

17 Octobre 2019

# Pourquoi l'IA ne peut avoir ni empathie, ni créativité ni conscience

Raja Chatila

Institut des Systèmes Intelligents et de Robotique (ISIR)

Sorbonne Université, Paris, France

[Raja.Chatila@sorbonne-universite.fr](mailto:Raja.Chatila@sorbonne-universite.fr)

Membre du groupe d'experts de haut niveau sur l'IA (UE)

Membre de la CERNA, Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique (FR)

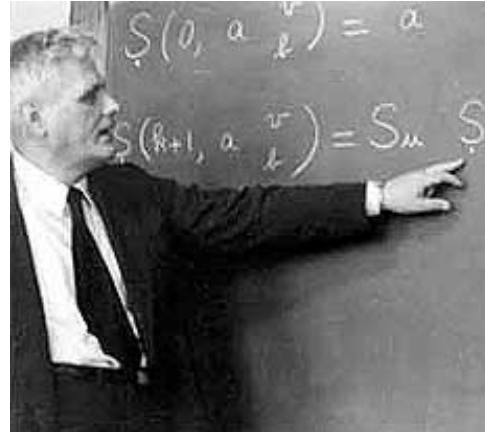
Chair, The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

# Réponse courte

Parce que l'IA est un ensemble d'algorithmes



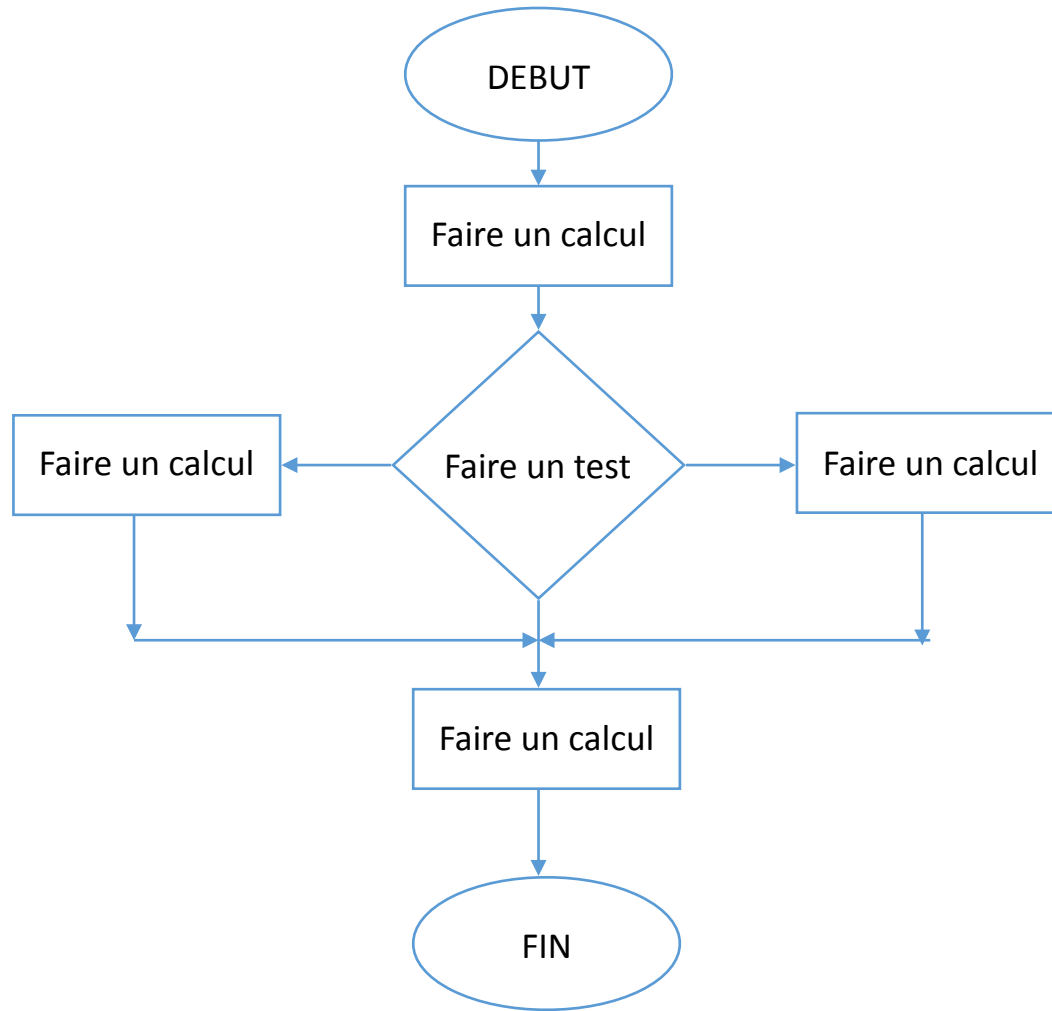
Kurt Gödel



Alonzo Church



Alan Turing



# Qu'est-ce que l'Intelligence Artificielle?

## La conférence de Dartmouth College 1956



John McCarthy

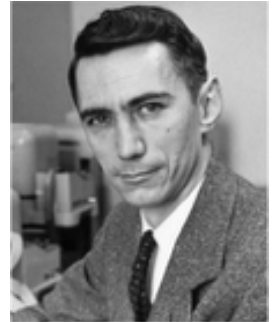


Marvin Minsky

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.



Nathaniel Rochester



Claude Shannon

# Processus Décisionnel Markovien

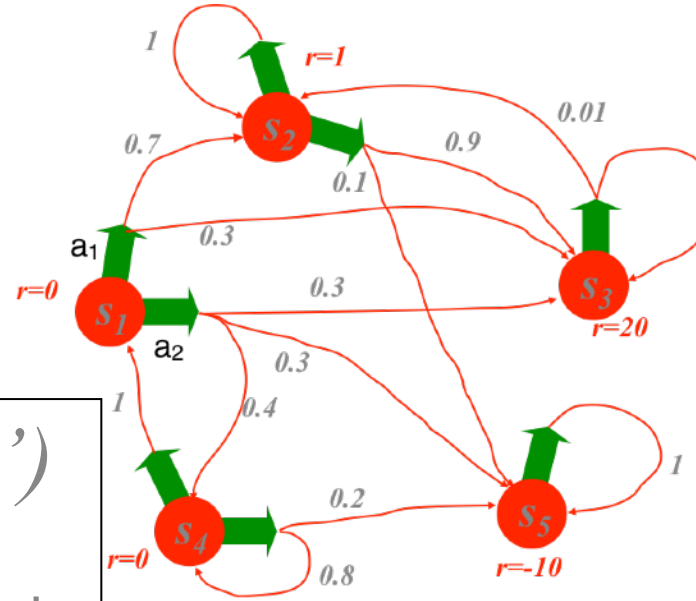
## Meilleure décision $a$ dans l'état $s$

$$U_{i+1}(s) = R(s) + \gamma \max_a \{ \sum_{s'} T(s, a, s') U_i(s') \}$$

$$\pi^*(s) = \text{Argmax}_a \sum_{s'} T(s, a, s') U(s')$$

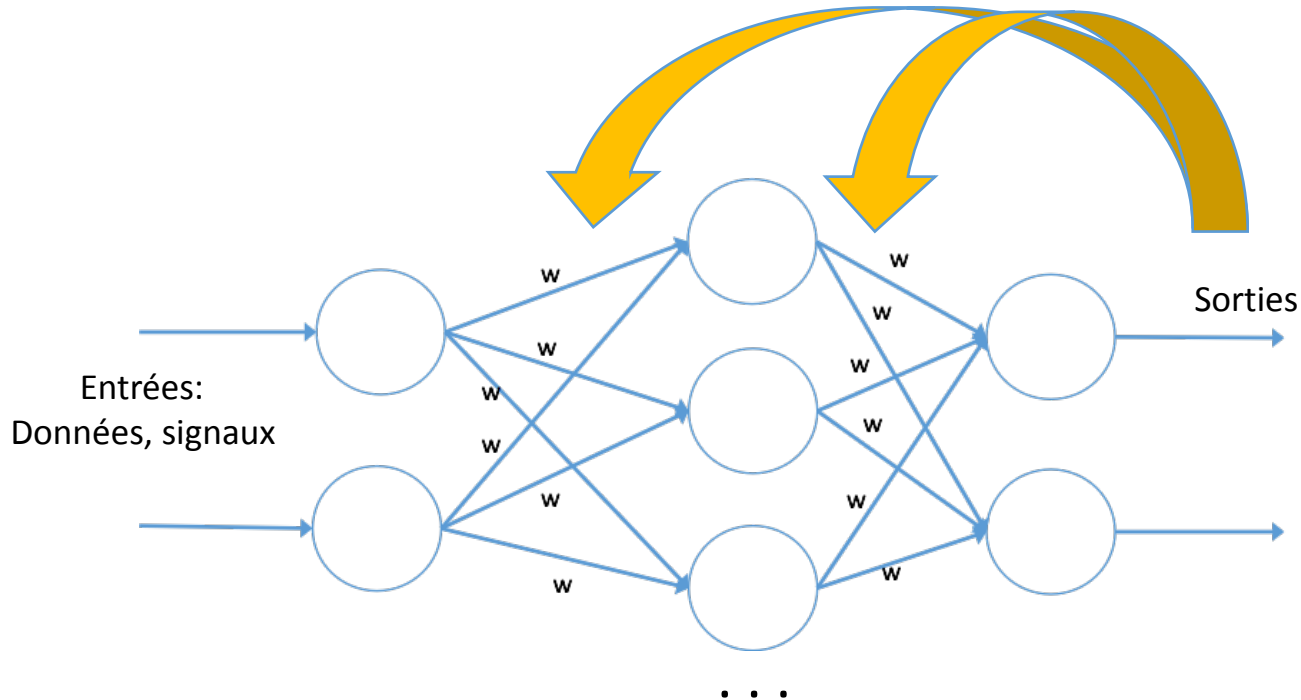
$T(s, a, s')$  = Probabilité d'atteindre un état  $s'$  à partir de l'état  $s$

$U(s')$  = Utilité de l'état  $s'$ .

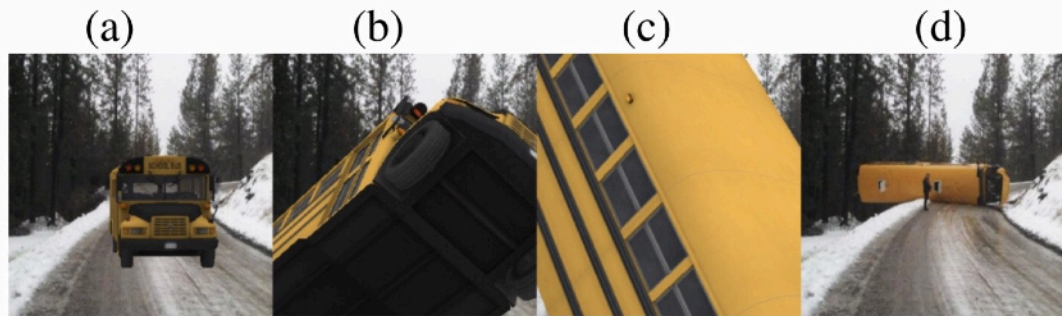


# Apprentissage (ou entraînement)

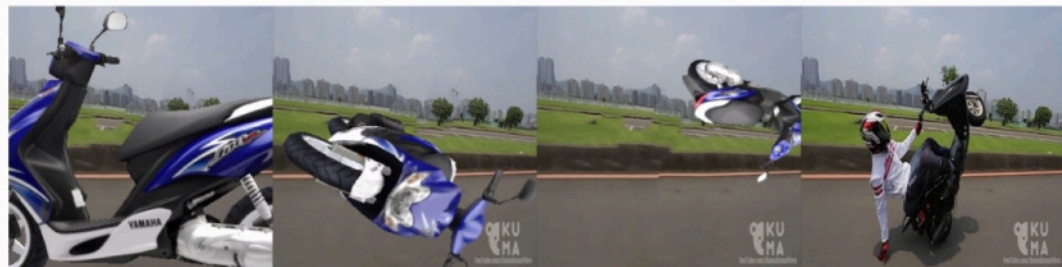
Optimisation: ajustement itératif des poids synaptiques pour obtenir les sorties désirées en réduisant les erreurs



# Exemple de limitations de l'apprentissage profond



**school bus** 1.0 **garbage truck** 0.99 **punching bag** 1.0 **snowplow** 0.92



**motor scooter** 0.99 **parachute** 1.0 **bobsled** 1.0 **parachute** 0.54



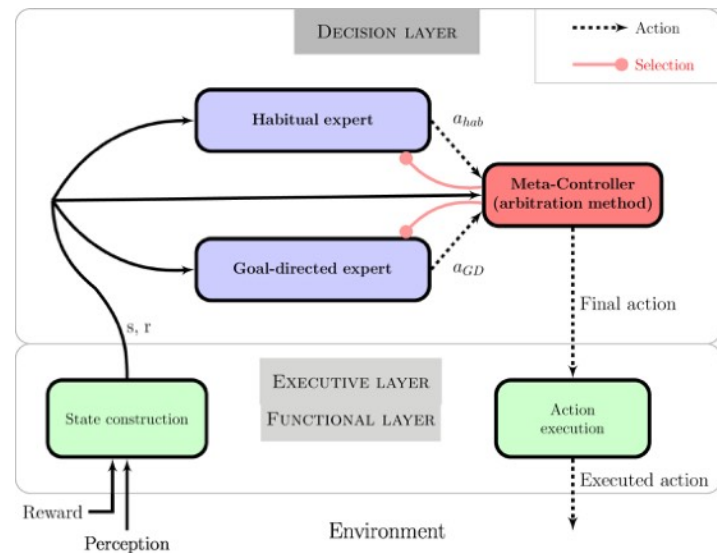
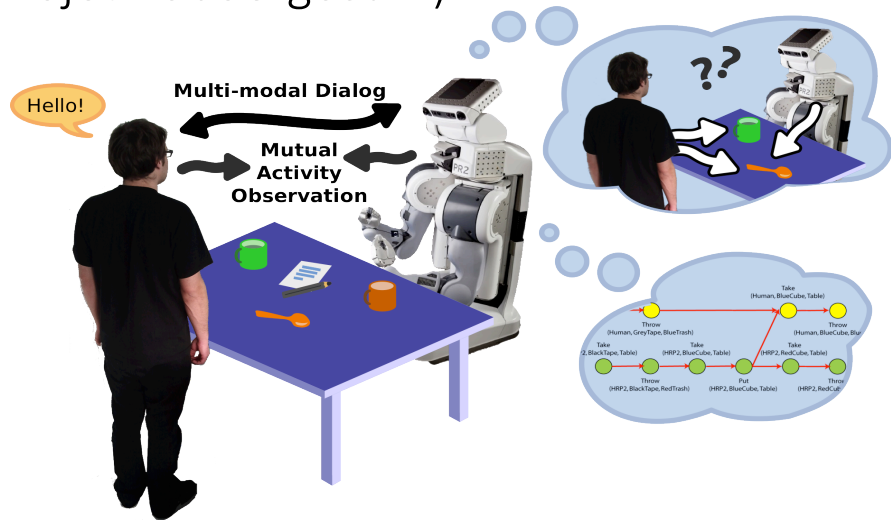
**fire truck** 0.99 **school bus** 0.98 **fireboat** 0.98 **bobsled** 0.79

Strike (with) a Pose: Neural Networks Are Easily Fooled by Strange Poses of Familiar Objects. Michael A. Alcorn et al., April 2019



Conjecture:

“La *conscience de soi* est fondée sur l’**interaction avec le monde** et sur une capacité d’**auto-évaluation** et de **méta-raisonnement**”  
(Projet Roboergosum)



# Autonomie, décision et conscience

- Les décisions prises par des systèmes informatiques se situent à un **niveau calculatoire** et sont inscrites dans les algorithmes conçus par les êtres humains.
- Toute “création” par un système informatique est le résultat d’un processus calculatoire.
- Les systèmes d’IA n’ont aucune autonomie de décision. Ils optimisent des fonctions.
- Les systèmes d’IA sont syntaxiques: ils n’ont aucune sémantique et ne comprennent pas ce qu’ils font.

- Utilisation métaphorique de termes relatifs aux humains par les experts pour designer le fonctionnement de programmes informatiques
- Exagération des résultats par les experts et par les médias
- Choix effectués pour résoudre des questions pratiques
- **La perception sur l'IA est faussée**
- **Responsabilité de clarification par la communauté.**

# A robot hand taught itself to solve a Rubik's Cube after creating its own training regime

Researchers at OpenAI have developed a new method for transferring complex manipulation skills from simulated to physical environments.

by Karen Hao

Oct 15, 2019