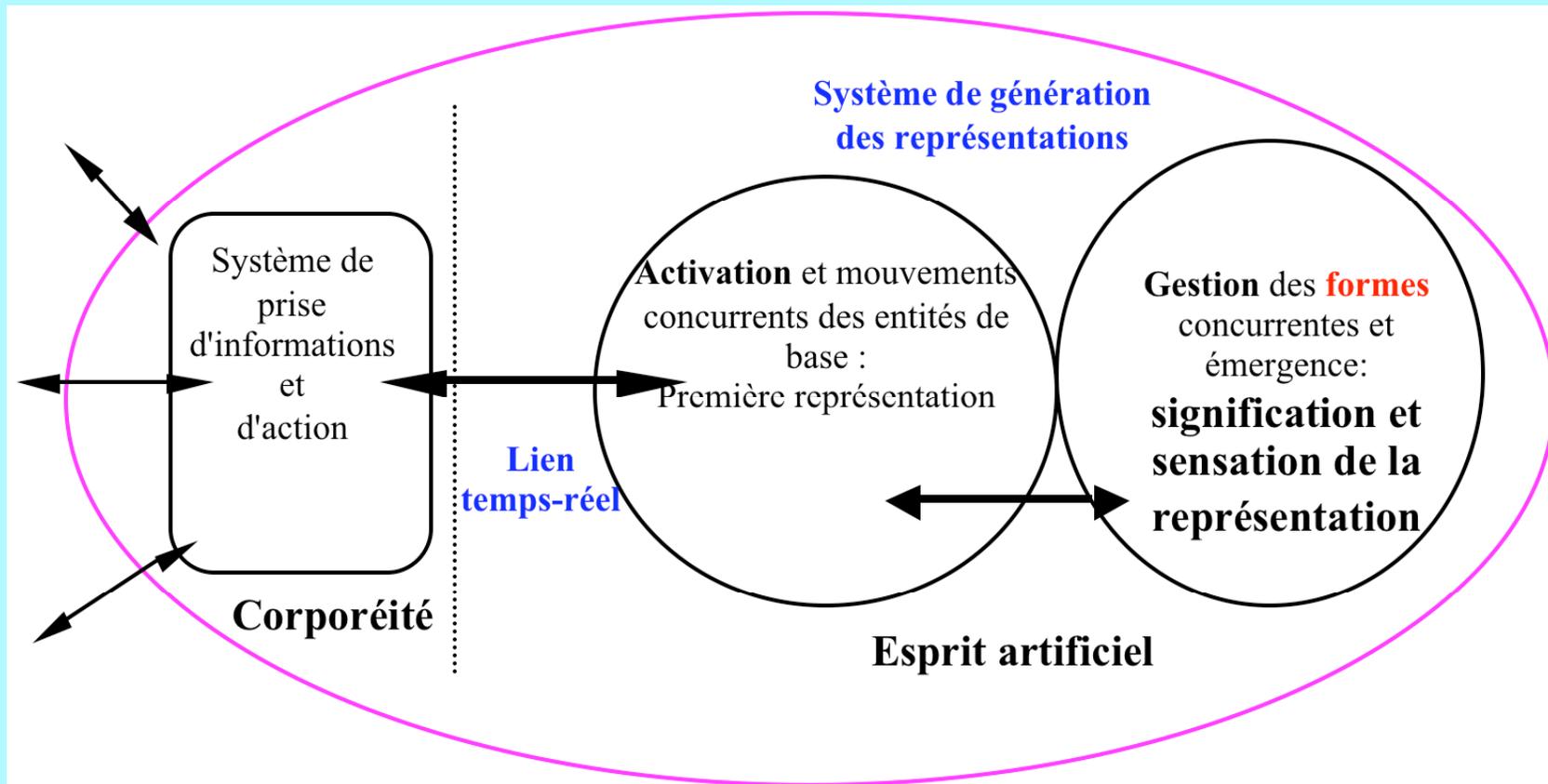
- Disposer d'un corps : robots autonomes multi-capteurs et multi-effecteurs.
- Concevoir un « système émotionnel »: architecture logicielle couplée au corps.
- Concevoir un « système générateur de pensées »: architecture logicielle fortement couplée à la précédente.
- Et poser les hypothèses de construction ...

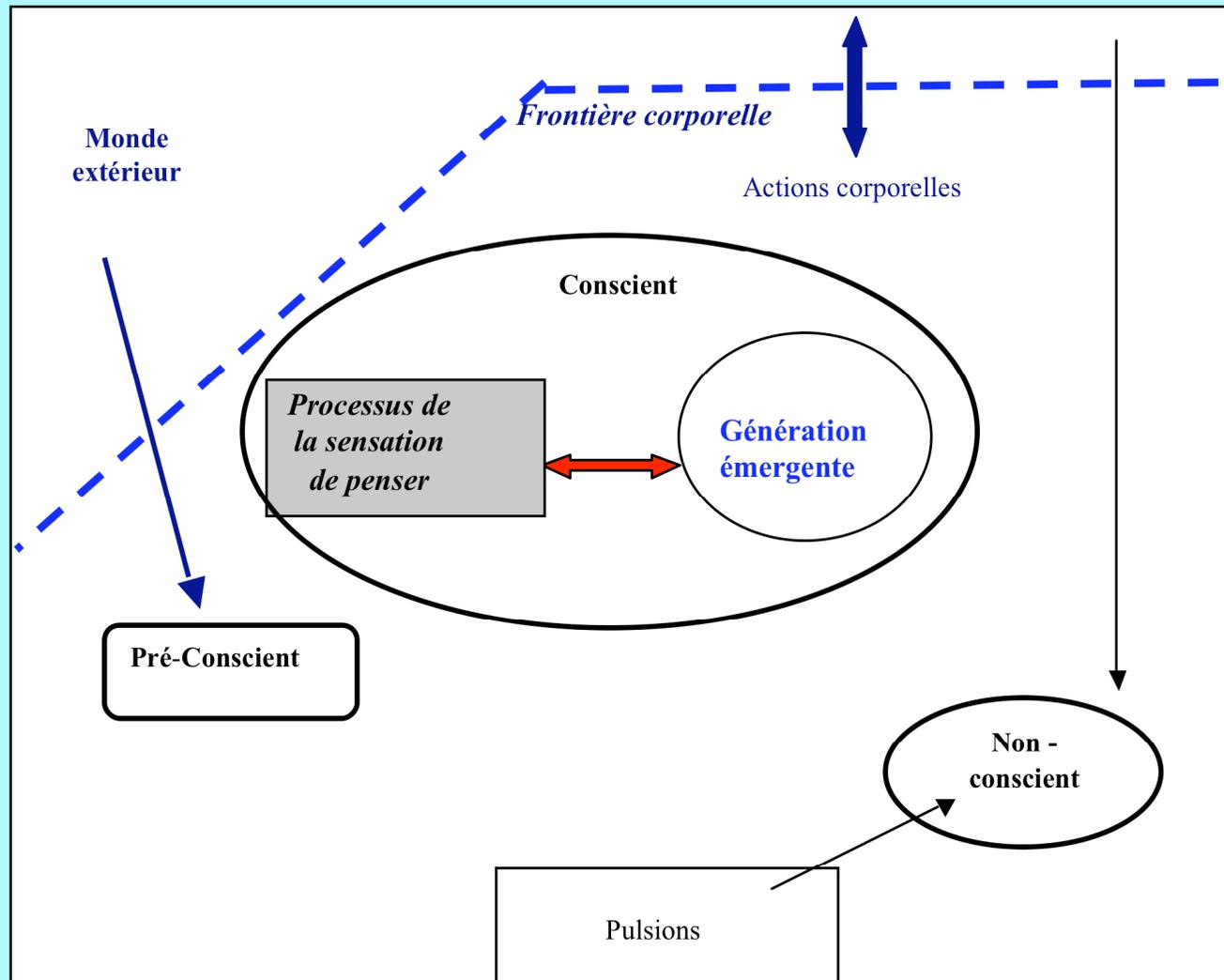
Point de vue constructiviste

- Hypothèse 1 : Définir une architecture générale basée sur les **première topiques freudiennes**.
- Hypothèse 2 : Les éléments de base ont une autonomie (**proactivité**), un pouvoir modificateur et ils peuvent s'agréger pour constituer des agrégats réifiés d'échelles plus grande.
- Hypothèse 3 : Il y a un **contrôle distribué** de tous ces éléments, selon l'architecture, ce qui est le point majeur : **nouvelle théorie du contrôle**.

Les hypothèses

- Le modèle est basé sur des éléments discrets : non prise en compte des théories du continu.
- Notion d'un fait de conscience éprouvé comme représentation géométrico-sémantique formée d'éléments dynamiques spatialement et temporellement organisés.
- Le modèle est basé sur une notion de contrôle très particulière tendant à produire des émergences manipulées par le système et valant pour le ressenti d'un fait de conscience.
- Il faut trouver une architecture pour permettre l'expression de ce contrôle et ce seront les premières topiques de Freud.
- Il faut disposer d'une mémoire organisationnelle valant pour un vécu.

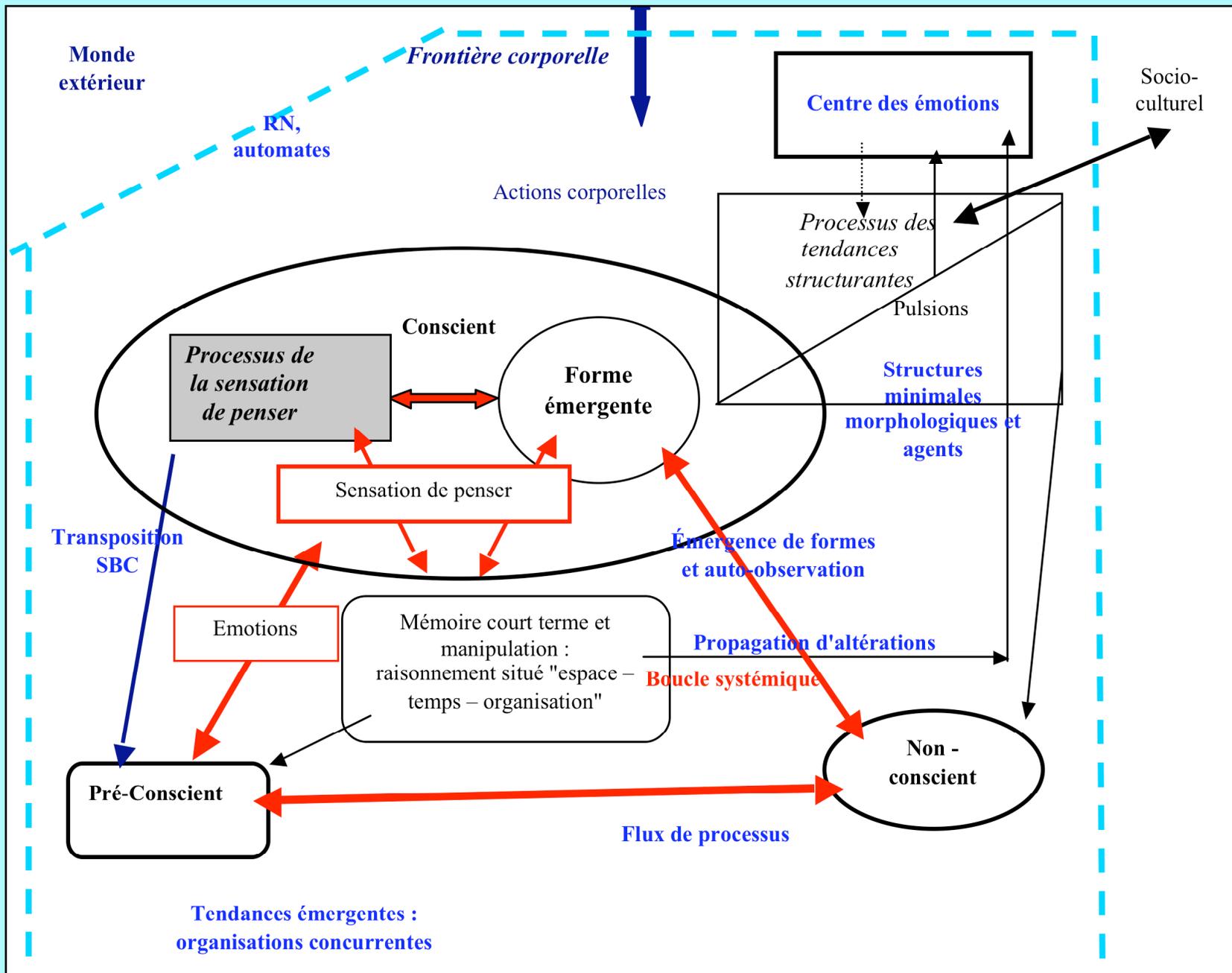




Les éléments majeurs

Représenter les trois instances et les émotions :

- 1 – Un **conscient** qui éprouve une forme de pensée.
- 2 – Un **pré-conscient** qui prépare les formes dont l'une sera éprouvée parmi d'autres.
- 3 – Un **inconscient** qui constitue une mémoire organisationnelle et qui loge des pulsions.
- 4 – Une **centre des émotions** reliant à la corporéité.
- 5 – Et une **boucle organisationnelle** qui coactive ces quatre instances.



Caractères généraux

- Corporéité effective avec flux d'informations continu venant des sens artificiels.
- Des **représentations** à construire sans cesse, tenant compte des éléments du passé et du présent disponibles dans la mémoire et engageant le futur du système par modification organisationnelle (sensation et mémorisation).
- Des tendances multiples et des contrôles permanents.
- **Émergences = processus continu de représentations appréciées selon leurs caractères géométriques et sémantiques.**

Émotions

États sensibles du corps artificiel, avec un système permettant l'interprétation des stimuli (externes et internes) **sans représentation consciente** de ce processus (sans ressenti).

Conceptuellement, c'est un état réactif correspondant à l'état dynamique d'un système particulier, en coactivité avec les autres instances : le **système générateur d'émotions**.

L'élément minimal : une clé

- Ce n'est pas l'élément neuronal, qui ne fait pas sens pour la pensée exprimée.
- Ce ne sont pas des structures de connaissances basées sur des objets dénotant des choses avec précision.
- Ce sera un élément réifiant une *indication de signification et appelant un contexte* pour s'exprimer.

Les agents aspectuels

Ce seront les éléments de base de la représentation idéale artificielle, au-dessus du neurone et en dessous du mot ou de la désignation. C'est principalement un élément *d'indication* et de *mise en relation implicite* utilisé pour former des amas dynamiques ayant des formes qui seront admissibles.

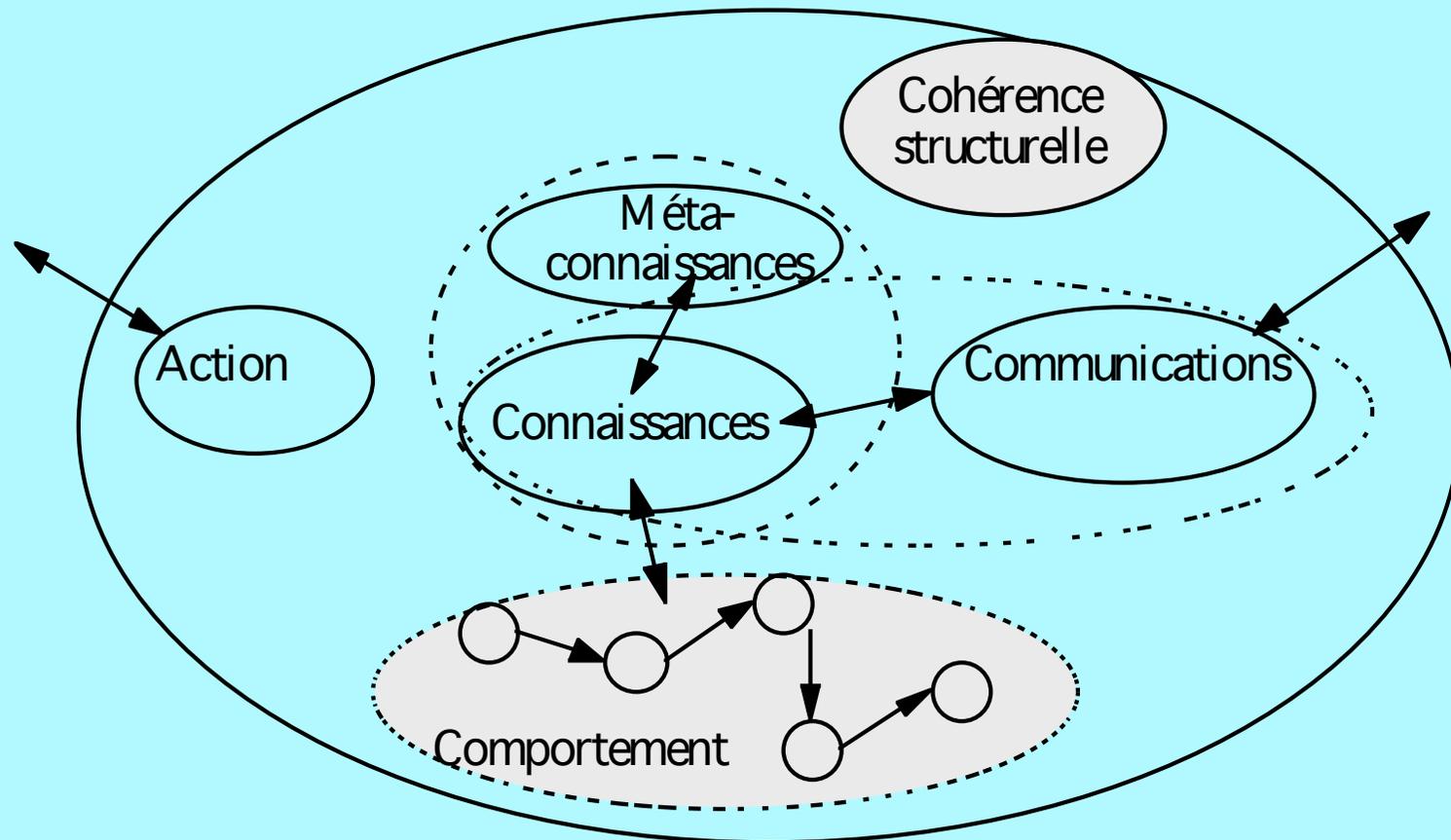
Remarque : Aucun rapport avec l'Idée de Platon ou la monade de Leibnitz.

Les agents aspectuels

L'agent réifie des aspects conceptualisables des choses formant des représentations :

- Il est auto-centré, c'est-à-dire qu'il a des buts propres (réactivité et pro-activité).
- Il est auto-motivé (pro-activité).
- Il est interagissant et communiquant.
- Il est structurellement hétérogène, c'est-à-dire formé de parties non nécessairement homogènes.
- Il est persistant.
- Il est très dépendant des autres agents.
- Il est combinable avec d'autres agents pour se fondre en un nouvel agent ou en générer plusieurs autres (changement d'échelle) et surtout former des agrégats.
- Il est le composant de base pour générer une forme dynamique.

L'agent aspectuel



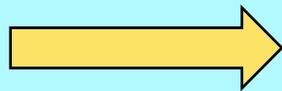
La mémoire organisationnelle : une autre clé

Constituer une mémoire des événements considérés comme vécus par le conscient artificiel :

- Pas des données ni des structures fixes.
- Mémoire qui produit des formes dynamiques valant pour des parties d'histoires.
- Mémoire qui se modifie à chaque utilisation.
- Mémoire qui se modifie en propre, constituant un inconscient artificiel.

L'objectif du fonctionnement

- 1 Former des agrégats d'agents aspectuels.
- 2 Faire se coactiver ces agrégats pour changer de niveau.
- 3 Dégager un préordre d'agrégats formant une émergence.



Donc un **contrôle** très particulier

Contrôle : attracteurs et régulateurs

- 1 - **Éléments de base** : agents logiciels proactifs et agrégations d'agents logiciels, et des contrôleurs opérant au niveau morphologico-sémantique de l'activité des éléments de base.
- 2 - **Régulateurs** : contrôleurs spécifiques des agrégations opérant selon les tendances (pulsions et raisons).
- 3 - **Attracteurs** : altérateurs, autonomes par rapport aux régulateurs et s'y opposant.
- 4 - Fonctionnement en **boucle systémique** : non-conscient, pré-conscient, conscient.

Rôle des contrôleurs

- 1 - Analyser la structure morphologique des agents et des agrégats actifs selon son caractère ontologique propre. Dégager les caractères pertinents de cette organisation en faisant une projection dans son thème. **Ce sera le rôle des agents de morphologie.**
- 2 – Chaque régulateur analyse l'état d'action et d'influence des autres régulateurs pour définir sa liberté d'action, selon sa classe ontologique.
- 3 - Tenter, sous conditions, d'amplifier ou de réduire certains agrégats, par appel d'éléments de base et aide de régulateurs associés étendant ou spécialisant sa sémantique.
- 4 – Conformer les agrégats pour générer une forme admissible.

Types de contrôleurs

1 - Contrôleurs établissant le climat psychologique, l'arrière plan de toute génération idéale dans une durée longue. Réification de tendances fondamentales, de pulsions.

2 - Contrôleurs établissant les caractères de toute génération idéale dans le moment de la génération : dans l'enveloppement des précédents, réalisent la visée et la génération émergente ressentie.

Catégories de contrôleurs

Classements et catégorisations, types de raisonnements, manipulations logiques, classifications et distinctions, types d'émotions et de sensations, éléments de posture, caractères temporels et spatiaux, abstractions, désignations, identifications et reconnaissances, formes et mouvements, catégories langagières, catégories géométriques, jugements et appréciations, anticipations, règles sociales, tendances profondes ou acquises...

Représentation : agents de morphologie.

Morphologie de paysage d'agents

L'auto-contrôle de l'organisation des agents

Aspectuels est assuré par une autre organisation d'agents :

L'organisation des **agents de morphologie**, analysant *on-line* les caractères organisationnels de l'organisation des agents aspectuels (ses mouvements typiques).

Caractères morphologiques

Considérer tout agent aspectuel comme soumis à des forces internes et externes :

- Vitesse de développement de l'agent : *vélocité*
- Activité et développement social de l'agent par rapport à son contexte : *facilité*
- Intensité et domination de l'agent par rapport aux autres : *suprématie*

.....

Les dimensions de l'espace morphologique

Aspect externe de l'agent	Suprématie
	Indépendance
	Persistance
Aspect interne de l'agent	Facilité
	Vélocité
Etat organisationnel de l'agent	Intensité du flux interne d'activation
	Complexification
	Fréquence communicationnelle
	Ecart organisationnel
Ouverture du système	Transport d'information

Fonctionnement (1)

Boucle systémique organisationnelle **en parallèle contrôlé** :

- 1 - Intention et / ou stimulus
- 2 - Action de l'inconscient (pulsions)
- 3 - Déploiements dans le pré-conscient avec l'arrière-plan psychologique générant des formes
- 4 – Emergence ressentie dans le conscient
- 5 - Mémorisation dans la mémoire organisationnelle...

Caractères : ceux de la boucle systémique = durées, collisions, synchronisations, retards, évitements...

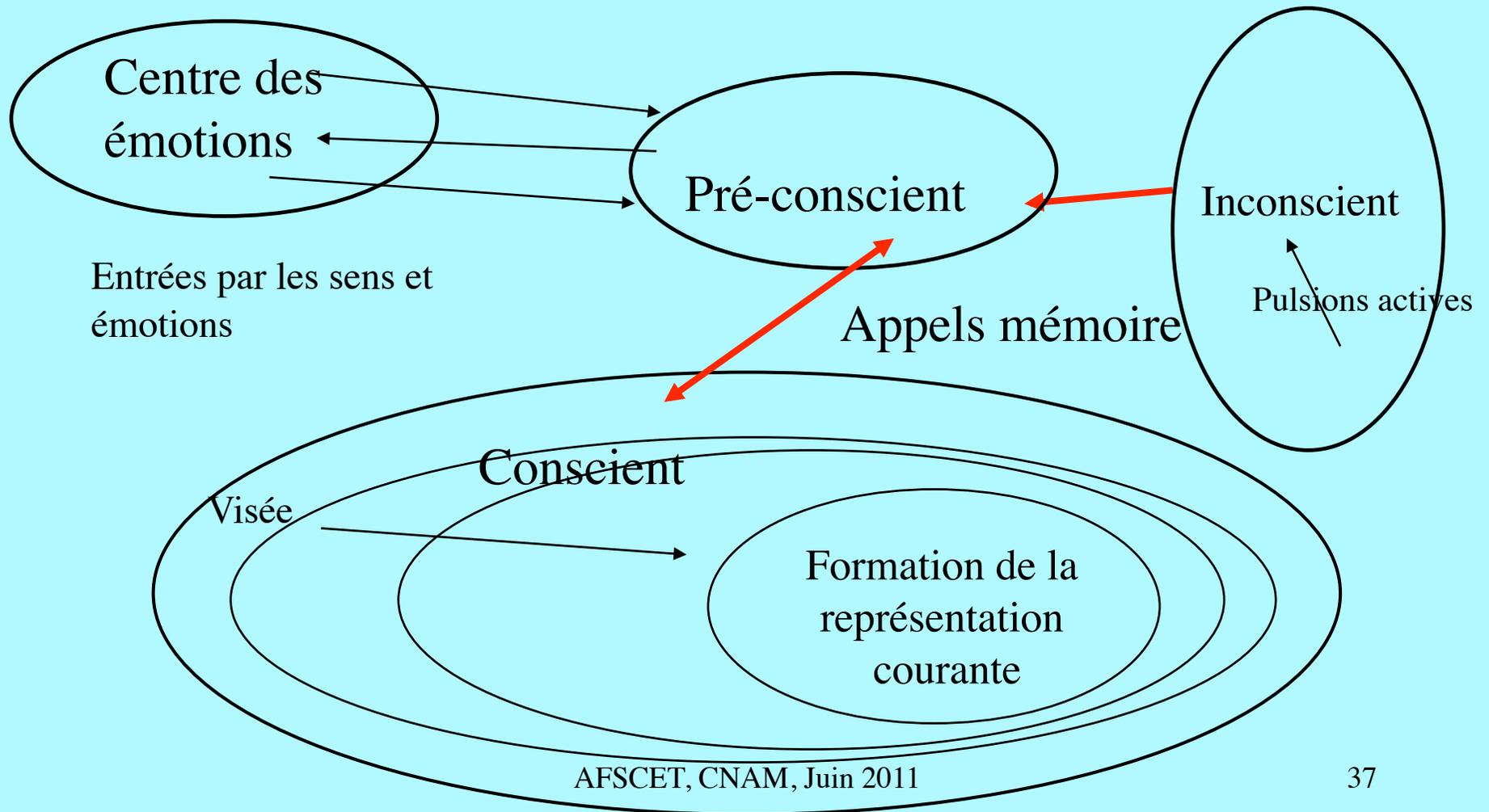
Fonctionnement (2)

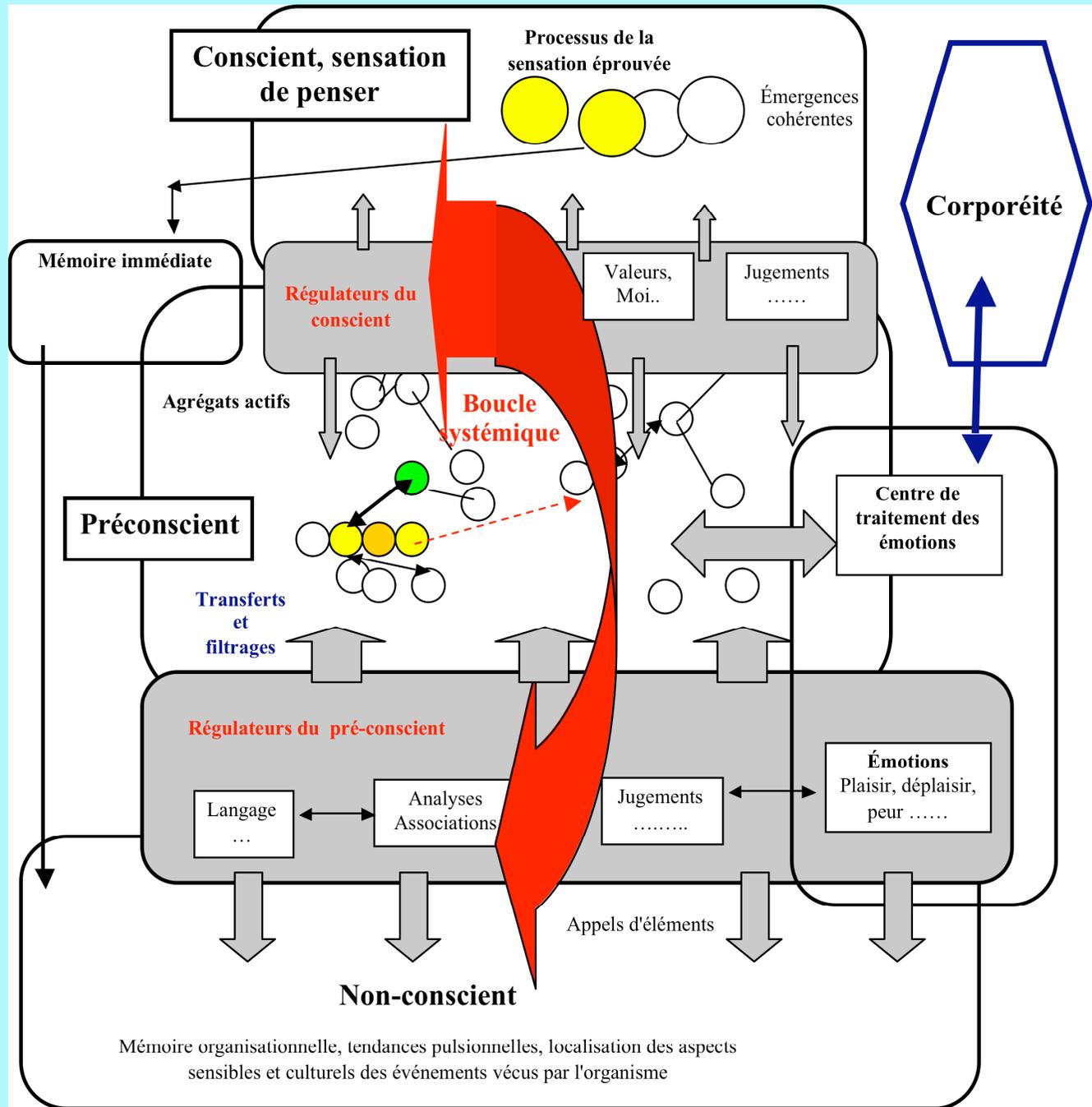
Possibilités de dysfonctionnements nombreuses, locales ou globales (pathologies) notamment par bifurcations spatiales ou temporelles, à de multiples niveaux dans les conformations des agrégations.

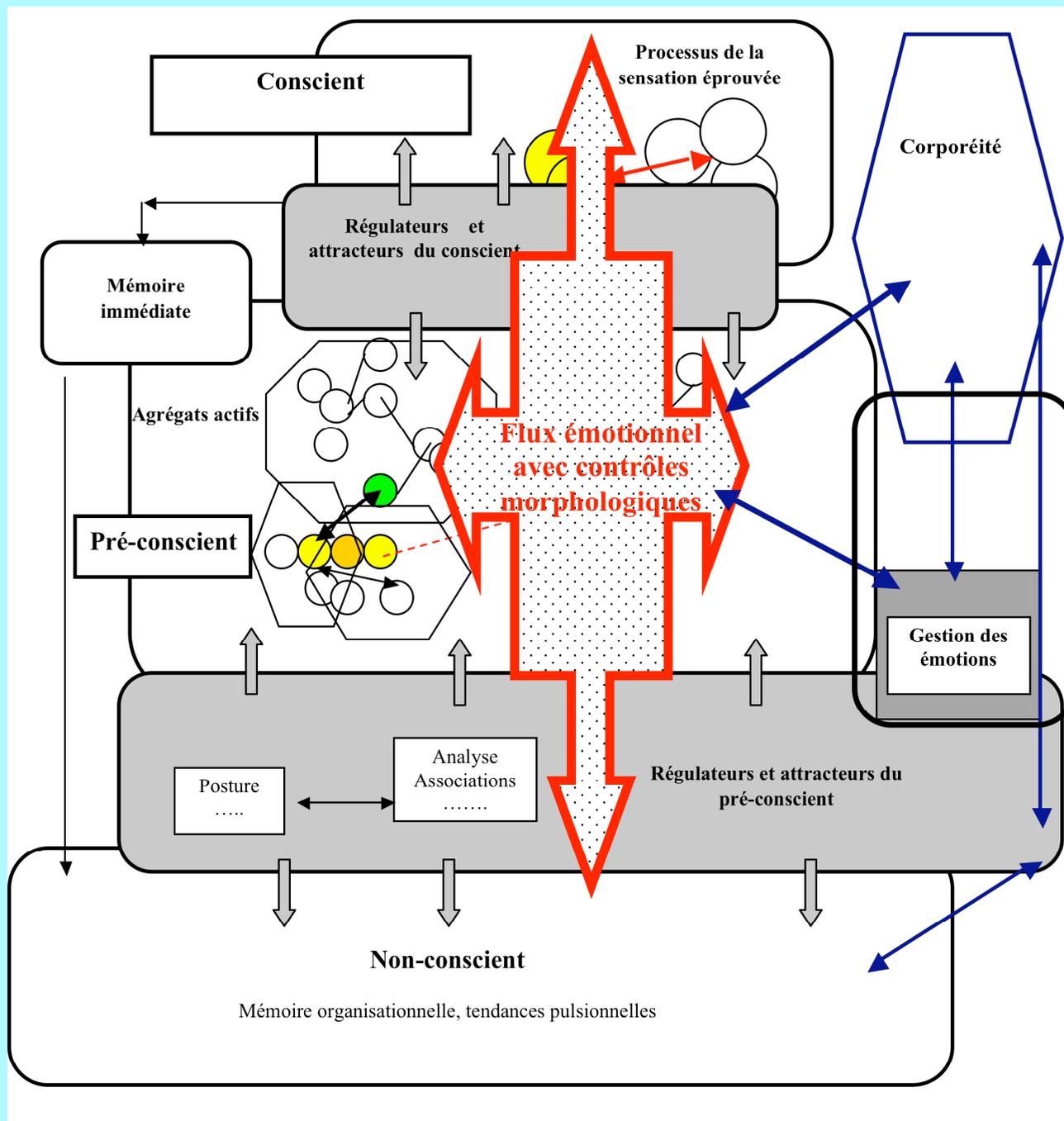
Plasticité maximale : nouveaux éléments de base et nouveaux contrôleurs pouvant se créer.

Espace des contrôleurs : autre morphologie, multi-échelle.

Processus de génération



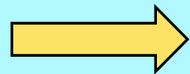




Ce qu'est penser artificiellement

Générer une scène et la ressentir :

- 1 - Activation de certaines structures de signification, selon un certain ordre, constituant une certaine forme.
- 2 - Avec certaines synchronisations, en liaison avec certains caractères de vitesse et d'ampleur de la boucle systémique...
- 3 - Référant à certaines catégories, certains rôles, certaines images.



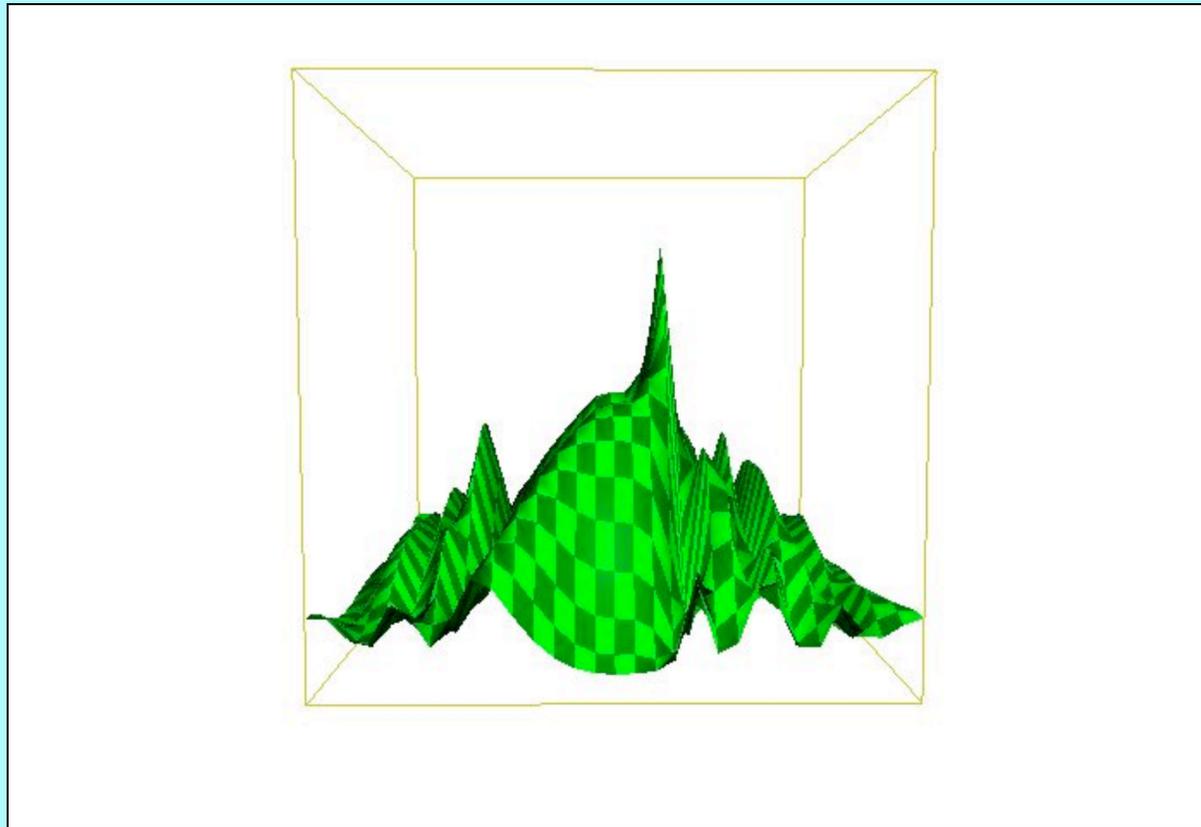
Le tout étant dynamique et apprécié à l'aide de morphologies manipulées ...

Forme d'une pensée artificielle

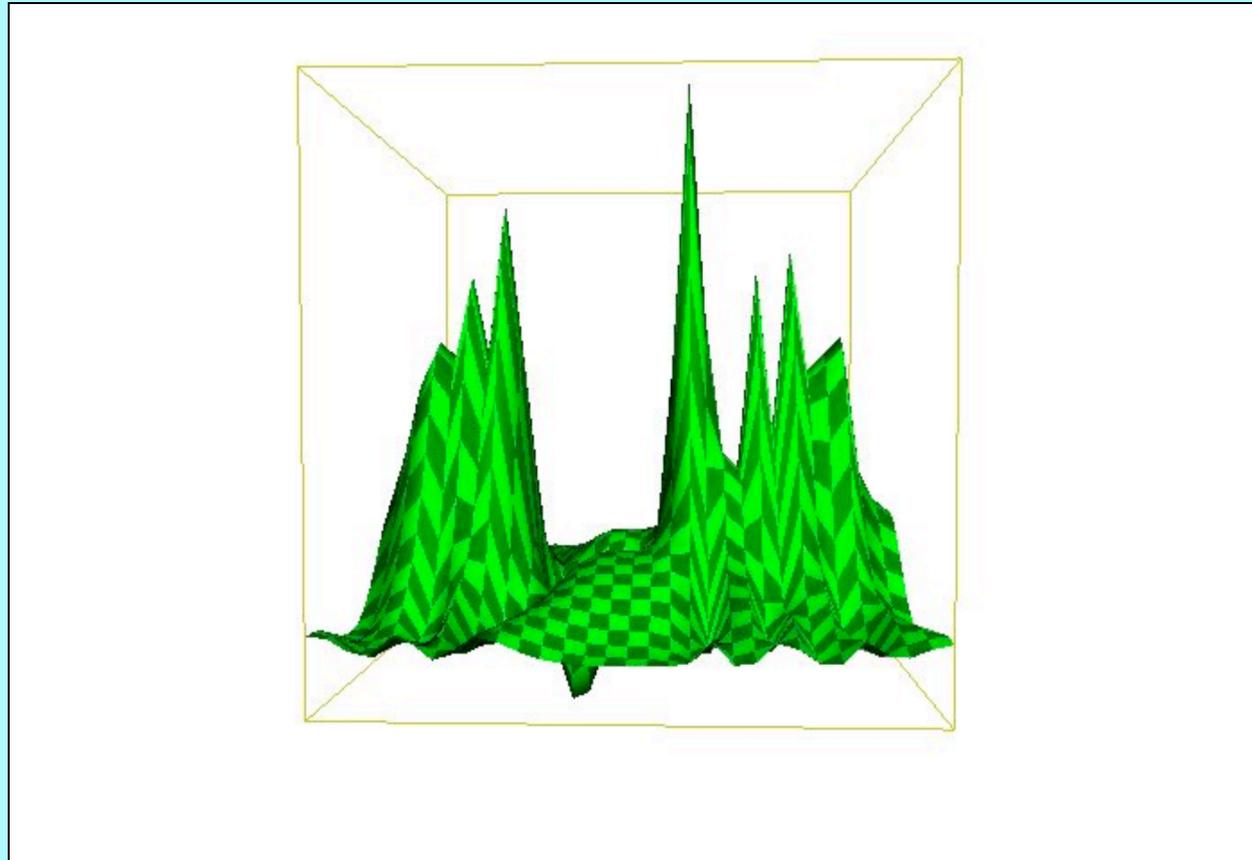
- Arrière-plan d'atmosphère, d'éclairage.
- Arrière-plan de suscitations (tendances multi-échelles).
- Avant-plan intrusif local de certains attracteurs.

- Construction de la sensation qui va donner la pensée éprouvée.
- Mémorisation très spécifique.

Espace des contrôleurs : cas centré



Espace des contrôleurs : cas de bifurcation



Relations de limitation

- **Relation de limitation entre profondeur de la forme émergente et continuité du flux idéal :**
 - Profondeur intentionnelle x Continuité organisationnelle \leq **Capacité informationnelle**

Relation de limitation entre la rapidité de production des émergences et la qualité des mémorisations :

- Rapidité d'émergence x Amplitude de mémorisation \leq **Capacité informationnelle**

Recherches...

- Trouver la forme de l'élément d'aspect d'une "*pensée artificielle éprouvée à propos de quelque chose*".
- Trouver les opérateurs d'agrégation conduisant à la pensée artificielle exprimée.
- Trouver la structure permettant de transposer la "*sensation de se représenter une pensée*".
- Trouver les opérateurs de mémorisation.
- Trouver les opérateurs permettant de faire se générer des pensées artificielles (intentionnalité).
- Trouver les qualifications de toute pensée artificielle.

Le problème majeur :
construire ou pas le système ?

L'enjeu de la construction

- Connaissance précise du système de pensées de l'homme :
 - Mieux comprendre et situer l'homme et ses comportements sociaux d'après ce qu'il est capable de penser.
 - Distinguer ce qui est inné de ce qui est acquis et comprendre l'inconscient.
 - Caractériser les défaillances et les pathologies.

La science dans son époque

- Surnatalité : 230 000 humains de plus chaque jour.
- Consommation en hectares globaux de la population : **1.4** planètes Terre en 2010.
- Budget d'entretien des armées supérieur à l'alimentation de la population mondiale.
- Finance : **3%** pour le PIB de la planète Terre dans le jeu des échanges financiers boursiers.
- L'histoire tragique du 20^{ème} siècle : un passé très inquiétant qui se poursuit.

La réalisation ?

- Une petite équipe d'informaticiens sachant programmer et comprenant le problème, pour réaliser un prototype du système à partir de spécifications.
- Équipe pluridisciplinaire importante pour **définir le vécu** et valider les productions : des recherches longues, difficiles, devant être publiques.

Le choix très probable

- Réalisation du système pour et par des militaires, un système bien plus simple, sans profondes recherches sur le vécu, avec un objectif pour ingénieurs : faire des systèmes d'armes et de contrôle totalement autonomes, dotés d'intentions et de tendances fondamentales précises et limitées.

L'enjeu de la construction

- Applications dans le monde actuel tel qu'il est aujourd'hui :
 - 1 Systèmes d'armes autonomes très coordonnées (coactives) gérant leurs usages sur les théâtres d'opérations urbains.
 - 2 Système de surveillance méta intégrant toutes les données informationnelles manipulées par les humains
 - 3 Les applications humanistes et humanitaires multiples ?

Les engagements

- Le projet est un thème de recherche en cours du Ministère de la Recherche : **CHIST-ERA**, dans le thème « Systèmes d'Information ».
- Informer les citoyens du problème majeur des conséquences de la réalisation, ce qui n'est pas fait.

Conclusion

- Choix personnel de ne pas réaliser le système : renoncement, abandon.
- Des développements du modèle général présentés dans :
 - <http://www.automatesintelligents.com/interviews/2011/interviewcardon.html>
- Écriture achevée d'un livre grand public sur le problème, transmis aux éditeurs.

Proposition de thème de recherche pluridisciplinaire AEIS

- Nous proposons de travailler sur la constitution et l'utilisation du vécu artificiel intervenant dans la production de pensées intentionnelles vues comme des formes manipulées, éprouvées et utilisées par le système qui les construisent. L'étude portera sur :
 - 1. Organisation générale d'un vécu (architecture et éléments de base)
 - 2. Organisations particulières d'un vécu
 - 3. Espace des attracteurs et des régulateurs formant le contrôle organisationnel du vécu
 - 4. Éléments hégémoniques produisant des tendances globales éventuellement altératrices
 - 5. Fracturation et bifurcation des flots idéels et formation des pathologies
 - 6. Boucle systémique utilisant le vécu : régulation spatio-temporelle avec les régularités et les irrégularités.