

IA & Cognition

Les approches émergentes

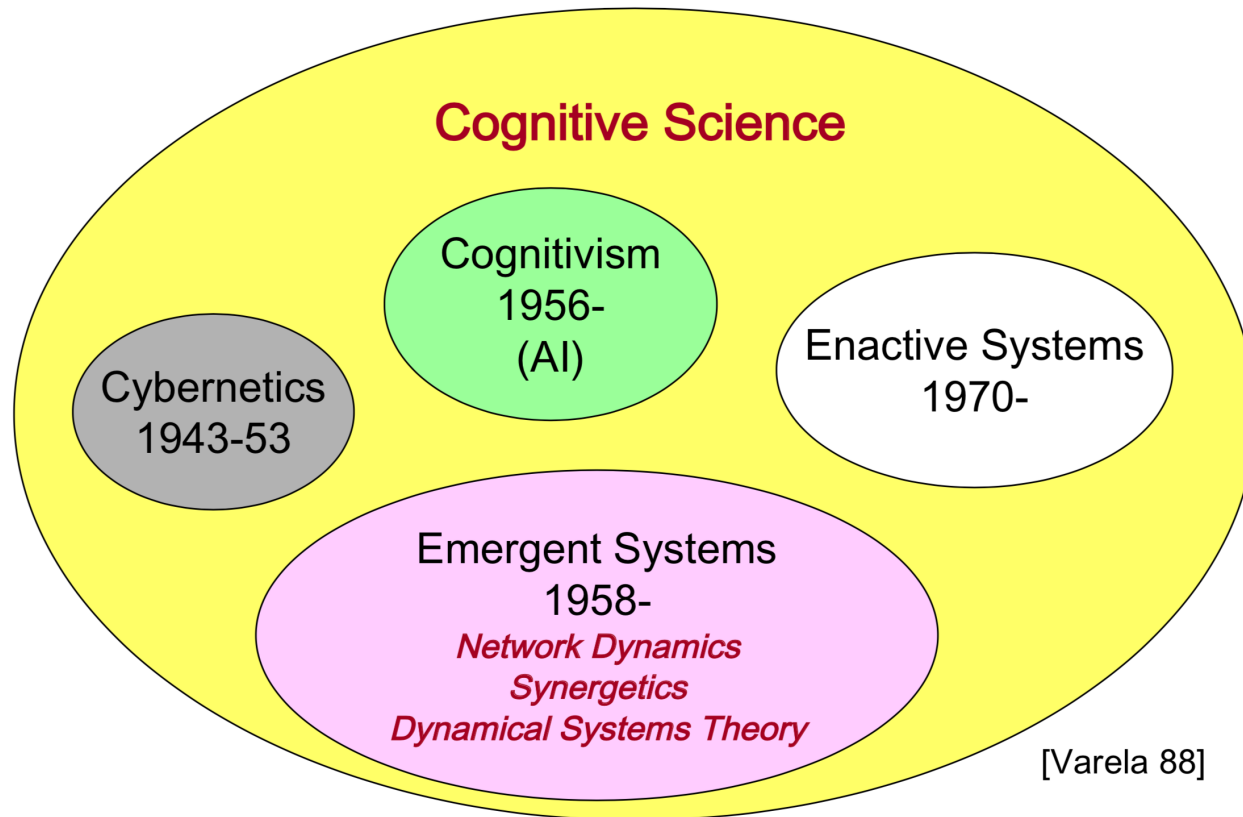
Marie Lefevre

Université Lyon 1 - Laboratoire LIRIS

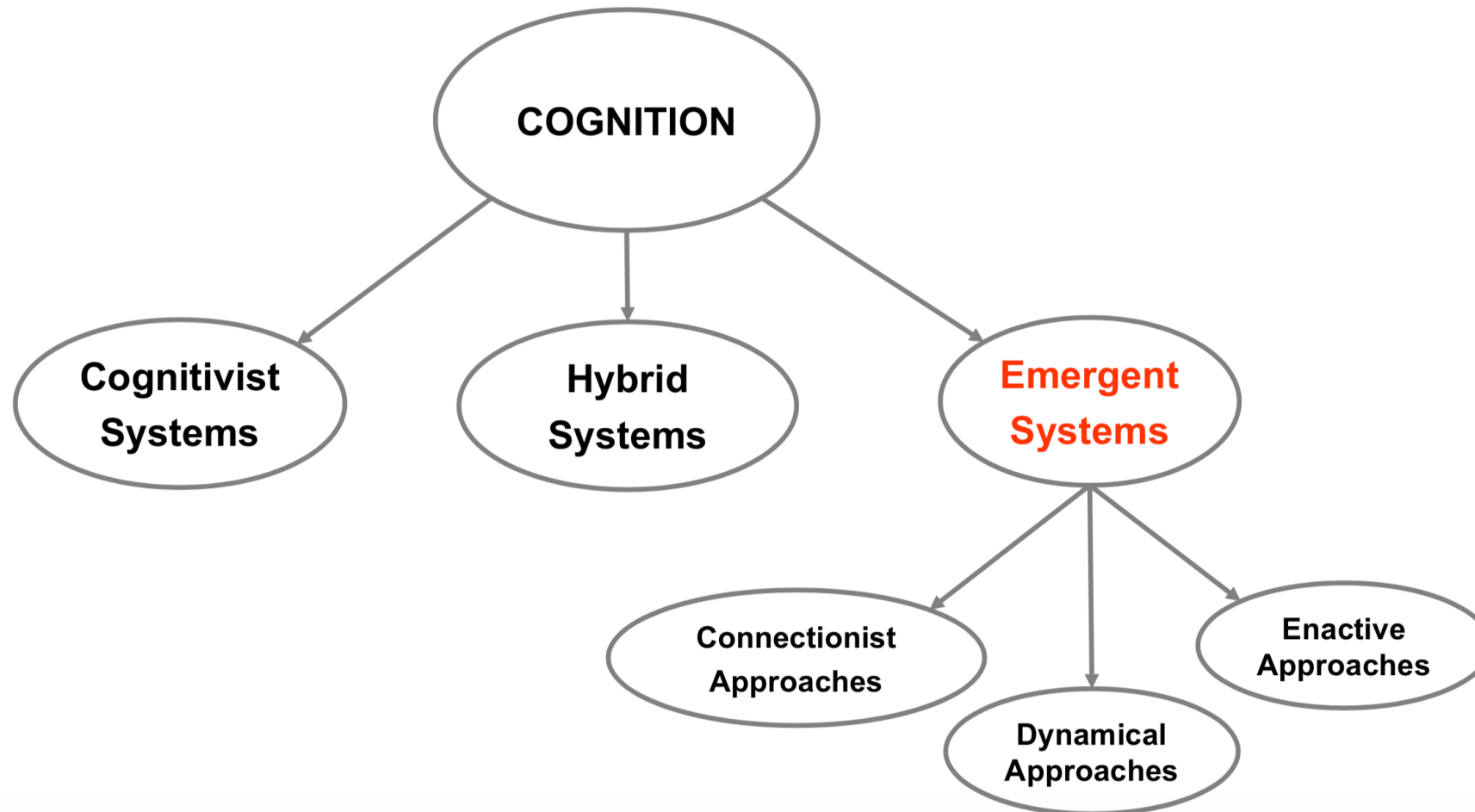
marie.lefevre@liris.cnrs.fr



Paradigmes / courants en Sciences Cognitives



Paradigmes / courants en Sciences Cognitives



Approches émergentes

- ▶ La cognition est le processus par lequel un système autonome devient viable et efficace dans son environnement
- ▶ Il le fait par le biais d'un processus d'auto-organisation
 - ▶ Le système se reconstitue continuellement
 - ▶ En temps réel
 - ▶ Pour maintenir son identité opérationnelle
 - ▶ *via* la modération de l'interaction système / environnement
 - ▶ *via* la co-détermination

Approches émergentes : co-détermination

- ▶ L'agent cognitif est spécifié par son environnement
- ▶ Le processus cognitif détermine ce qui est réel ou significatif pour l'agent
- ▶ Le système **construit sa réalité** (son monde) à la suite de son fonctionnement dans ce monde
- ▶ La perception fournit des données sensorielles pour permettre une action efficace, mais toujours en conséquence des actions du système
- ▶ La cognition et la perception dépendent fonctionnellement de la richesse de l'interface d'actions

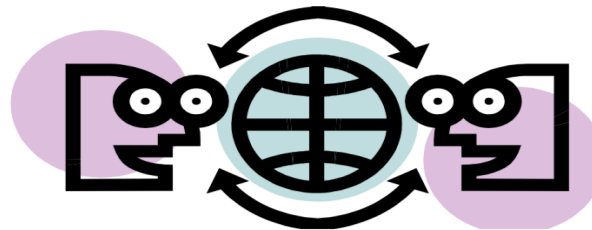
Approches émergentes : cognition & perception

- ▶ *La cognition est le complément de la perception [Sandini]*
 - ▶ La perception traite de l'immédiat
 - ▶ La cognition traite des délais plus longs
- ▶ Le modèle principal de l'apprentissage cognitif est la construction anticipée des compétences, et non l'acquisition des connaissances
- ▶ Le point de départ de l'intelligence est d'agir efficacement, d'anticiper la nécessité d'agir et de faire évoluer ses possibilités d'actions
- ▶ Tout cela en étant incarné dans un système physique capable d'interactions physique avec le monde

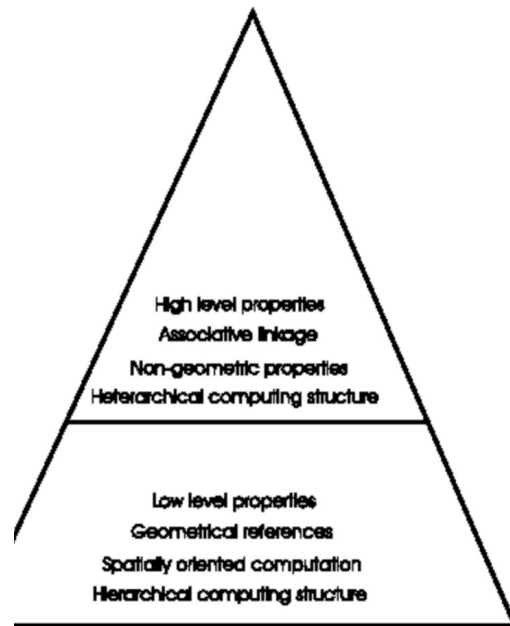
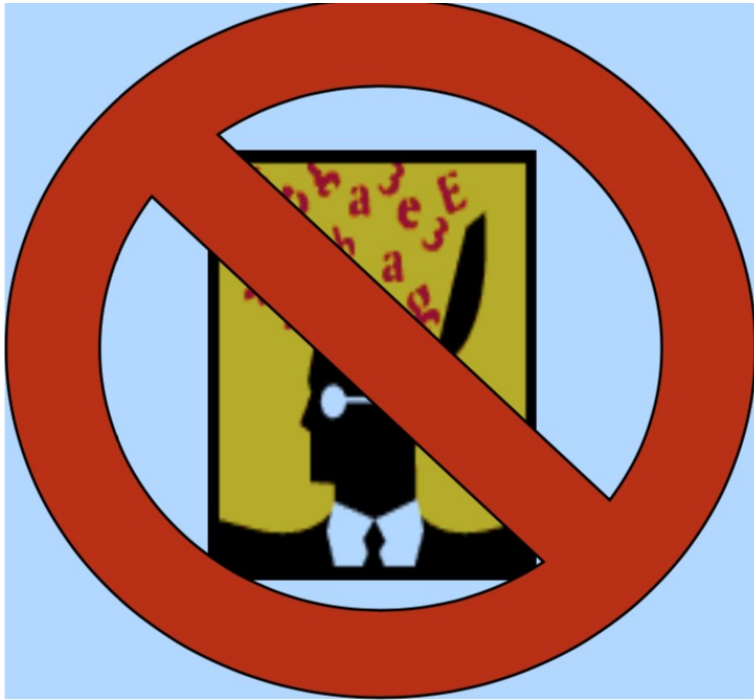
Approches émergentes

'Cognitive systems need to acquire information about the external world through learning or association'

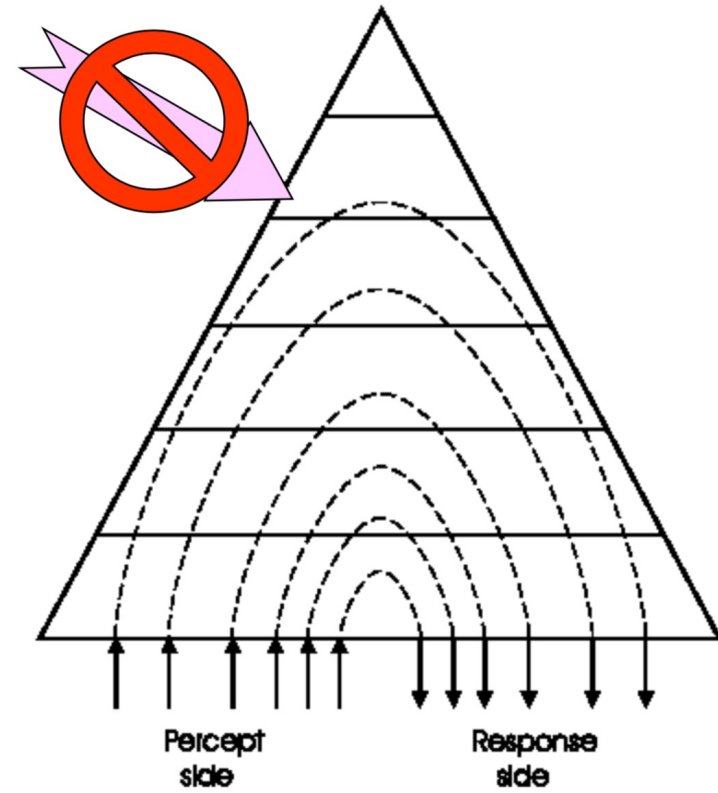
[Granlund'02]



Approches émergentes



Systems



iversib

Approches émergentes

▶ Auto-organisation

- ▶ Les systèmes auto-organisés sont des systèmes physiques et biologiques dans lesquels le modèle et la structuration au niveau du monde proviennent uniquement des interactions entre les composants de niveau inférieur du système
- ▶ Les règles spécifiant les interactions entre les composants du système ne sont exécutées qu'à l'aide d'informations locales, sans référence au modèle global

▶ Emergence

- ▶ Un processus par lequel un système d'éléments en interaction acquiert un modèle et une structure qualitativement nouveaux qui ne peuvent être compris simplement comme la superposition des contributions individuelles
- ▶ « *Le tout est plus grand que la somme des parties* » [Aristote]

Approches émergentes : auto-organisation



Trafic à Hanoi

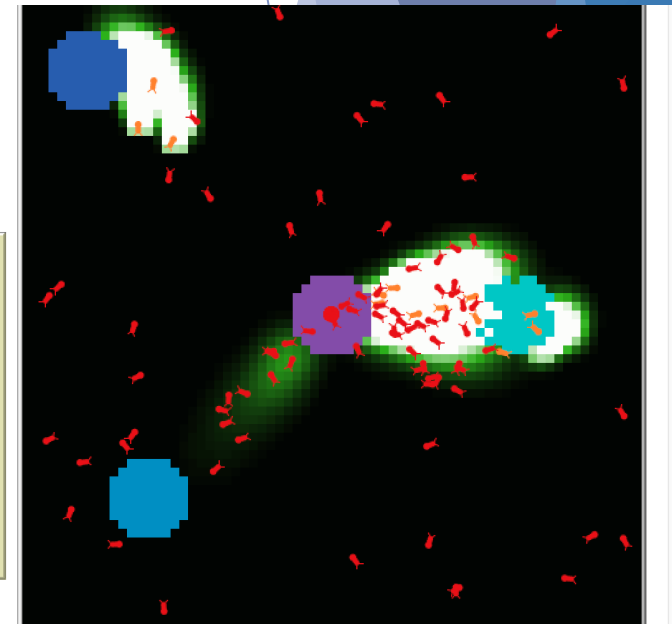
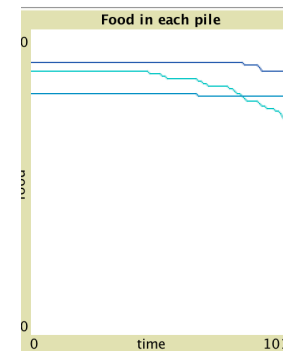
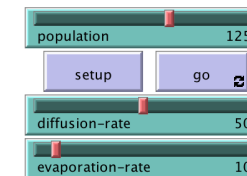


Fourmis



Vol d'étourneaux

Septembre 2023



Paradigmes naturo-inspirés : l'auto-organisation

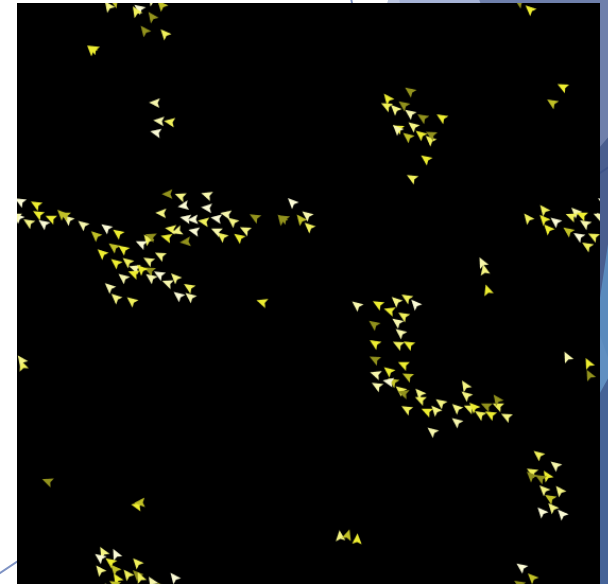
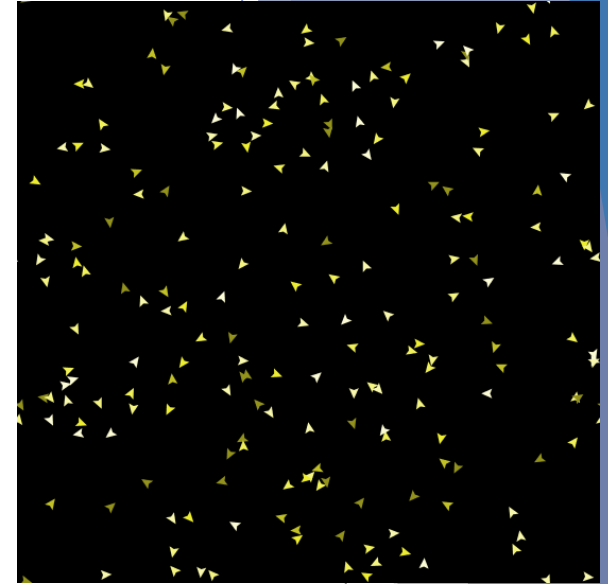
L'auto-organisation: une définition (Bonabeau)

- ▶ Un ensemble de **mécanismes dynamiques** via lequel des **structures émergent** à un niveau **global** d'un système, à partir d'**interactions** entre ses composants de **plus bas niveau**. Les règles spécifiant ces interactions, sont exécutés sur la base **d'informations purement locales**, **sans aucune référence aux structures émergentes du niveau global**, qui sont considérées comme une propriété émergente du système **plutôt qu'une propriété imposée au système par une influence (ordonnée) externe**.

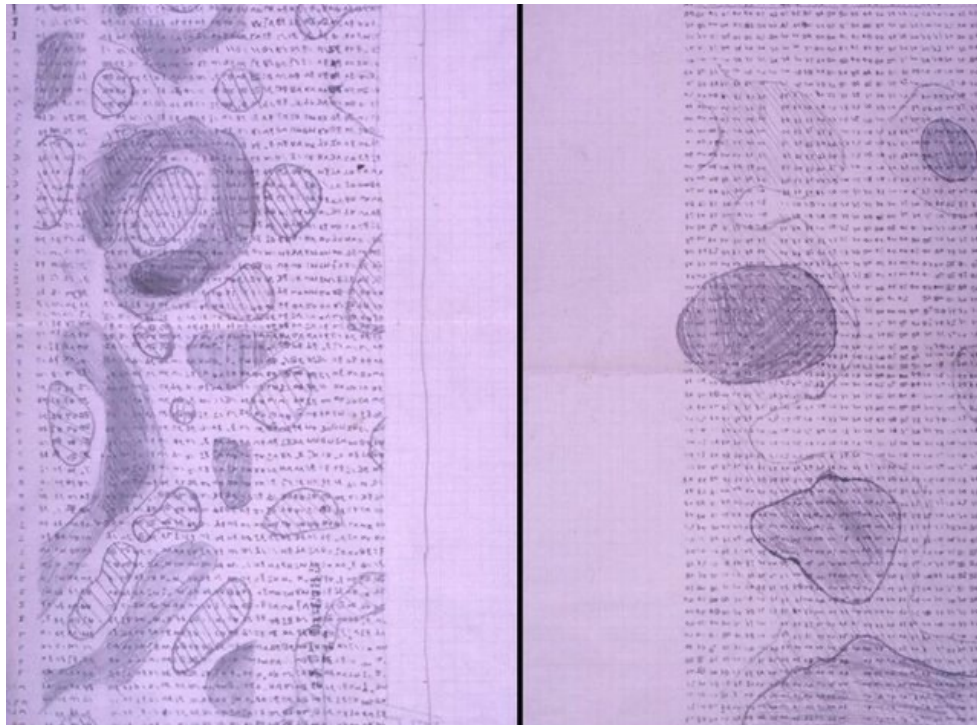
Paradigmes naturo-inspirés : l'auto-organisation

L'auto-organisation: autre définition (plus simple) (Palut 2004)

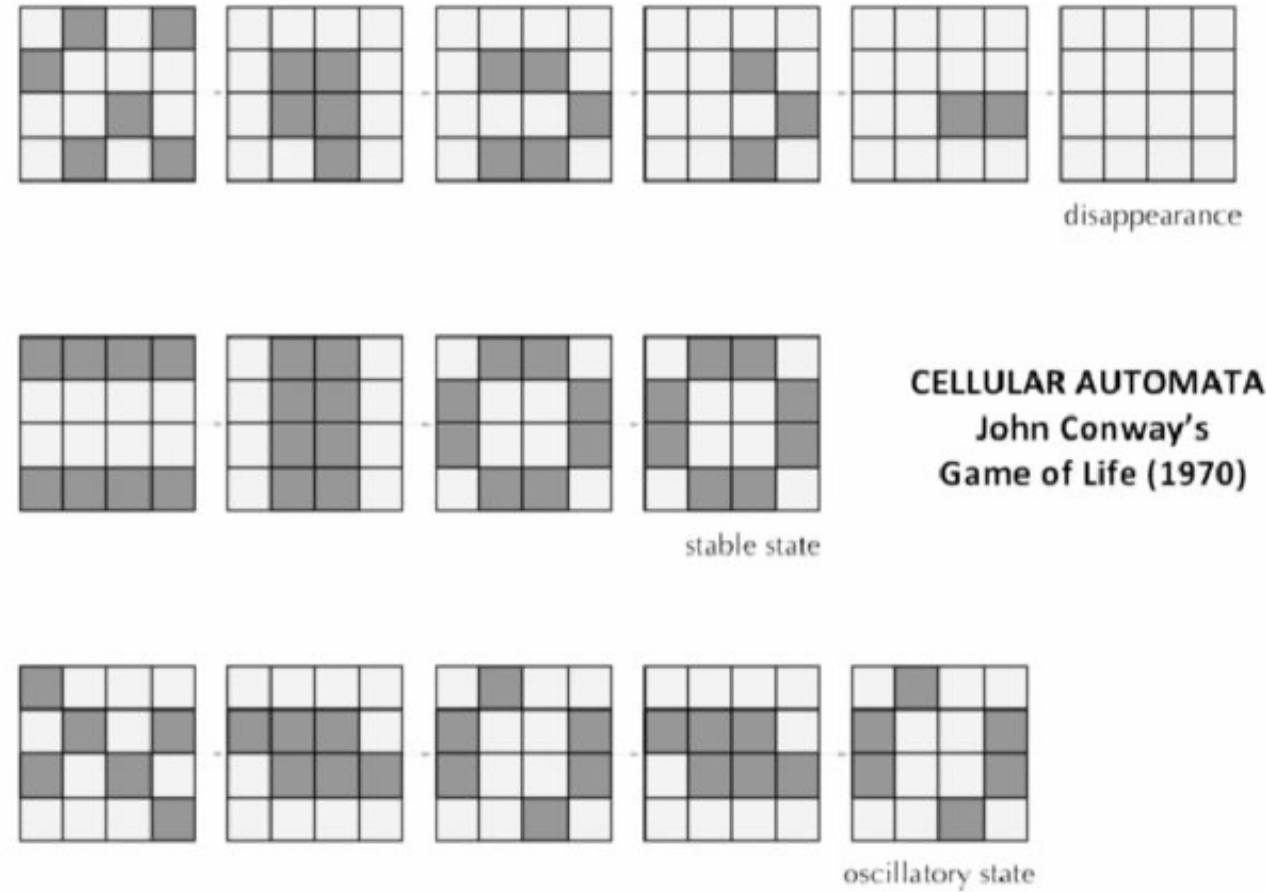
- ▶ L'auto-organisation est un processus selon lequel des structures collectives émergent à partir des multiples interactions entre les éléments d'un système sans que ces structures collectives ne soient explicitement codées ni au niveau des éléments, ni au niveau collectif.



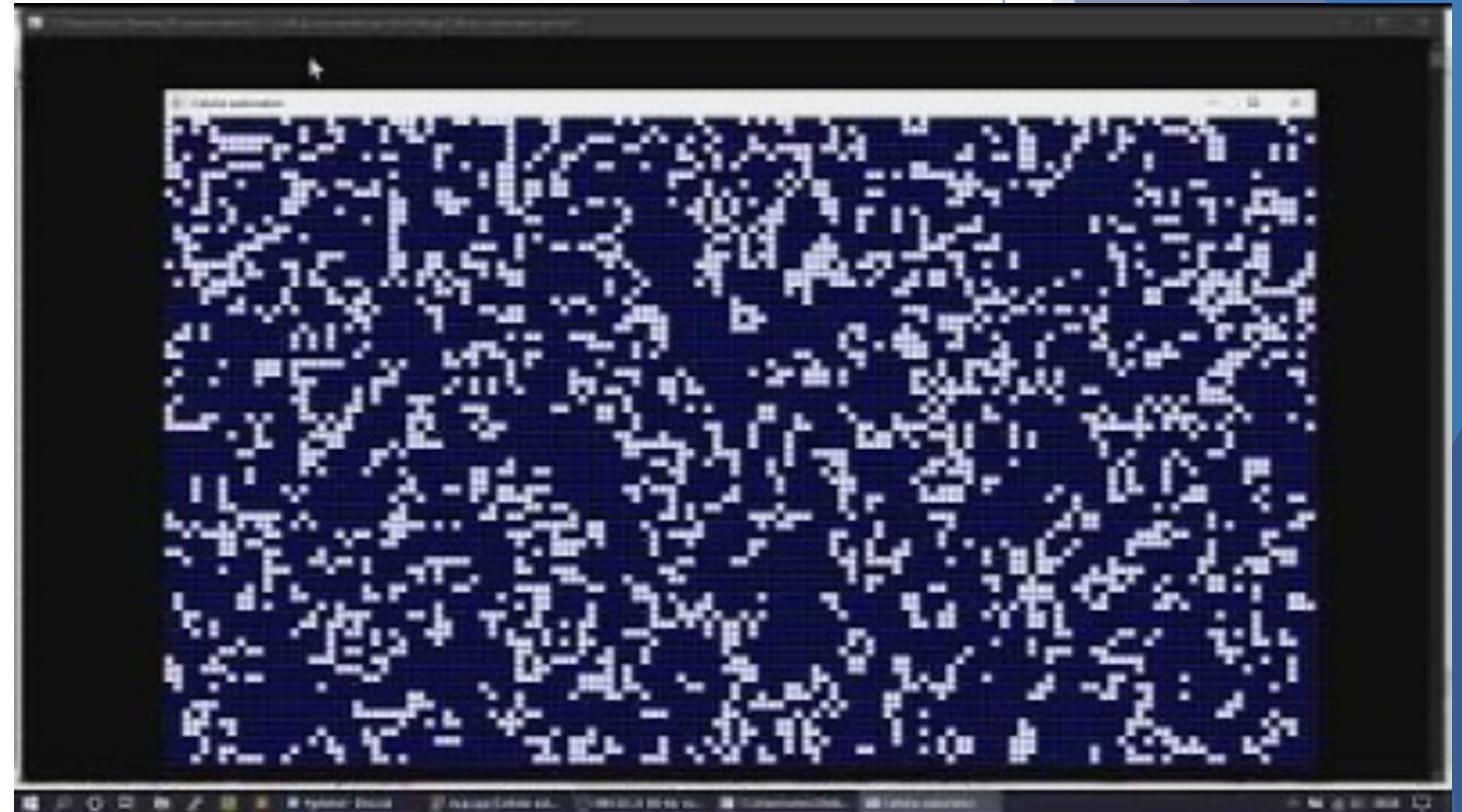
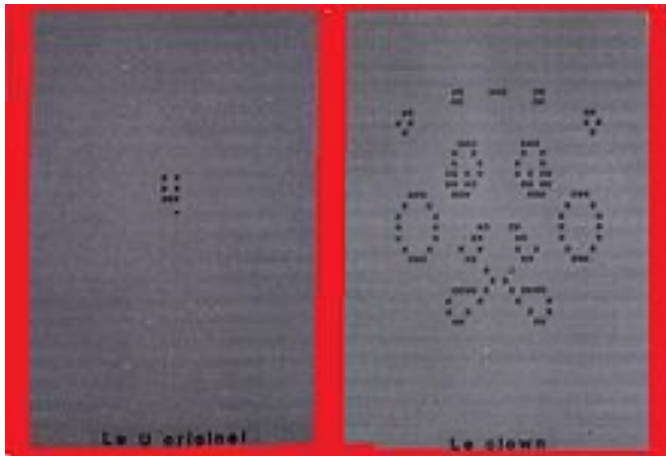
Exemple : le jeu de la vie



Motifs calculés avec le Manchester Electronic Computer puis dessinés par Alan Turing



Exemple : le jeu de la vie



Marie Lefevre



Paradigmes naturo-inspirés : l'auto-organisation

- ▶ La coordination entre les éléments est uniquement déterminée par les contraintes (flux d'énergie, de matière, d'information) qui agissent sur les interactions des éléments du système
- ▶ Les structures collectives produites résultent exclusivement du couplage entre les individus ou entre les individus et leur environnement (couplage rétroactions + et -)

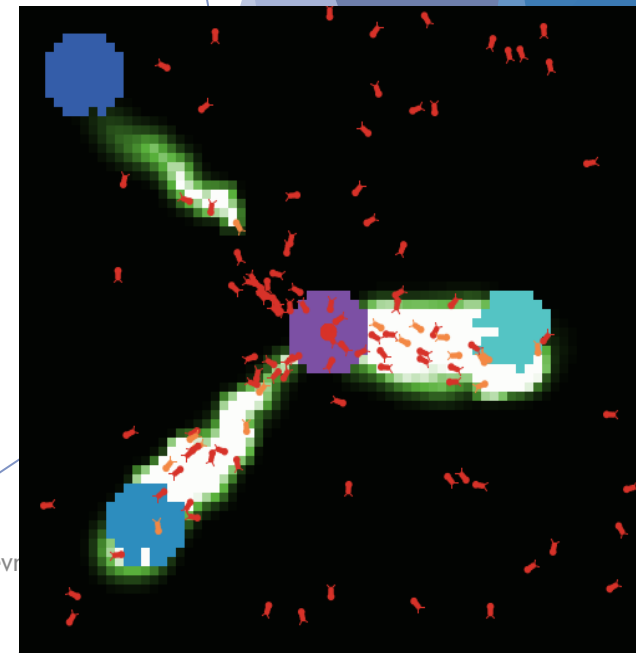
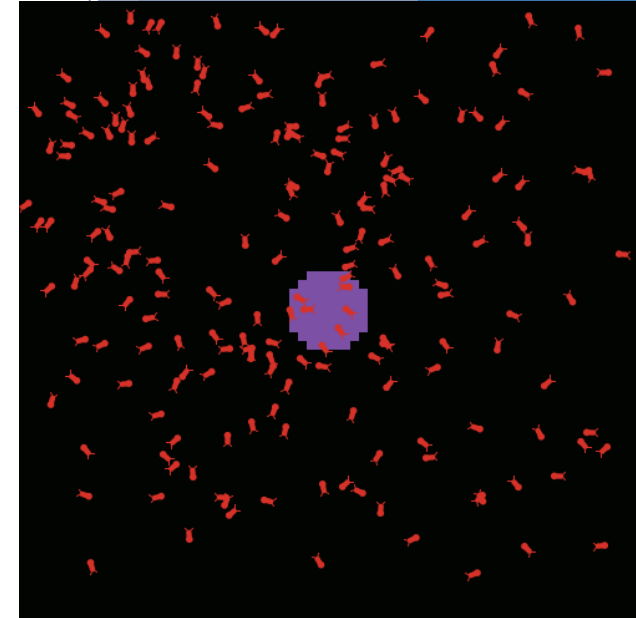
Paradigmes naturo-inspirés : l'auto-organisation

- ▶ 4 ingrédients de base
 - ▶ Interactions multiples
 - ▶ Feedback positif (Amplification / Renforcement)
 - ▶ Feedback négatif (Atténuation / Contrainte)
 - ▶ Amplification des fluctuations
- ▶ 3 signatures caractéristiques
 - ▶ Création de structures spatio-temporelles dans un médium initialement homogène
 - ▶ Multi-stabilité : possibilité de plusieurs états stables atteignables
 - ▶ Existence de bifurcations paramétrables déterminées

Paradigmes naturo-inspirés : l'auto-organisation

Description de la dynamique [Palut 2004]

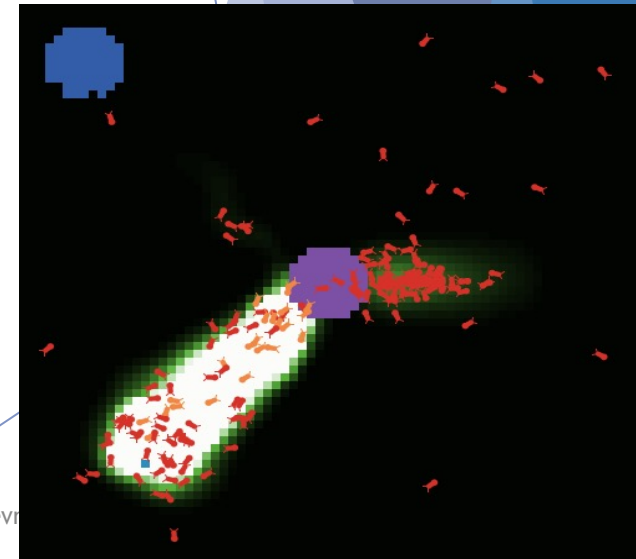
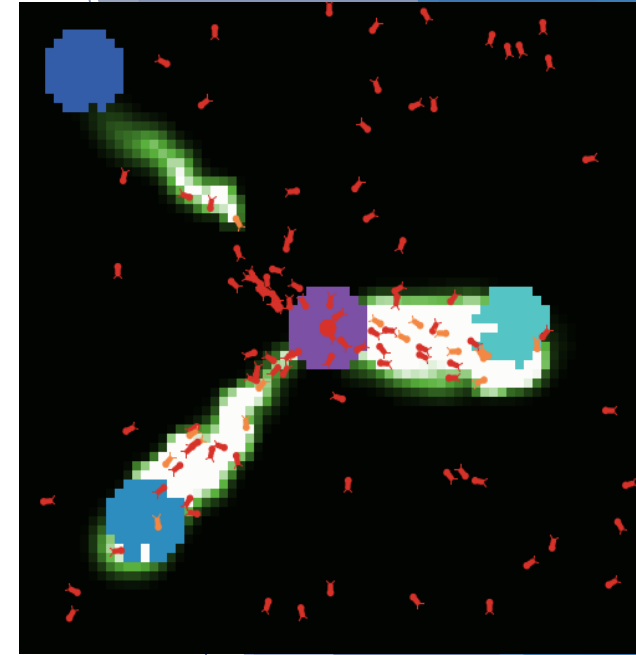
- ▶ Structures collectives décrites par 1 (ou +) variable collective X_t
 - ▶ $X_{t+1} = \text{fonction non linéaire}(X_t)$
- ▶ X_t : caractérise les relations spatio-temporelles entre composants
 - ▶ indicateur macroscopique de coordination
- ▶ Variabilité des comportements micro
 - ▶ reflète des fluctuations au niveau macro de la variable collective
 - ▶ Ex: apparition du chemin phéromonal



Paradigmes naturo-inspirés : l'auto-organisation

Description de la dynamique [Palut 2004]

- ▶ Structures collectives décrites par 1 (ou +) variable collective X_t
 - ▶ $X_{t+1} = \text{fonction non linéaire}(X_t)$
- ▶ X_t : caractérise les relations spatio-temporelles entre composants
 - ▶ indicateur macroscopique de coordination
- ▶ **Perturbations de la structure collective**
 - ▶ fluctuations de la variable collective
 - ▶ Ex: disparition de la source



Paradigmes naturo-inspirés : l'auto-organisation

Stabilité, Equilibre, Transition de Phase

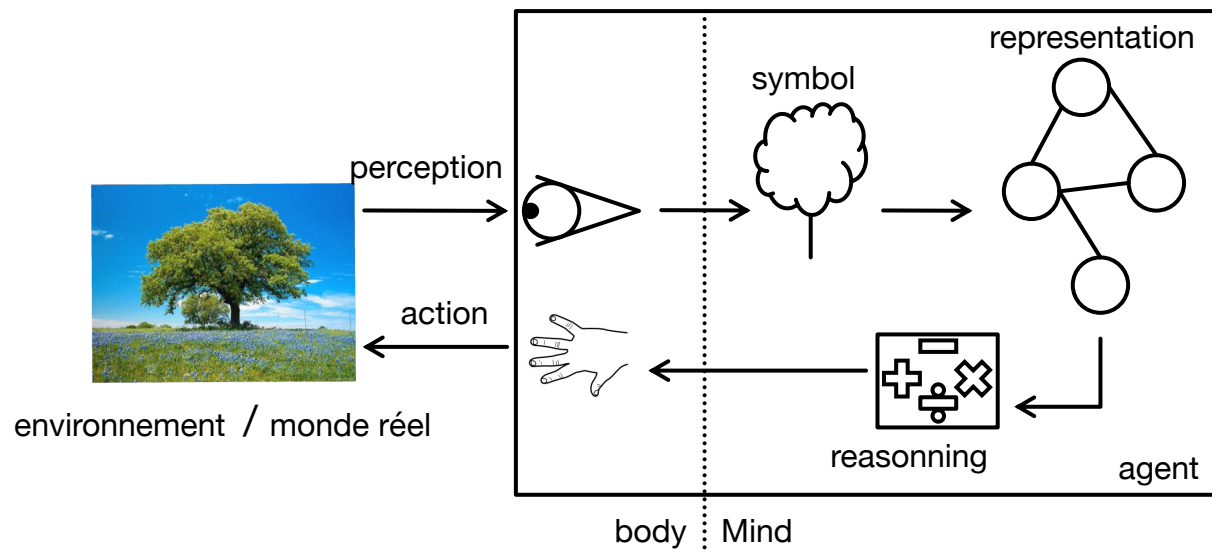
- ▶ Si X_t = constante ou peu variable au cours du temps
 - ▶ état d'équilibre
- ▶ 2 états d'équilibre
 - ▶ Attracteur : état dans lequel converge le système quelque soit les conditions de départ
 - ▶ Répulseur : état dans lequel le système ne peut se maintenir (sauf force externe)
- ▶ Transition de phase
 - ▶ Passage d'un état stable à un autre état stable
 - ▶ sous la variation d'un paramètre de contrôle
 - ▶ ou sous l'influence des fluctuations inhérentes au système

Approches émergentes

- ▶ Auto-organisation
 - ▶ Activateur à courte portée
 - ▶ Auto-catalyse : promouvoir ses propres actions
 - ▶ Augmenter la production d'un inhibiteur (antagoniste)
 - ▶ Inhibiteur
 - ▶ Diffuse rapidement
 - ▶ Résultat
 - ▶ Augmentation locale de l'activation
 - ▶ Inhibition antagoniste à longue portée qui permet de localiser la réaction d'auto-amélioration

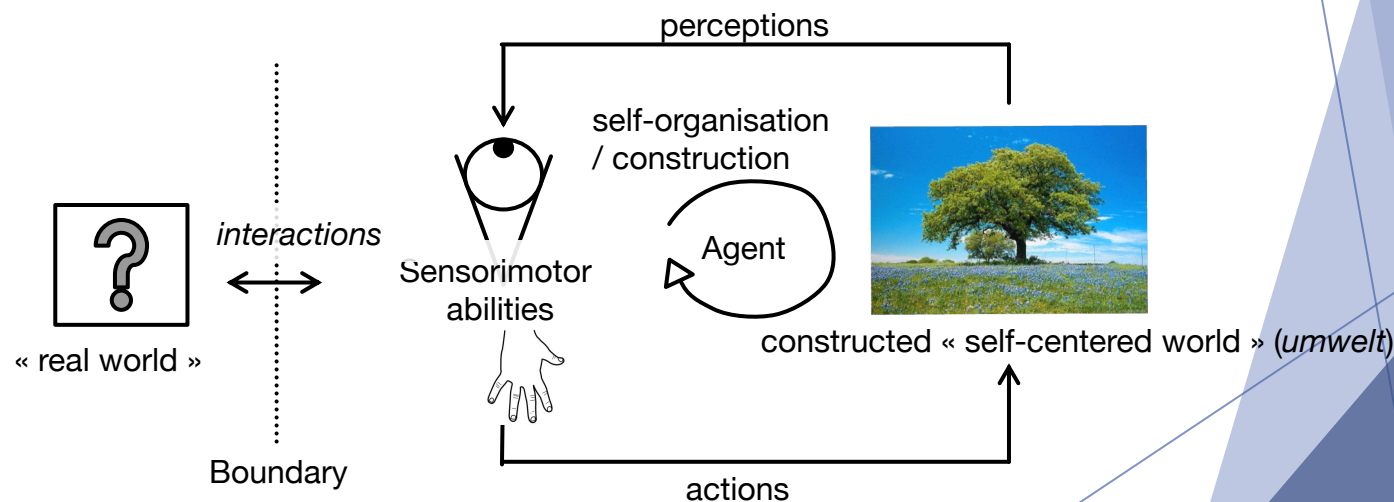
Cognitivist Approaches

- *Moravec Paradox*
- *Grounding Symbol Problem*

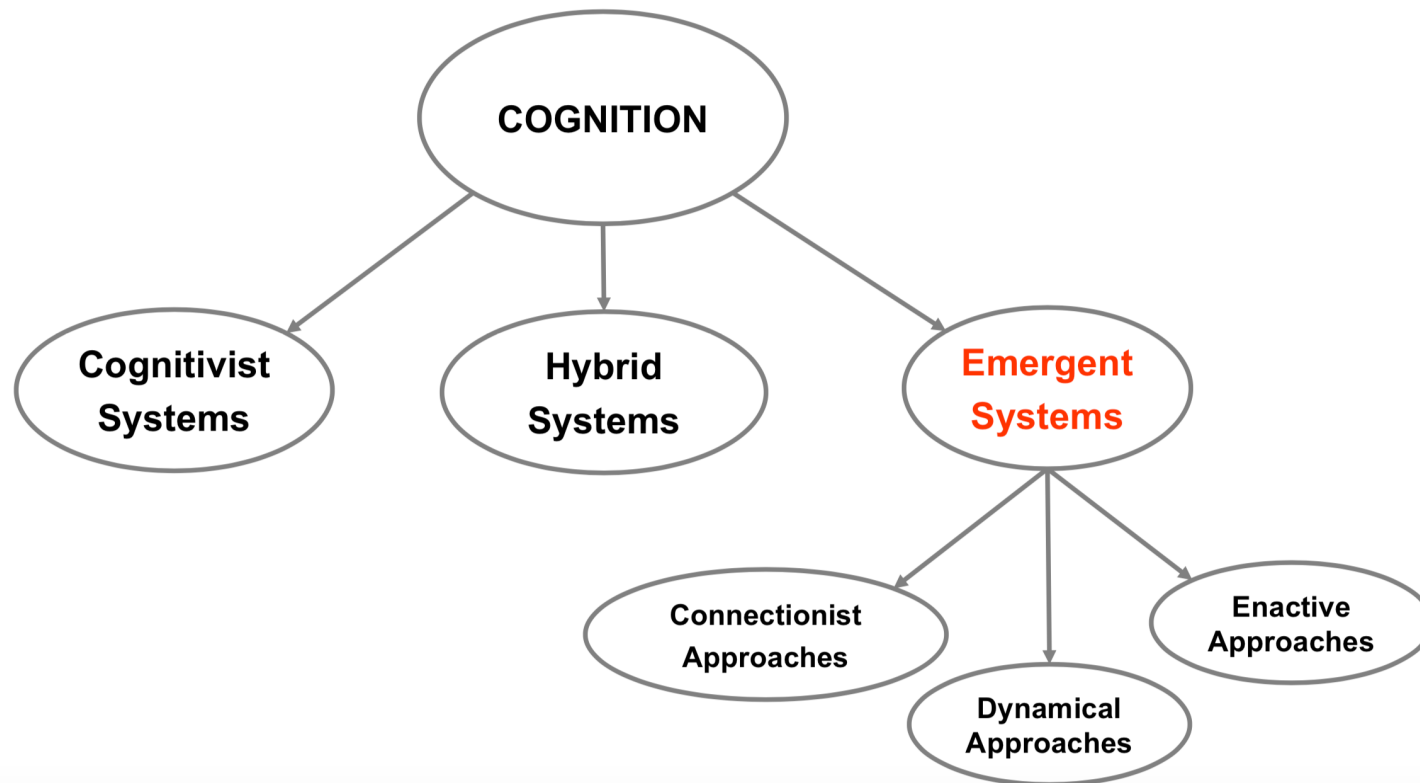


Constructivist Approaches

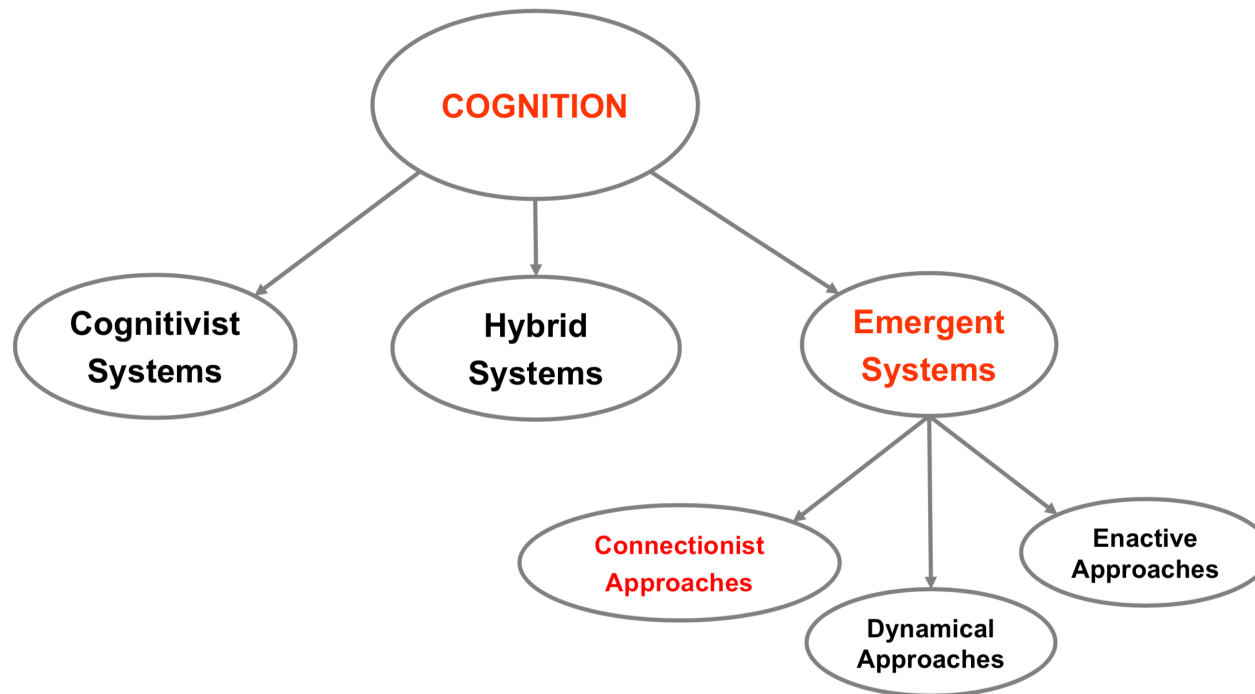
- *Embodiment* : Embodied AI
- *Developmental Robotics*



Approches émergentes : des systèmes connexionnistes, dynamiques et éenactifs



Approches émergentes : systèmes connexionnistes



Approches émergentes : systèmes connexionnistes

- ▶ Le connexionnisme
 - ▶ Approche utilisée en sciences cognitives, neurosciences, psychologie et philosophie de l'esprit
 - ▶ Théorie proposée à la fin des années 1980 en tant qu'alternative au computationnalisme
 - ▶ Modélise les phénomènes mentaux ou comportementaux comme des processus émergents de réseaux d'unités simples interconnectées
- ▶ Les modèles connexionnistes utilisent des réseaux de neurones formels :
 - ▶ Organisation et fonctionnement : abstraction des systèmes neuronaux physiologiques
 - ▶ Cherchent à faire le lien entre le fonctionnement du cerveau et celui de l'esprit

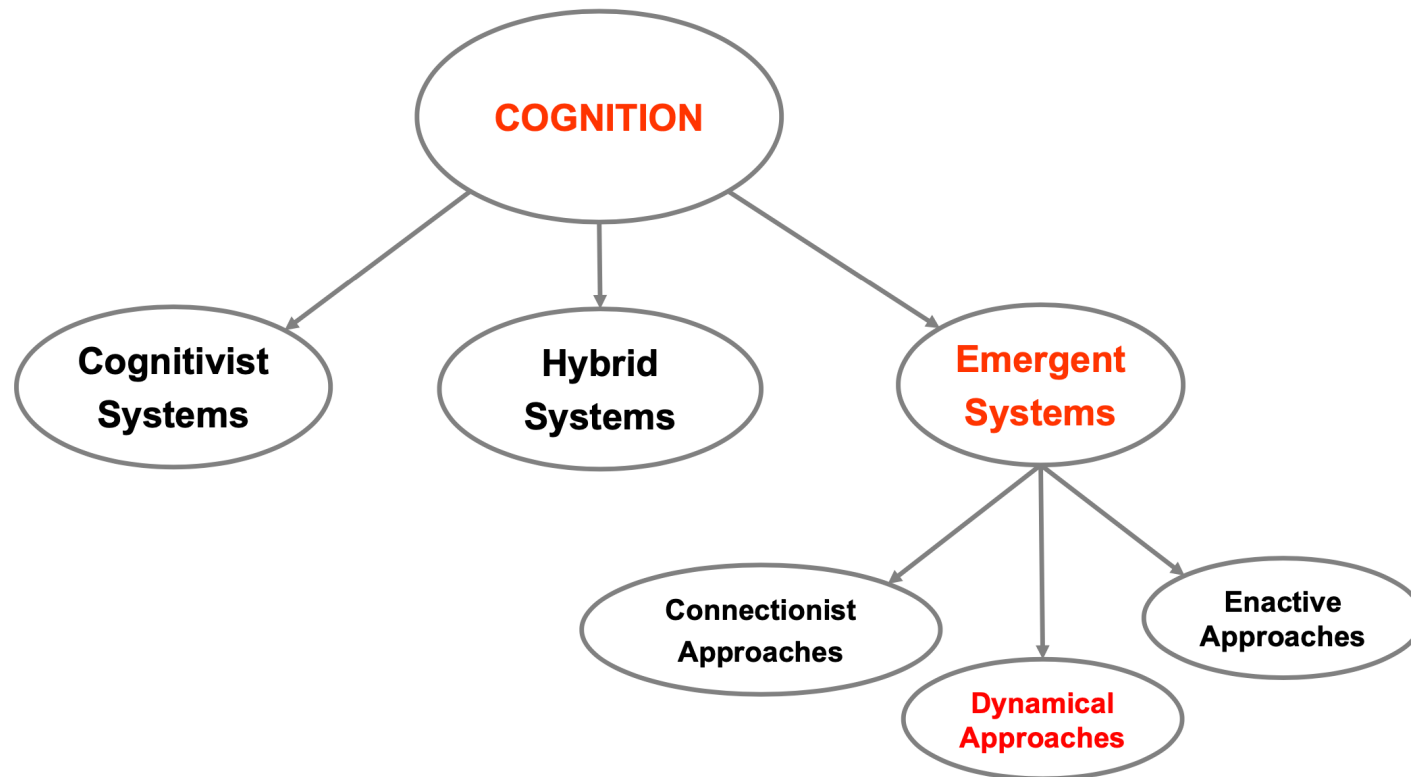
Approches émergentes : systèmes connexionnistes

- ▶ Principes de base
 - ▶ Les phénomènes mentaux peuvent être décrits à l'aide de réseaux d'unités simples interconnectées
 - ▶ La forme des connexions et des unités peut varier selon les modèles
 - ▶ Ex : unités du réseau = des neurones, connexions = des synapses
 - ▶ Ex : unités du réseau = des mots, connexions = la similarité sémantique
- ▶ Concepts clés : l'activation et la diffusion
 - ▶ Traitement parallèle de l'information
 - ▶ Modèle d'activation distribués non symboliques, règles non logiques
 - ▶ À chaque instant, une unité du réseau a un certain niveau d'activation
 - ▶ Ex: si les unités du modèle sont des neurones, l'activation pourrait représenter la probabilité que le neurone génère un potentiel d'action. Si ce modèle est un modèle à diffusion d'activation, alors une fois le seuil d'activation atteint, l'activation va se propager à toutes les unités connectées.

Approches émergentes : systèmes connexionnistes

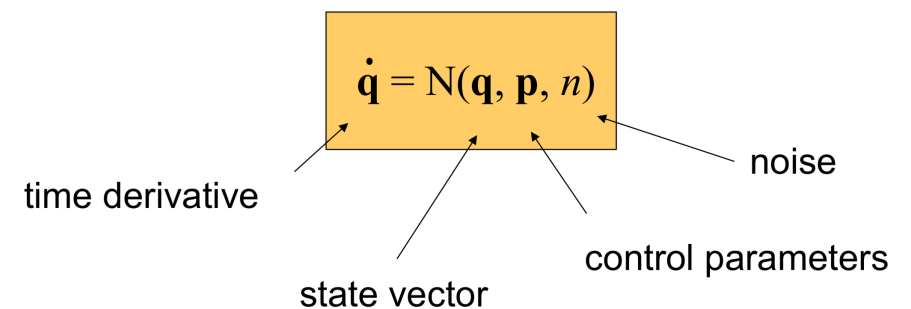
- ▶ Réseaux neuronaux : les modèles connexionnistes les plus répandus
 - ▶ Systèmes dynamiques qui capturent les régularités ou les associations statistiques
- ▶ Les réseaux connexionnistes répondent à deux caractéristiques :
 - ▶ Chaque état mental peut être représenté comme un vecteur à n dimensions représentant les valeurs d'activation des unités neuronales
 - ▶ Le réseau peut apprendre en modifiant le poids des connexions entre ses unités. La force des connexions (ou poids) est généralement représentée comme une matrice à n dimensions
- ▶ Les modèles peuvent différer de par :
 - ▶ **La définition de l'activation** : l'activation peut être définie de différentes façons (seuil d'activation, fonction d'activation...)
 - ▶ **L'algorithme d'apprentissage** : définit la façon dont les poids des connexions vont changer au fur et à mesure du temps

Approches émergentes : systèmes dynamiques

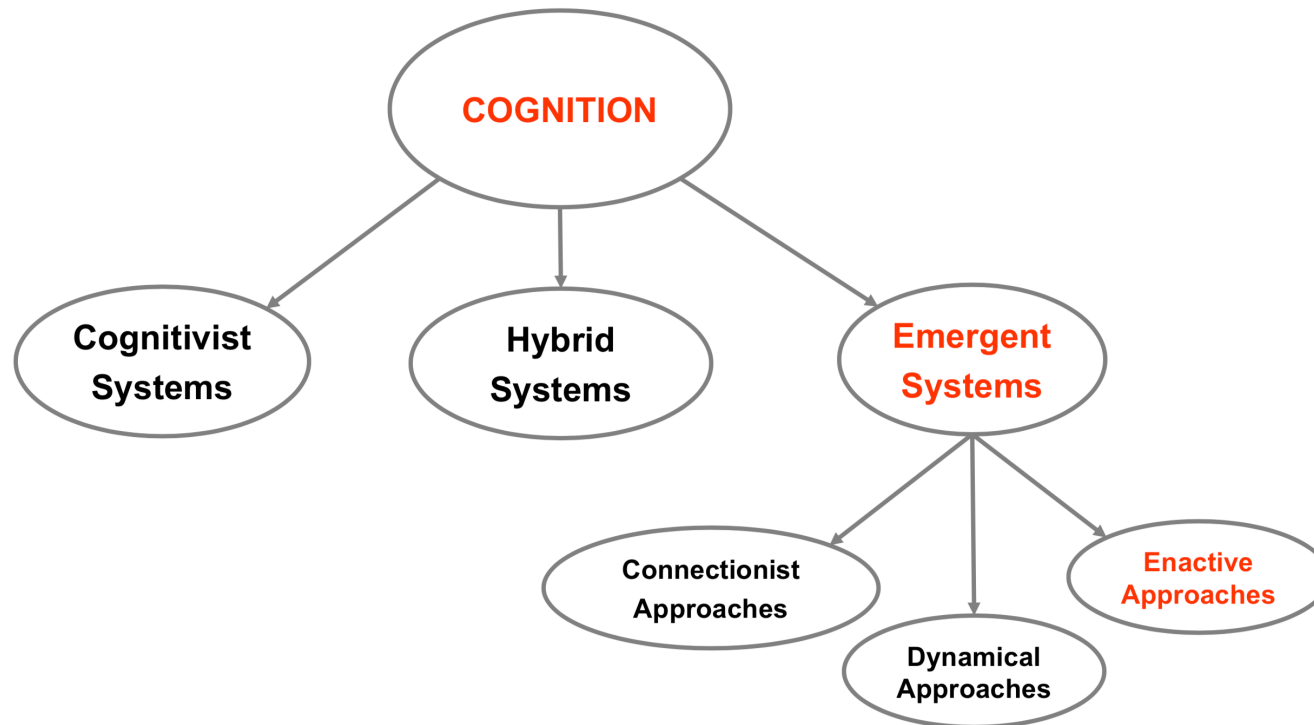


Approches émergentes : systèmes dynamiques

- ▶ Un système dynamique est un système dissipatif ouvert, non linéaire et non équilibré
 - ▶ Système : un grand nombre de composants en interaction et un grand nombre de degrés de liberté
 - ▶ Dissipatif : à énergie diffuse, i.e. l'énergie diminue au cours du temps
 - ▶ Non-équilibré : incapable de maintenir la structure sans sources externes d'énergie, de matériel, d'information (donc ouvert)
 - ▶ Non-linéaire : la dissipation de l'énergie n'est pas uniforme, et un petit nombre de degrés de liberté du système contribuent au comportement



Approches émergentes : systèmes énaectifs



L'énaction ou l'énactivisme

- ▶ Manière de concevoir la cognition qui met l'accent sur la manière dont les organismes et esprits humains s'organisent eux-mêmes en interaction avec l'environnement
- ▶ Approche théorique de la cognition proche de la cognition située et de la cognition incarnée (**embodiment**)
 - ▶ Manière dont un agent (humain ou artificiel) évalue un objet
 - ▶ En fonction de la position de son corps et/ou de ses mouvements
 - ▶ En fonction de ses expériences sensori-motrices
- ▶ Conçue comme une alternative au cognitivisme et au computationnalisme
 - ▶ Le monde tel que le système l'expérimente est indépendant du système cognitif



ATELIER DEDIÉ

L'énaction

- ▶ 5 éléments clefs des systèmes énatifs
 - ▶ Autonomie
 - ▶ Incarnation
 - ▶ Emergence
 - ▶ Expérience
 - ▶ Création de sens

L'énaction

- ▶ L'autonomie
 - ▶ Auto-entretien du système
 - ▶ Homéostasie : stabilisation, réglage de ses caractéristiques
 - ▶ Ex : chez les êtres vivants, réglages des caractéristiques physiologiques comme la pression artérielle, la température, etc.
 - ▶ Non contrôlé par des entités extérieures
 - ▶ Se distingue de son environnement

L'énaction

- ▶ L'incarnation
 - ▶ Existe en tant qu'entité physique
 - ▶ Interagit directement avec son environnement
 - ▶ Couplage structurel
 - ▶ Perturbation mutuelle
 - ▶ Constitue une part du processus cognitif
- ▶ L'émergence
 - ▶ Le comportement cognitif découle d'un jeu dynamique entre les composants
 - ▶ Dynamique interne
 - ▶ Maintient de l'autonomie
 - ▶ Conditionne les expériences du système à travers son incarnation

L'énaction

- ▶ Les expériences
 - ▶ Histoire de l'interaction avec le monde
 - ▶ Les interactions ne contrôlent pas les expériences
 - ▶ L'interaction déclenchent des changements dans l'état du système
 - ▶ Les changements sont déterminés structurellement
 - ▶ Phylogénie (liens d'évolution / de parenté des agents)
 - ▶ Couple structurel

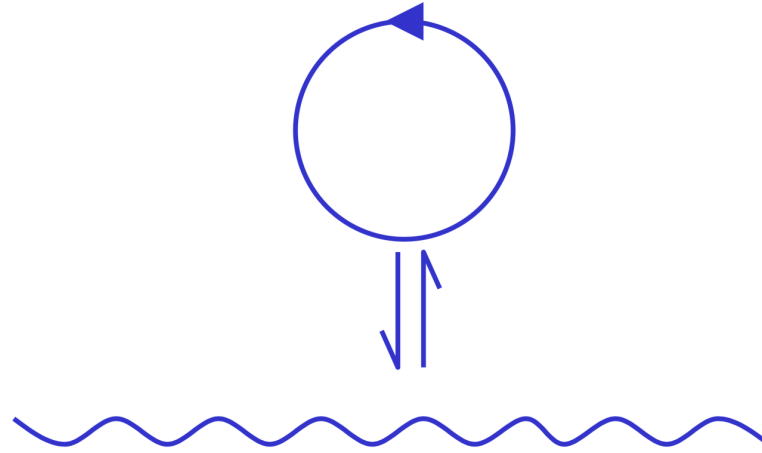
L'énaction

- ▶ La création de sens
 - ▶ La connaissance est générée par le système lui-même
 - ▶ *via* l'observation d'une certaine régularité ou légalité dans les interactions
 - ▶ Le « sens » dépend de la façon dont l'interaction peut se dérouler
 - ▶ Importance de la perception et des actions
 - ▶ Permet de modifier son propre état pour améliorer
 - ▶ La capacité prédictive
 - ▶ Les possibilités d'actions
 - ▶ **Processus d'auto-modification autonome génératif**

L'énaction : création du sens

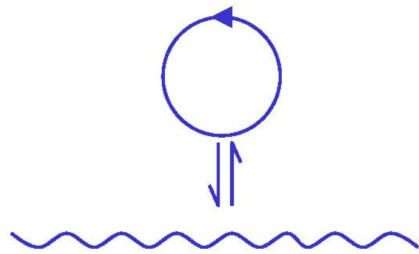
- ▶ Acquisition autogénique progressive des capacités anticipatoires
 - ▶ La cognition ne peut pas court-circuiter l'ontogène
 - ▶ Priorité au développement de l'agent
 - ▶ **Nécessairement** le produit d'un processus de **développement incarné**
 - ▶ Il faut d'abord traiter les événements immédiats
 - ▶ Puis acquérir de plus en plus de capacités prédictives
- ▶ La cognition et la perception dépendent fonctionnellement de la richesse d'interaction avec l'environnement

L'énaction : création du sens

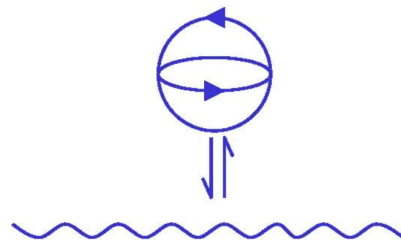


- ▶ Co-détermination / couplage structurel
- ▶ Préservation de l'autonomie *via* les interactions mutuelles entre le système et l'environnement
- ▶ La perturbation du système n'est effectuée que par l'environnement

L'énaction : création du sens



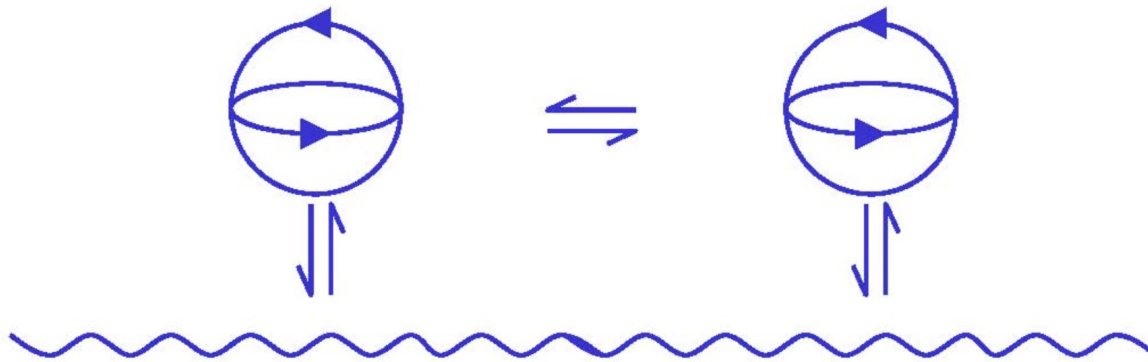
→ t



→ t

Anticipation / Planning / Explanation / Prediction

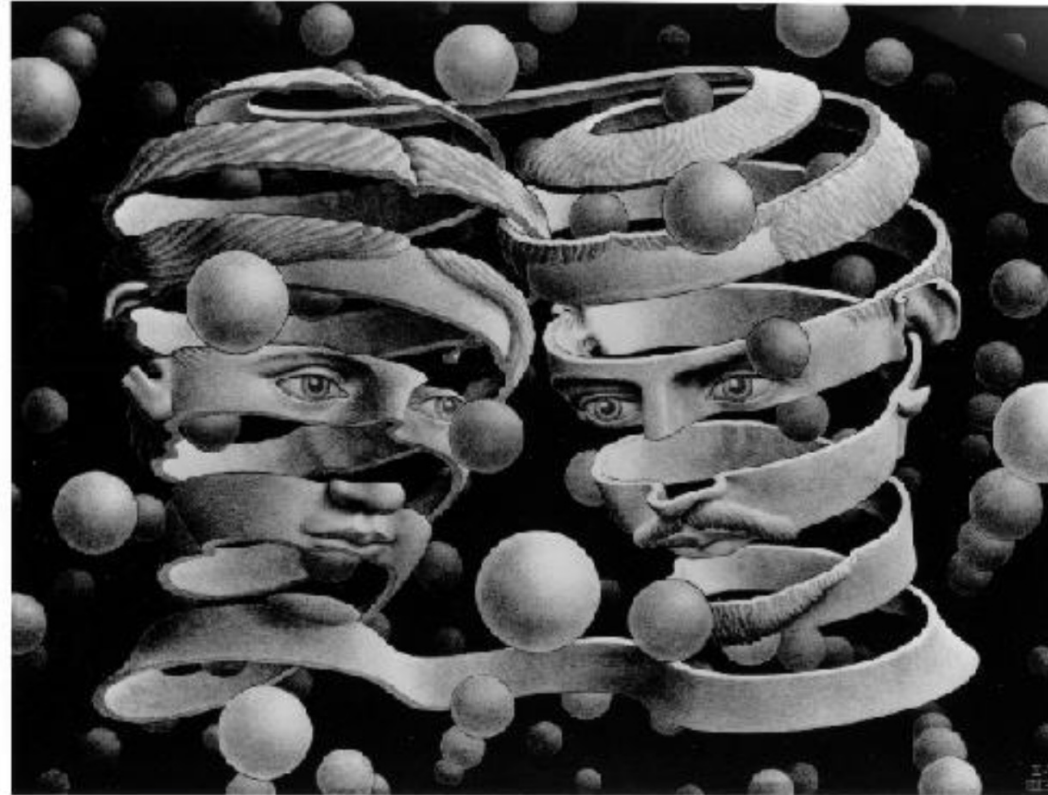
L'énaction : l'interaction



- ▶ Une activité partagée dans laquelle les actions de chaque agent influencent les actions des autres agents
- ▶ Résultant d'un modèle de comportement partagé mutuellement construit

L'énaction : l'interaction

- ▶ Le sens peut émerger d'une expérience consensuelle partagée médiée par l'interaction



Bond of Union
M. C. Escher, 1956

“Interaction is a shared activity in which the actions of each agent influence the actions of the other agents engaged in the same interaction, resulting in a mutually-constructed patterns of shared behaviour”

Ogden, Dautenhahn, Stribling 2002

Differences between Cognitivist & Emergent Paradigms



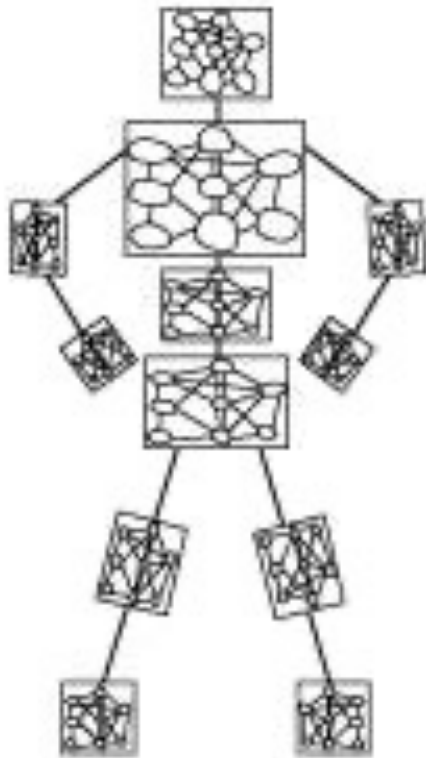
1. Computational operation
2. Representational framework
3. Semantic grounding
4. Temporal constraints
5. Inter-agent epistemology
6. Embodiment
7. Perception
8. Action
9. Anticipation
10. Adaptation
11. Motivation
12. Autonomy
13. Cognition
14. Philosophical foundation

[Vernon, Von Hofsten, Fadiga 2010]

The Cognitivist Paradigm vs. the Emergent Paradigm

Characteristic	Cognitivist	Emergent
Computational Operation	Syntactic manipulation of symbols	Concurrent self-organization of a network
Representational Framework	Patterns of symbol tokens	Global system states
Semantic Grounding	Percept-symbol association	Skill construction
Temporal Constraints	Atemporal	Synchronous real-time entrainment
Inter-agent epistemology	Agent-independent	Agent-dependent
Embodiment	No role implied: functionalist	Direct constitutive role: non-functionalist
Perception	Abstract symbolic representations	Perturbation by the environment
Action	Causal consequence of symbol manipulation	Perturbation by the system
Anticipation	Procedural or probabilistic reasoning	Traverse of perception-action state space
Adaptation	Learn new knowledge	Develop new dynamics
Motivation	Criteria for goal selection	Increase space of interaction
Autonomy	Not entailed	Cognition entails autonomy
Cognition	Rational goal-achievement	Self-maintenance and self-development
Philosophical Foundation	Positivism	Phenomenology

Symbolic vs. Analogical Man: Top-Down vs. Bottom Up [*Minsky*]



Pour aller plus loin

- ▶ Quelques pistes de lecture sur la représentation et le contrôle dans une approche constructiviste
 - ▶ [Intelligence without representation](#) (R. Brooks)
 - ▶ [Philosophy of Artificial Intelligence: Neats vs. Scruffies](#) (M. Minsky)
 - ▶ [Artificial Intelligence Meets Natural Stupidity](#) (D. McDermott)
- ▶ Sources pour construire ce cours
 - ▶ Les cours de Salima Hassas et David Vernon