

---

---

# Multi-agent deep reinforcement learning in mobile robotics

**Maxime TOQUEBIAU**  
1ère année de thèse

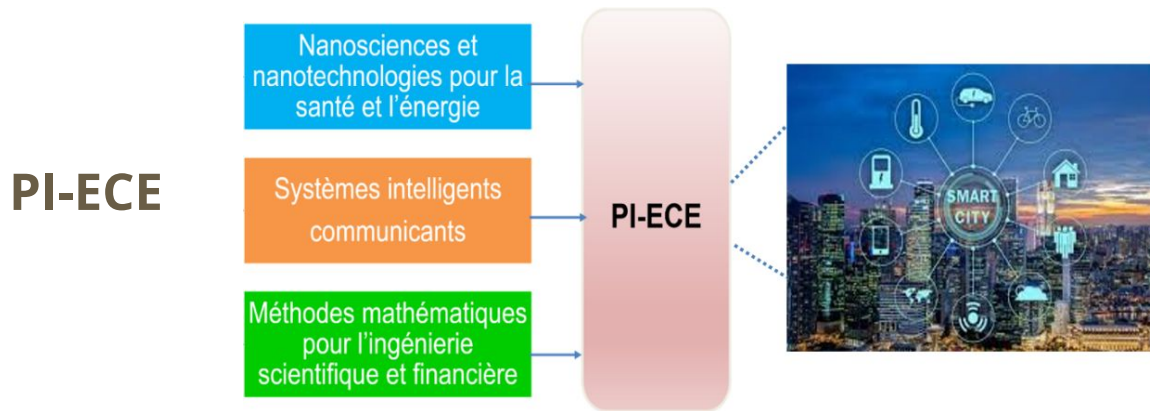
---

---

# Présentation

## Mon parcours

- ECE Paris - Diplôme d'ingénieur, Majeure Big Data & Analytics
- Stages:
  - Openvalue - **Reconnaissance faciale** (Computer Vision, Machine Learning)
  - Ministère des Armées - **Création d'un chatbot** (Génération automatique de texte)
- Thèse: janvier 2021 → décembre 2023



# Sujet de la thèse

Multi-agent deep reinforcement learning in mobile robotics

# Sujet de la thèse

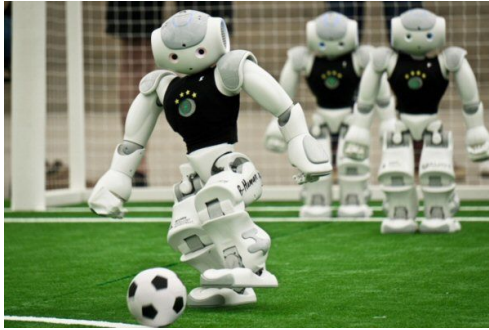
DEEP REINFORCEMENT LEARNING

Multi-agent deep reinforcement learning in mobile robotics

MULTI-AGENT SYSTEMS

MOBILE ROBOTICS

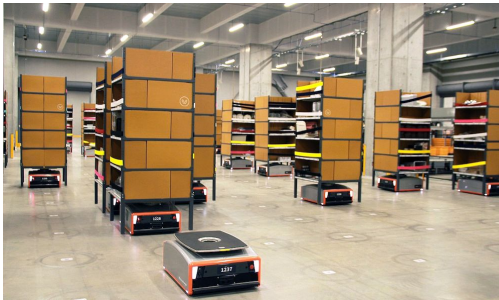
# Exemples d'applications



Soccer - RoboCup



Surveillance/Exploration -  
SMP Robotics



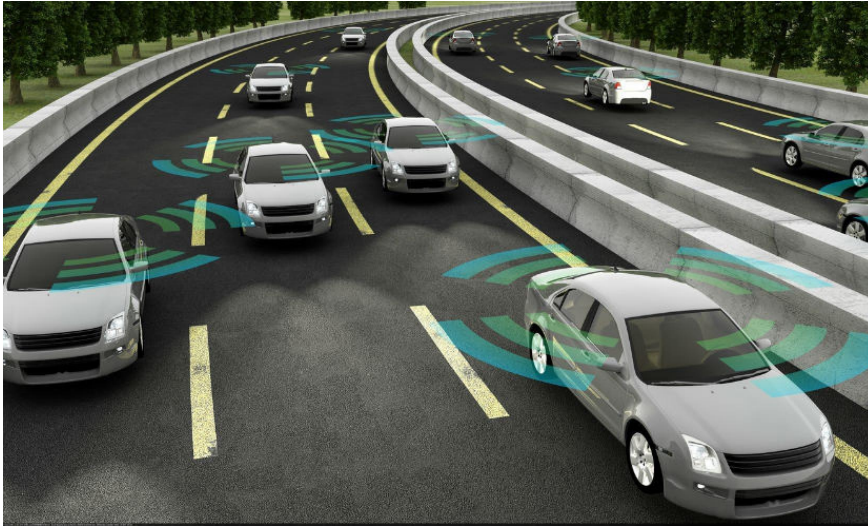
Logistique - GreyOrange



Sécurité/Militaire -  
Icarus Swarms

# Exemple d'application

Flotte de véhicules autonomes:



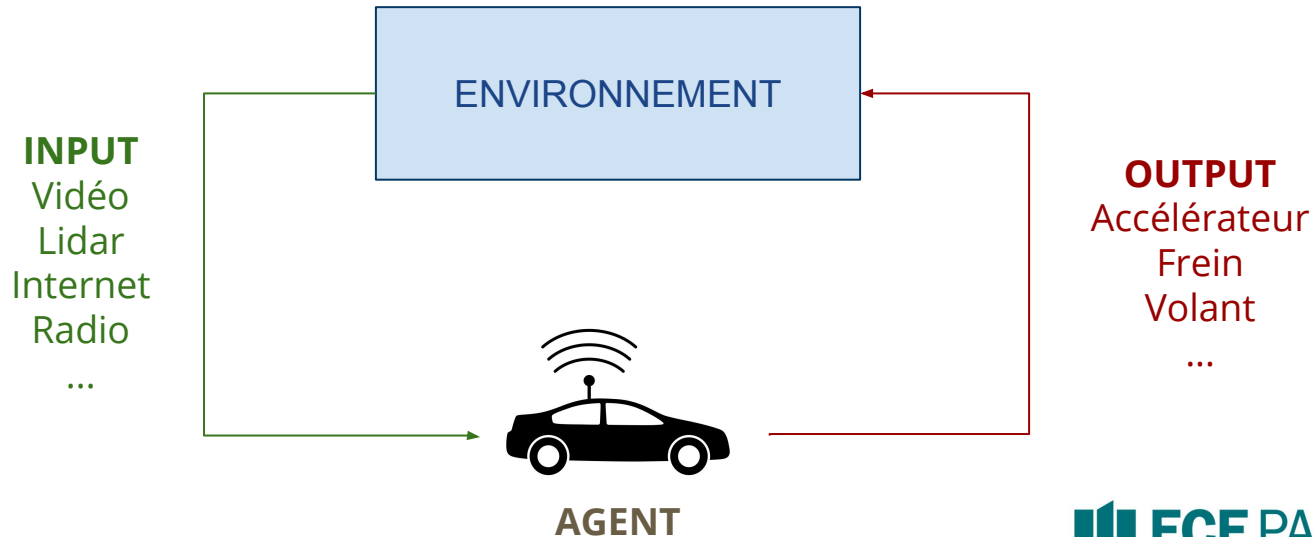
Compétences requises:

- Autonomie
- Planification
- Communication
- Négociation
- Adaptation

# Multi-agent Systems

## Définitions (1/2)

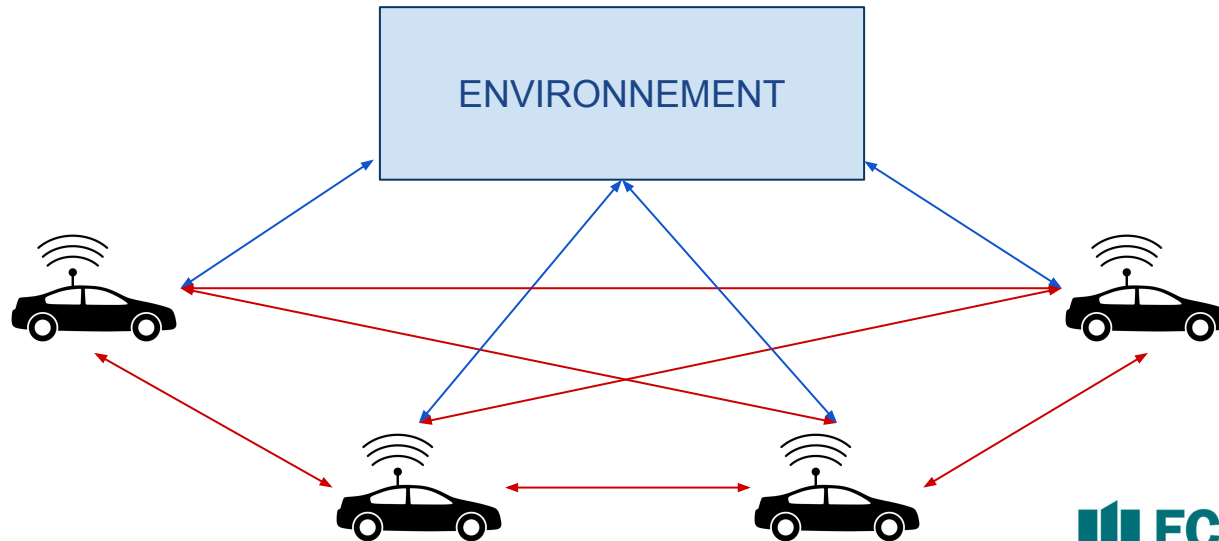
**Agent intelligent** = entité autonome capable de percevoir son environnement et d'agir sur celui-ci.



# Multi-agent Systems

## Définitions (2/2)

**Système multi-agent** = système composé de plusieurs agents intelligents, capables d'interagir entre eux et avec différentes parties de leur environnement

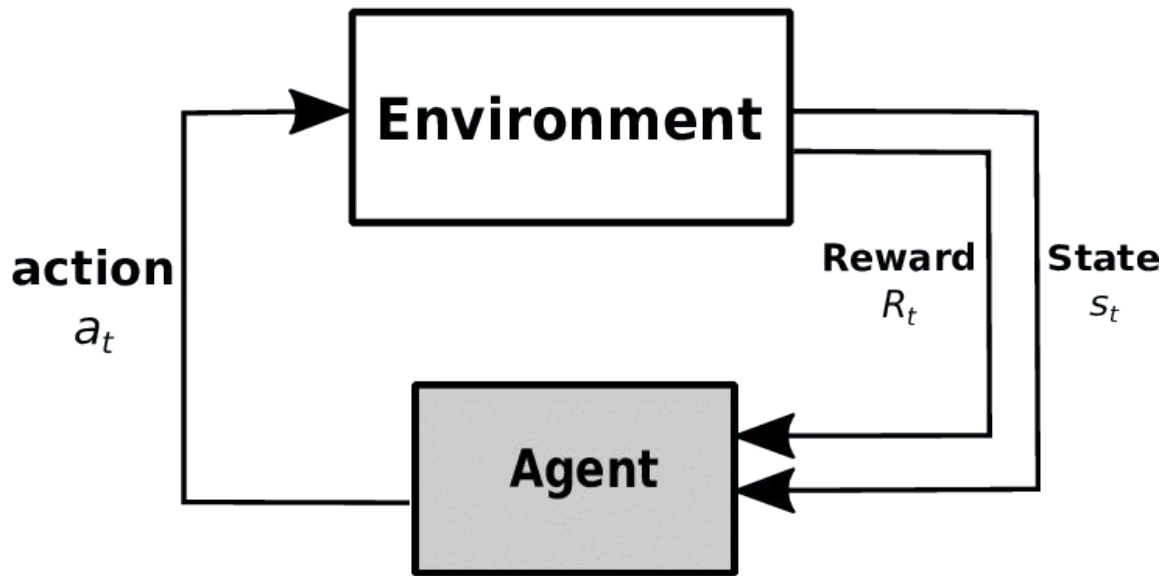




# Reinforcement learning

## Définitions (1/2)

**Apprentissage par renforcement:** Apprendre par l'interaction



# Reinforcement learning

## Définitions (2/2)

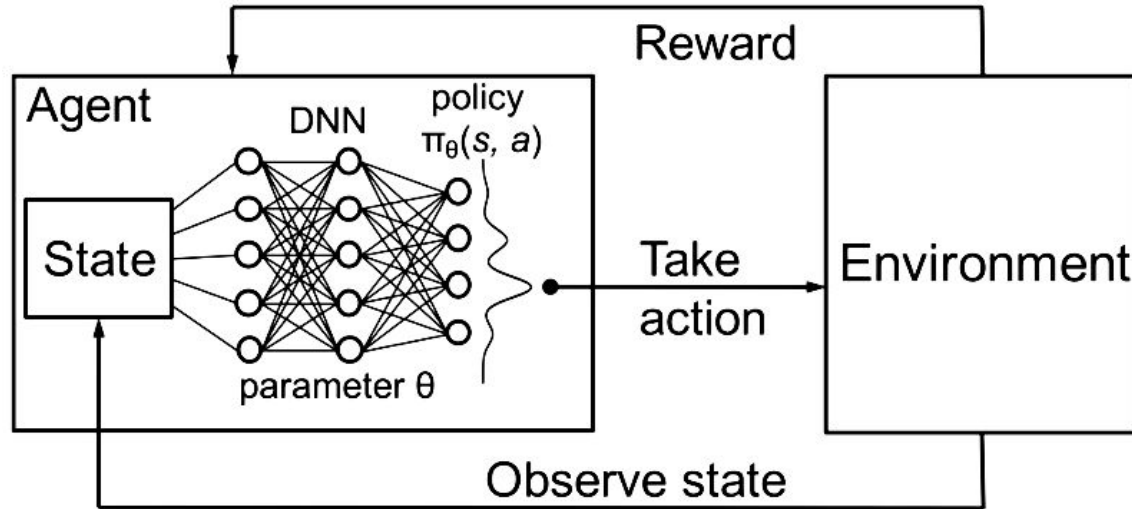
- Policy function = Stratégie d'un agent

$$\begin{array}{lcl} \rightarrow \text{Déterministe:} & \begin{array}{l} \pi : S \rightarrow \mathbb{R} \\ \pi(s) = a \end{array} & \left| \right. \rightarrow \text{Stochastique:} \begin{array}{l} \pi : S \times A \rightarrow [0, 1] \\ \pi(s, a) = \mathbb{P}[a|s] \end{array} \end{array}$$

- Value function = Récompense espérée en partant d'un état  $s$

$$V_{\pi}(s) = E_{\pi} \left[ \sum_{t=0}^{\infty} \gamma^t r_t | S_t = s \right], \gamma \in [0, 1]$$

# Deep reinforcement learning



Publications majeures:

- **Deep Q-Learning:** *Playing Atari with Deep Reinforcement Learning*, Mnih et al., 2013
- **Deep Deterministic Policy Gradient:** *Continuous control with deep reinforcement learning*, Lillicrap et al., 2015
- **Muzero:** *Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model*, Schrittwieser et al., 2019

# Deep reinforcement learning

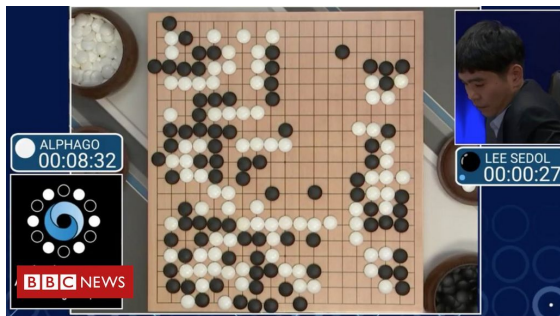
## Avantages du Deep learning:

- **Extraction automatique des informations utiles pour le contrôle d'un agent**
- Possibilité de traiter des espaces état-action de plus grandes dimensions
- Mieux généraliser

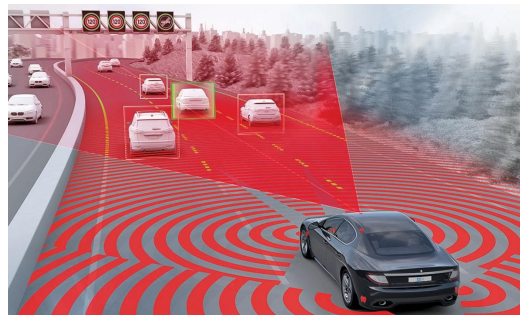
## Applications possibles:

Jeux à grandes dimensions:

**AlphaGo**  
(Silver et al., 2016)

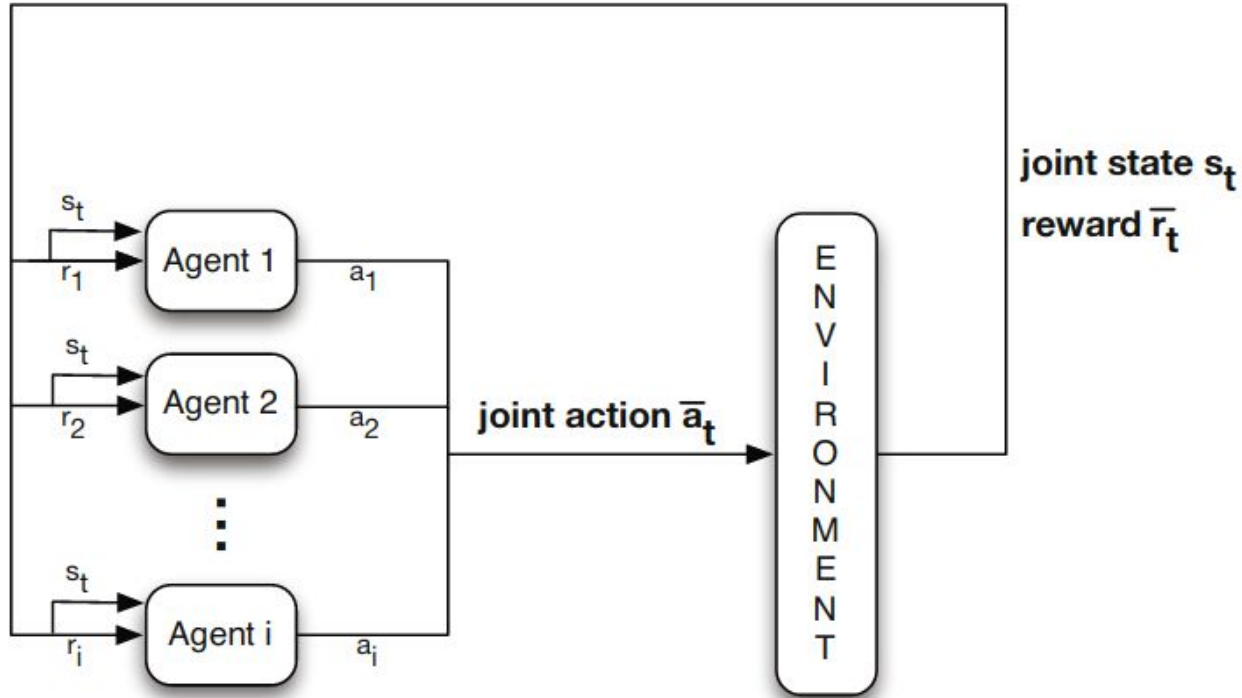


Monde réel:



**Conduite autonome**

# Multi-agent deep reinforcement learning



Nowé, Vrancx and De Hauwere, 2012

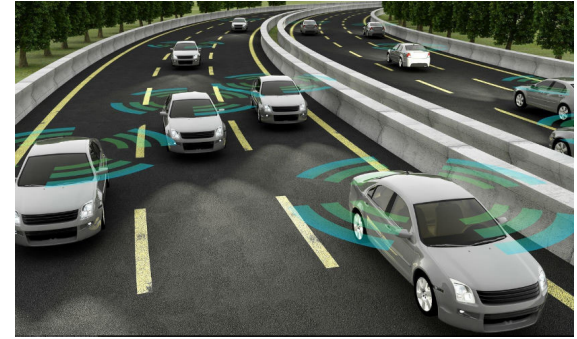
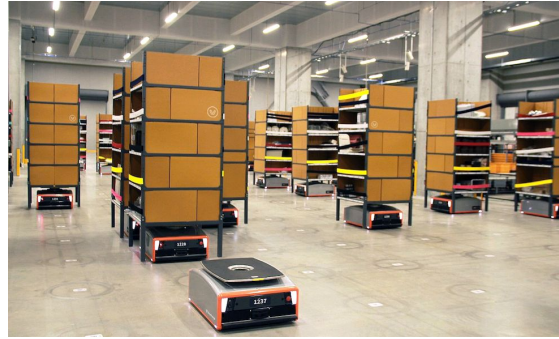
# Multi-agent deep reinforcement learning

Défis:

- Apprendre des interactions sociales
- Plusieurs agents => espace de recherche plus grand
- Plusieurs agents en apprentissage => non-stationnarité
- Communiquer quoi ? A qui ? Quand ?
  - ▲ Communiquer des informations non-pertinentes peut altérer la performance ou la rapidité d'apprentissage d'un agent (Tan, 1993)

# Mobile robotics

Def: Un robot mobile est un robot capable de se déplacer de manière autonome dans un environnement.



## Contraintes:

Complexité de  
l'environnement

Imperfection des  
capteurs/actionneurs

Contraintes de  
communication

Environnement  
dynamique

Sécurité

# Problématiques de recherche

- Partage des intentions d'action:

Y-a-t-il un lien entre le niveau de connaissance sur la stratégie des autres agents et la performance ?

- Robustesse par la coopération et la communication:

Comment mettre à profit la communication entre agents pour compenser aux contraintes du monde réel ?



# Encadrement

## Encadrant

Jae Yun JUN KIM

Poste: Enseignant chercheur

Affiliation: ECE Paris

Spécialité: Robotique,  
*Reinforcement Learning*

## Co-directeur

Faïz BEN-AMAR

Poste: Professeur

Affiliation: ISIR, Sorbonne  
Université

Spécialité: Robotique

## Co-directeur

Nicolas BREDECHE

Poste: Professeur

Affiliation: ISIR, Sorbonne  
Université

Spécialité: Systèmes  
multi-agent

**Merci pour votre attention!**  
Des questions ?