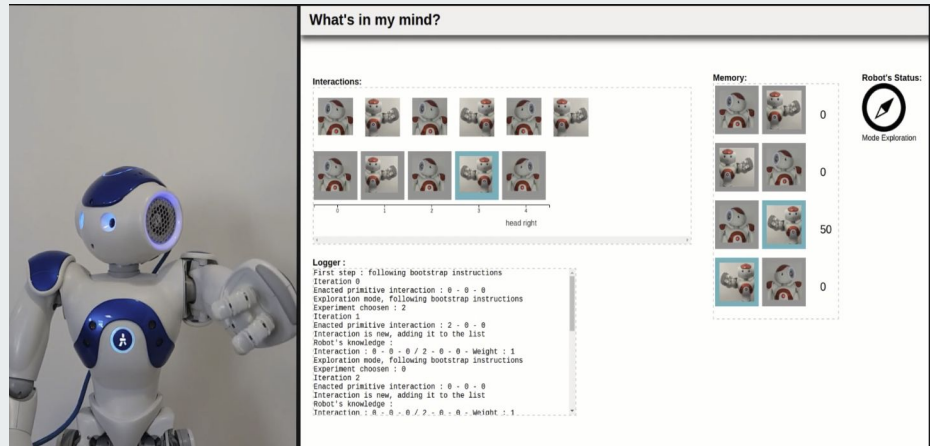


# Introduction à l'intelligence artificielle développementale



**What's in my mind?**

**Interactions:**

1	1	2	3	4	

head right

**Memory:**

		0
		0
		50
		0

**Robot's Status:**

Mote Exploration

**Logger:**

```
First step : following bootstrap instructions
Iteration 0
Enacted primitive interaction : 0 - 0 - 0
Exploration mode, following bootstrap instructions
Experiment chosen : 2
Iteration 1
Enacted primitive interaction : 2 - 0 - 0
Interaction is new, adding it to the list
Robot's knowledge :
Interaction : 0 - 0 / 2 - 0 - 0 - weight : 1
Exploration mode, following bootstrap instructions
Experiment chosen : 0
Iteration 2
Enacted primitive interaction : 0 - 0 - 0
Interaction is new, adding it to the list
Robot's knowledge :
Interaction : 0 - 0 - 0 / 2 - 0 - 0 - weight : 1
```



# Remerciements

Ce cours est très très très très fortement inspiré du cours d'Olivier Georgeon accessible à l'adresse ci-dessous

<http://www.oliviergeorgeon.com/radical-interactionism/developmental-learning.htm>



# Bibliographie

Le MOOC sur l'apprentissage développemental (EN) : <https://liris.cnrs.fr/ideal/mooc>

Le site du projet IDEAL (*Implementing DEvelopmentAI Learning*) : <https://liris.cnrs.fr/ideal>

Une FAQ (que vous pouvez enrichir) : <http://www.oliviergeorgeon.com/radical-interactionism/faq.htm>

Expérimentez pas vous-mêmes le *small loop problem* : <http://liris.cnrs.fr/ideal/demo/small-loop/>

Le blog scientifique associé au projet : <http://e-ernest.blogspot.fr/>

Quelques contributions d'étudiants stagiaires dans la suite du projet : <http://liris.cnrs.fr/devlearn/>

*Biologically Inspired Cognitive architectures* : <http://bicasociety.org/>



# À l'issue de ce cours, vous saurez...

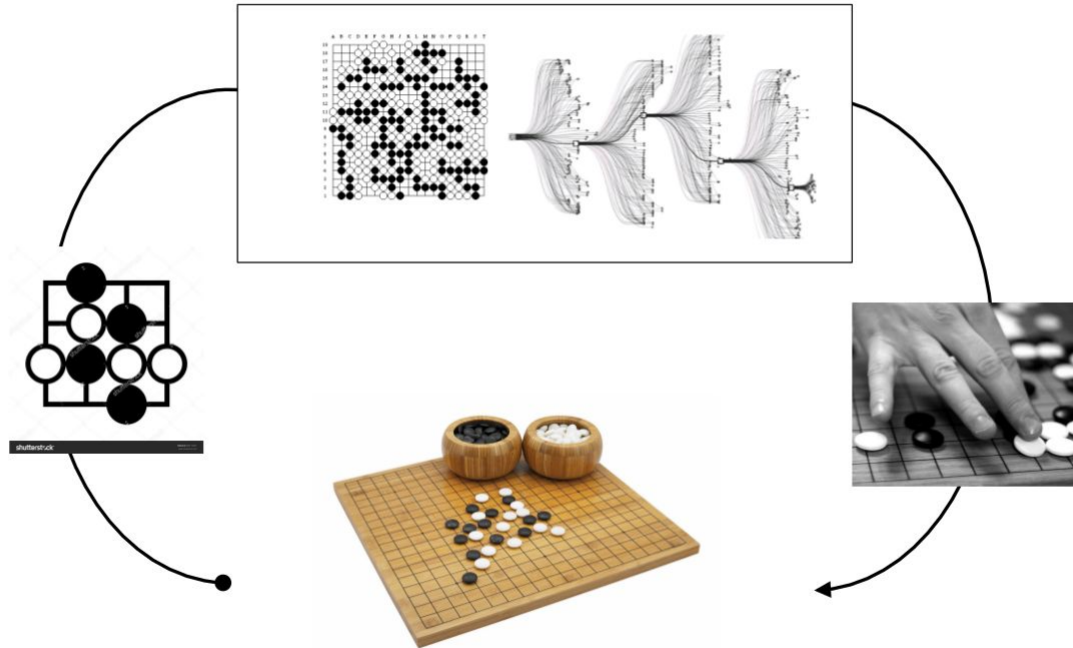
Faire la **différence** entre une IA en environnement modélisé a priori et en environnement non modélisé a priori

Donner une **définition de l'Intelligence Artificielle Développementale** (IAD)

- Citer d'autres disciplines scientifiques, sources d'inspiration pour les recherches en IAD
- Expliquer certains concepts clés qui fondent l'IAD
- Nommer les principaux contributeurs aux origines de ces concepts clés

**Implémenter un agent** capable d'apprentissage rudimentaire sans connaissance a priori du domaine

# IA dans un domaine modélisé a priori





# Modélisation du domaine

Selon Newell et Simon (1976). *Computer Science as empirical inquiry: Symbols and Search*.

*“A physical symbol system exercises its intelligence in problem solving by search - that is, by generating and progressively modifying symbol structures until it produces a solution structure.”*

Conséquence : on doit pouvoir construire un **espace de problèmes** et un **but** (ou état objectif)



## Modélisé vs. non modélisé

	Modélisé	Non modélisé
Conception	Espace de problèmes Objectif à atteindre	Domaine interactionnel Système de valeurs
Application	Atteindre l'objectif <b>Résoudre un problème</b>	<b>Bonne question !</b>
Évaluation	Maximiser une valeur Minimiser un temps d'exécution Trouver une solution	Évaluation par un humain Observation / analyse <b>Trace d'activité</b>

# L'IAD

intelligence artificielle développementale





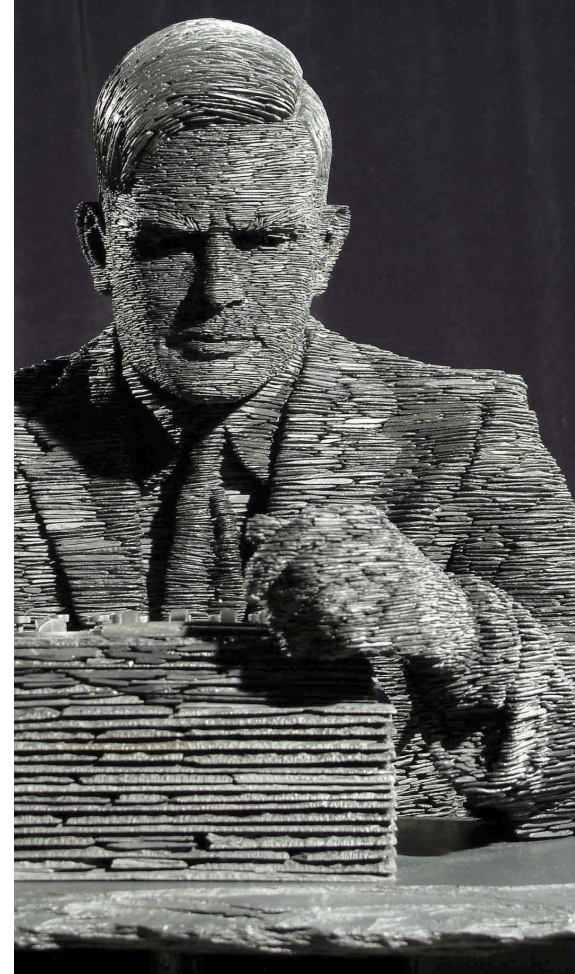
---

## Une idée pas si nouvelle...

*Instead of trying to produce a program to simulate the adult mind, why not rather **try to produce one which simulates the child's**? If this were then subjected to an appropriate course of education one would obtain the adult brain.*

*Presumably, the child brain is something like a notebook [...]. Rather little mechanism, and lots of blank sheets. [...]. **Our hope is that there is so little mechanism in the child brain that something like it can be easily programmed.** The amount of work in the education we can assume, as a first approximation, to be much the same as for the human child.*

**Computing machinery and intelligence**  
(Alan Turing, 1950, Mind, philosophy journal)





# Est-ce possible ?

Petit sondage...





# Est-ce possible ?

## Les arguments en faveur du **NON** :

- Théories spiritualistes de la conscience
- Théories de l'ouverture causale de la réalité physique
- Trop complexe
- On ne peut pas programmer la subjectivité
- Comment représenter les émotions ?
- ...

## Les arguments en faveur du **OUI** :

- Théorie matérialiste de la conscience (Julien Offray de la Mettrie, 1709-1751)
- La conscience comme processus computationnel (Chalmers, 1994 [1])

[1] <http://consc.net/papers/computation.html>



# Quelle démarche pour comprendre ?

**Observer** des exemples d'agents développementaux

Mettre en place des outils théoriques :

- **Formuler** clairement le problème
- S'interroger sur le concept d'**auto-programmation**

Travaux pratiques pour **développer votre propre agent**

# Exemple 1. Couplage agent / environnement

6 expériences



2 résultats

0 1

0

0

1 0

0 1

0 1

10 interactions (avec valeurs)

$i_1(5)$   $i_2(-10)$

$i_3(-3)$

$i_4(-3)$

$i_5(-1)$   $i_6(-1)$

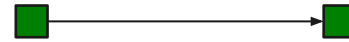
$i_7(-1)$   $i_8(-1)$

$i_9(-1)$   $i_{10}(-1)$

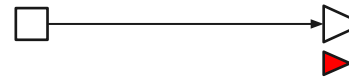
# Exemple 1. Régularités d'interactions

L'environnement présente des régularités séquentielles d'interactions :

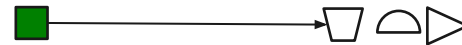
- Après  $i_8$ ,  $i_8$  peut à nouveau être effectué



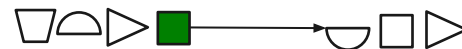
- Après  $i_7$ , essayer  $i_1$  ou  $i_2$  provoque plus souvent  $i_1$  que  $i_2$



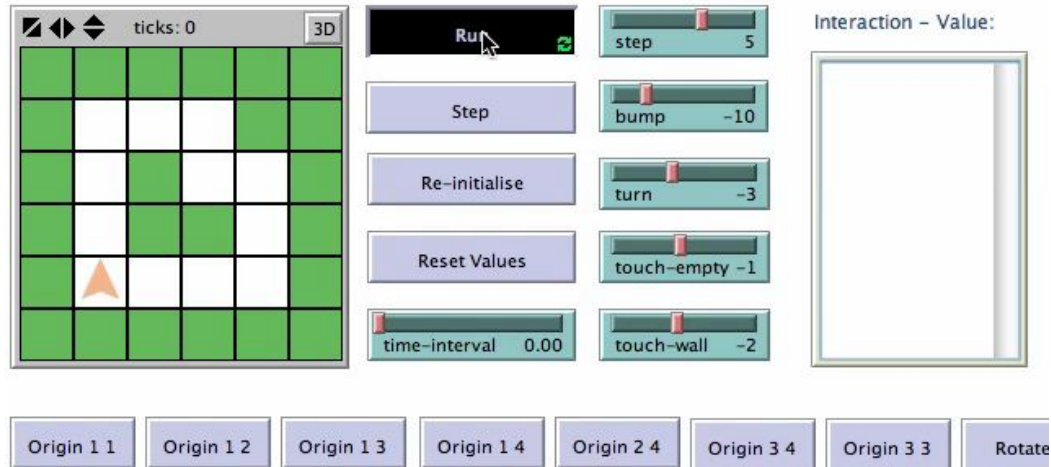
- Après  $i_8$ , la séquence  $\langle i_9, i_3, i_1 \rangle$  peut souvent être effectuée



- Après  $\langle i_9, i_3, i_1, i_8 \rangle$ ,  $\langle i_4, i_7, i_1 \rangle$  peut souvent être effectuée



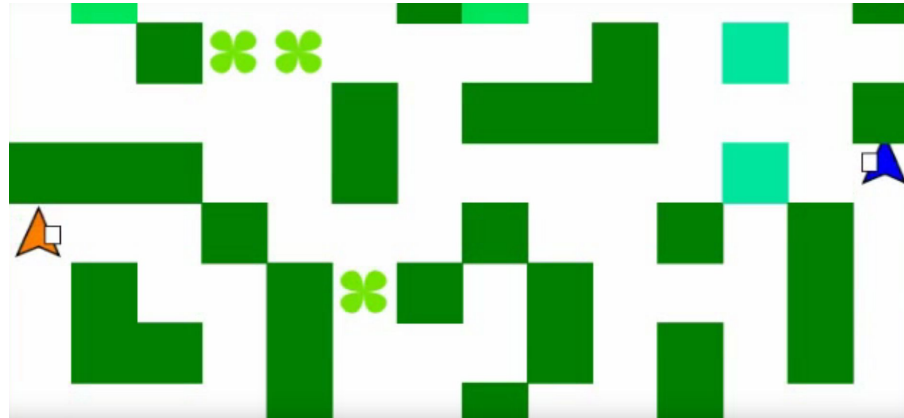
# Exemple 1. Babillage



<https://www.youtube.com/watch?v=MstI8BRCbSI>



## Exemple 2. Différentiation



<https://www.youtube.com/watch?v=q8WkYbt2BxM>





# Outils théoriques

Philosophie de l'esprit

Epistémologie (théorie de la connaissance)

Psychologie développementale

Biologie (autopoïèse, enaction)

Neurosciences



# Arguments philosophiques

John Locke (1632-1704). **Tabula Rasa**

La Mettrie (1709-1751). **La matière peut penser**

David Chalmers. **A Computational Foundation for the Study of Cognition** (1994)

Daniel Dennett. Consciousness explained (1991), **libre arbitre**, choix individuel, motivation propre, **déterminisme**





# Concepts philosophiques clés

Considérer la **cognition** comme une **computation**, au sens large

- Dispositif causal : exemple des réseaux de neurones (chimie, neurotransmetteurs, hormones, etc.)

Le déterminisme n'est pas incompatible avec le libre arbitre

- Ne pas confondre **déterminisme** et **prédictibilité** (Hervé Zwirn, Les systèmes complexes, 2006)

# Épistémologie (que peut-on connaître ?)

À propos des **ontologies**

L'étude de l'être en tant qu'être

- Aristote (384 - 322 av. JC)
- **Onto** (étant) / **Logos** (discours)
- C'est l'étude des propriétés générales de tout ce qui est

La réalité en tant que telle est **inconnaissable**

- Emmanuel Kant (1724 - 1804)
- David Eagleman - *What is reality?* [1]

[1] <https://www.youtube.com/watch?v=3MSw2irv0-A>





# Concepts épistémologiques clés

Implémenter des systèmes **sans présupposés ontologiques**

- Agents **agnostiques** (Georgeon 2012)
- La perception qu'a l'agent de l'environnement est différente de la nôtre

Implémenter des systèmes **avec des présupposés interactionnels**

- Prédéfinir les possibilités d'interaction entre l'agent et son environnement
- Laisser l'agent construire sa propre ontologie en fonction de ses **expériences d'interaction**

# Psychologie développementale (comment puis-je connaître ?)

**Apprentissage développemental** - Jean Piaget (1896 - 1980)

- Téléologie / principes motivationnels
- L'individu s'auto-finalise de manière réursive
- Il ne faut pas séparer perception et action a priori
- Notion de schème sensori-moteur

**Épistémologie constructiviste**

- Jean-Louis Le Moigne (1931 - ) - Le constructivisme et les épistémologies constructivistes
- Ernst von Glasersfeld - La connaissance est une adaptation fonctionnelle



# Inspiration chez les (jeunes) humains



<https://www.youtube.com/watch?v=8vNxjw2AqY>



# Étapes développementales (approximatives)

4<sup>ème</sup> mois : prédictions Bayésiennes

5<sup>ème</sup> mois : modèles des mouvements des mains

6<sup>ème</sup> mois : reconnaissance des objets et des visages

7<sup>ème</sup> mois : persistance des objets

8<sup>ème</sup> mois : modèles dynamiques des objets

9<sup>ème</sup> mois : usage d'outils, pointage d'objets

10<sup>ème</sup> mois : imite les mouvements, rampe

11<sup>ème</sup> mois : marche avec l'aide d'un adulte

15<sup>ème</sup> mois : marche seul

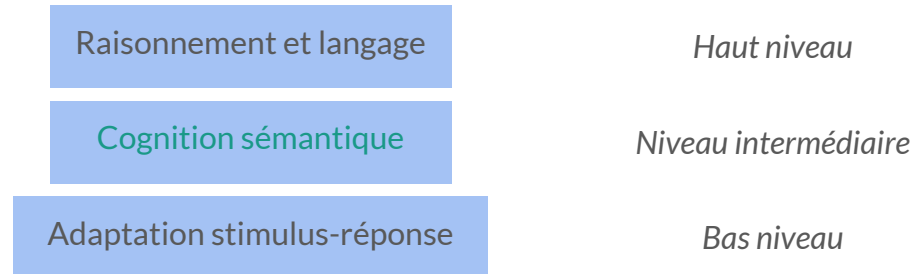




# Concepts psychologiques clés

Raisonner sur les **interactions** plutôt que de séparer perception et action

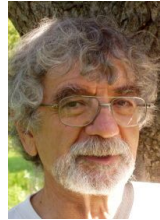
Définir un niveau “intermédiaire” d’intelligence, comme l’idée de **cognition sémantique** de (Manzotti et Chella, 2012)



# Théorie de la vie (pourquoi connaître ?)

## Autopoïèse (création de soi)

- Maturana 1972
- Couplage structurel agent / environnement
- **Domaine relationnel** (espace des possibilités d'interaction)



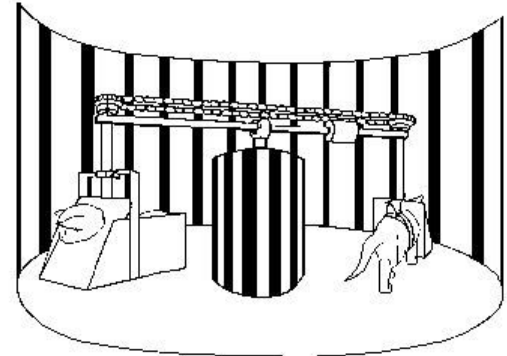
## Homéostasie

- Régulation de l'état interne
- **Motivation propre**



## Théorie de l'énaction (*to enact* / mettre en oeuvre)

- Auto-organisation avec interaction avec l'environnement
- Enactive artificial intelligence (Froese and Ziemke, 2009)





# Concepts enactivistes clés

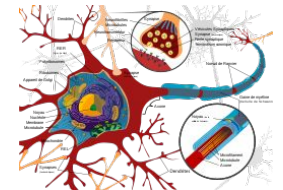
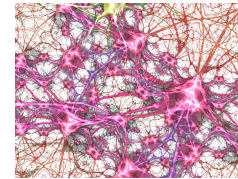
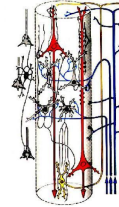
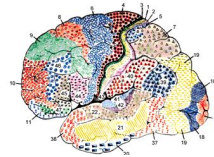
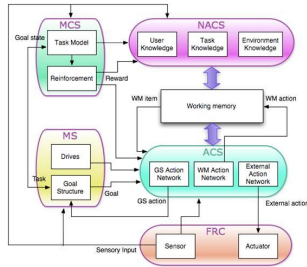
L'**autonomie constitutive** est nécessaire à la construction de sens

- Évolution des possibilités d'interaction au fil du temps et de la vie du système
- **Individuation** de la façon de (perce)voir le monde

Enjeu de la création de systèmes capables de s'**auto-programmer**

- Les “données” apprises ne sont pas des simples valeurs de paramètres ou des pondérations... il s'agit de **données exécutables**

# Inspiration des neurosciences



Multiples **niveaux d'analyse**

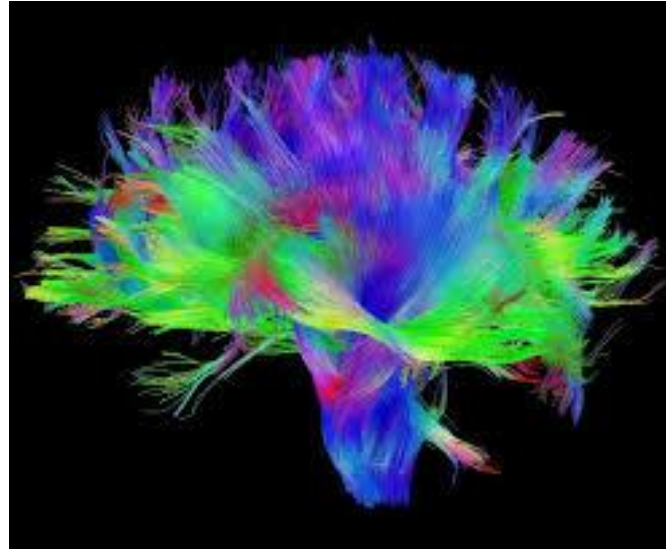
Beaucoup de **plasticité**

Beaucoup de **pré-câblage**



---

# Human Connectome Project



---

# *The Human Brain Project*

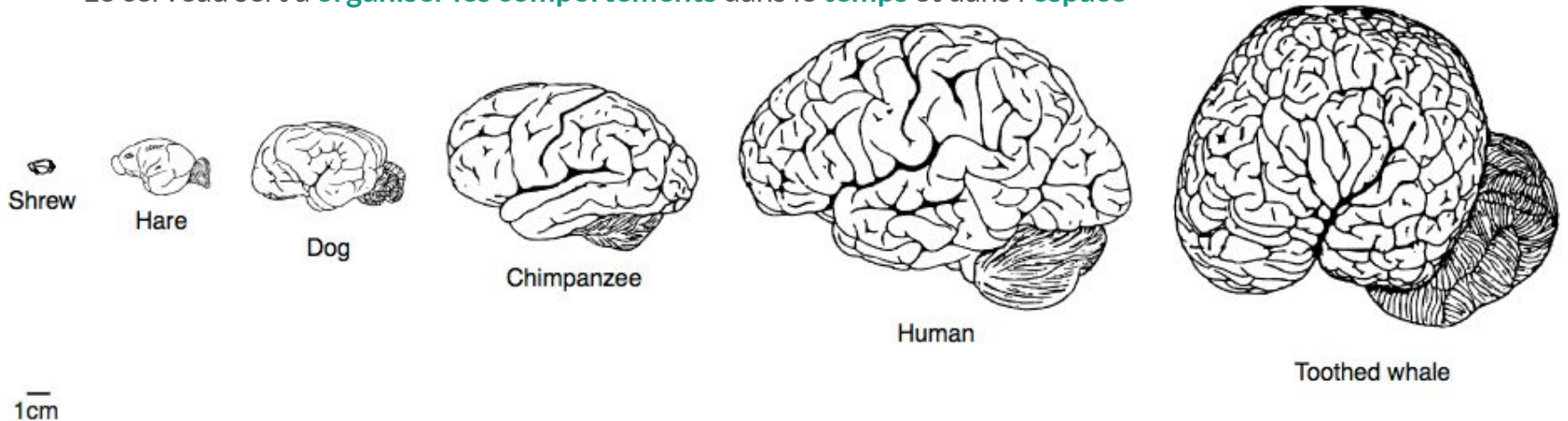


<https://www.humanbrainproject.eu/en/>

# Inspiration des neurosciences

Pas de **rupture qualitative**, mais des différences innées

Le cerveau sert à **organiser les comportements** dans le **temps** et dans l'**espace**







# Idées clés des neurosciences

Renoncer à l'idée que ce sera simple

S'interroger sur le **niveau de modélisation** auquel commencer

Le vivant peut être source d'inspiration (**architectures cognitives bio-inspirées**)

Importance de la **simulation interne de comportements**

---

## *Take away message*

Objectif : **apprendre** (découvrir, organiser et exploiter) des **régularités d'interaction** dans le temps et l'espace pour favoriser des critères innés (survie, etc.)

Contrainte : **ne pas encoder de connaissances ontologiques présumées**

Permettre une certaine autonomie constitutive par le biais de l'**auto-programmation**



# Exercice

Implémentez votre premier agent développemental





# Exercice : un premier algorithme rudimentaire

Deux expériences :  $E = \{e_1, e_2\}$

Deux résultats :  $R = \{r_1, r_2\}$

Quatre interactions :  $E \times R = \{i_{11}, i_{12}, i_{21}, i_{22}\}$

Deux environnements :

- $env_1 : e_1 \rightarrow r_1, e_2 \rightarrow r_2$   
(donc  $i_{12}$  et  $i_{21}$  ne se produisent jamais)
- $env_2 : e_1 \rightarrow r_2, e_2 \rightarrow r_1$   
(donc  $i_{11}$  et  $i_{22}$  ne se produisent jamais)

Des système motivationnel :

$$v(i_{11}) = 1, v(i_{12}) = 1, v(i_{21}) = -1, v(i_{22}) = -1$$

$$v(i_{11}) = -1, v(i_{12}) = -1, v(i_{21}) = 1, v(i_{22}) = 1$$

$$v(i_{11}) = 1, v(i_{12}) = -1, v(i_{21}) = 1, v(i_{22}) = -1$$

Consigne :

Implémenter un agent qui **apprend** à effectuer les interactions positives **sans connaître a priori** son **système motivationnel** ni son **environnement**





## Ce qu'il ne faut pas faire...

```
if (env == env2 and motiv == motiv1) or (env == env2 and motiv == motiv2)
    return e1
else
    return e2
```



## Ce à quoi cela devrait ressembler...

```
while ...  
    e = agent.chooseExperience(r)  
    r = environment.getResut(e);  
    print(e, r, value)
```

# Annexes

---



# Exemple de trace

