CNRS - INP - UT3 - UT1 - UT2J

Institut de Recherche en Informatique de Toulouse

L’intelligence artificielle : qu’est-ce que c’est ?

Dominique Longin
Chercheur CNRS,
Département d’intelligence artificielle, équipe Logique, Interaction, LAngage, et Calcul

Une définition générale de l'intelligence

- Intelligence :
  - « Ensemble des processus retrouvés dans des systèmes, plus ou moins complexes, vivants ou non, qui permettent de comprendre, d'apprendre ou de s'adapter à des situations nouvelles. » [wikipedia]

Une définition générale de l'intelligence

- Intelligence :
  - « Ensemble des processus retrouvés dans des systèmes, plus ou moins complexes, vivants ou non, qui permettent de comprendre, d'apprendre ou de s'adapter à des situations nouvelles. » [wikipedia]

Intelligence naturelle

Intelligence artificielle

Une définition générale de l'intelligence

- Intelligence :
  - « Ensemble des processus retrouvés dans des systèmes, plus ou moins complexes, vivants ou non, qui permettent de comprendre, d'apprendre ou de s'adapter à des situations nouvelles. » [wikipedia]

Intelligence naturelle

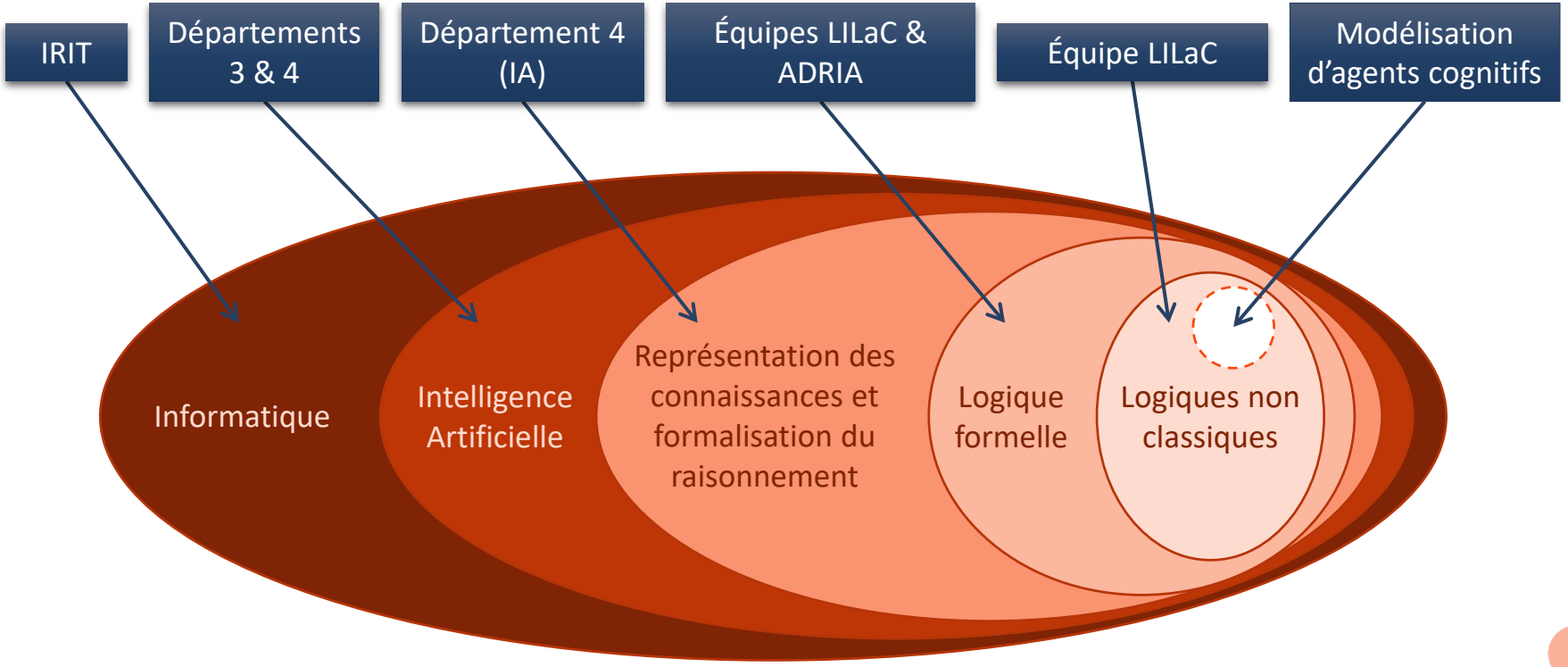
- psychologie,
- l'anthropologie (évolution),
- l'éthologie cognitive (intelligence animale),
- les neurosciences (biologie)
- ou encore la génétique …

Intelligence artificielle

???

L'intelligence artificielle (IA) : une partie de l'informatique…

- Exemple : cas de l'IRIT



L'intelligence artificielle (IA) : une partie de l'informatique…

… mais pas QUE !

- L'IA est abordée sous deux aspects :
  - Processus de raisonnement
  - Processus d'apprentissage

- Vision
  - Issue des Sciences Humaines (psycho, philo, socio…)
  - Et les Sciences du Vivant (neurosciences…)

⇨L'informatique s'inspire pleinement de ces domaines (théoriquement et expérimentalement)

L'intelligence artificielle (IA) : une partie de l'informatique…

… mais pas QUE !

- L'IA est abordée sous deux aspects :
  - Processus de raisonnement
  - Processus d'apprentissage

Cognition

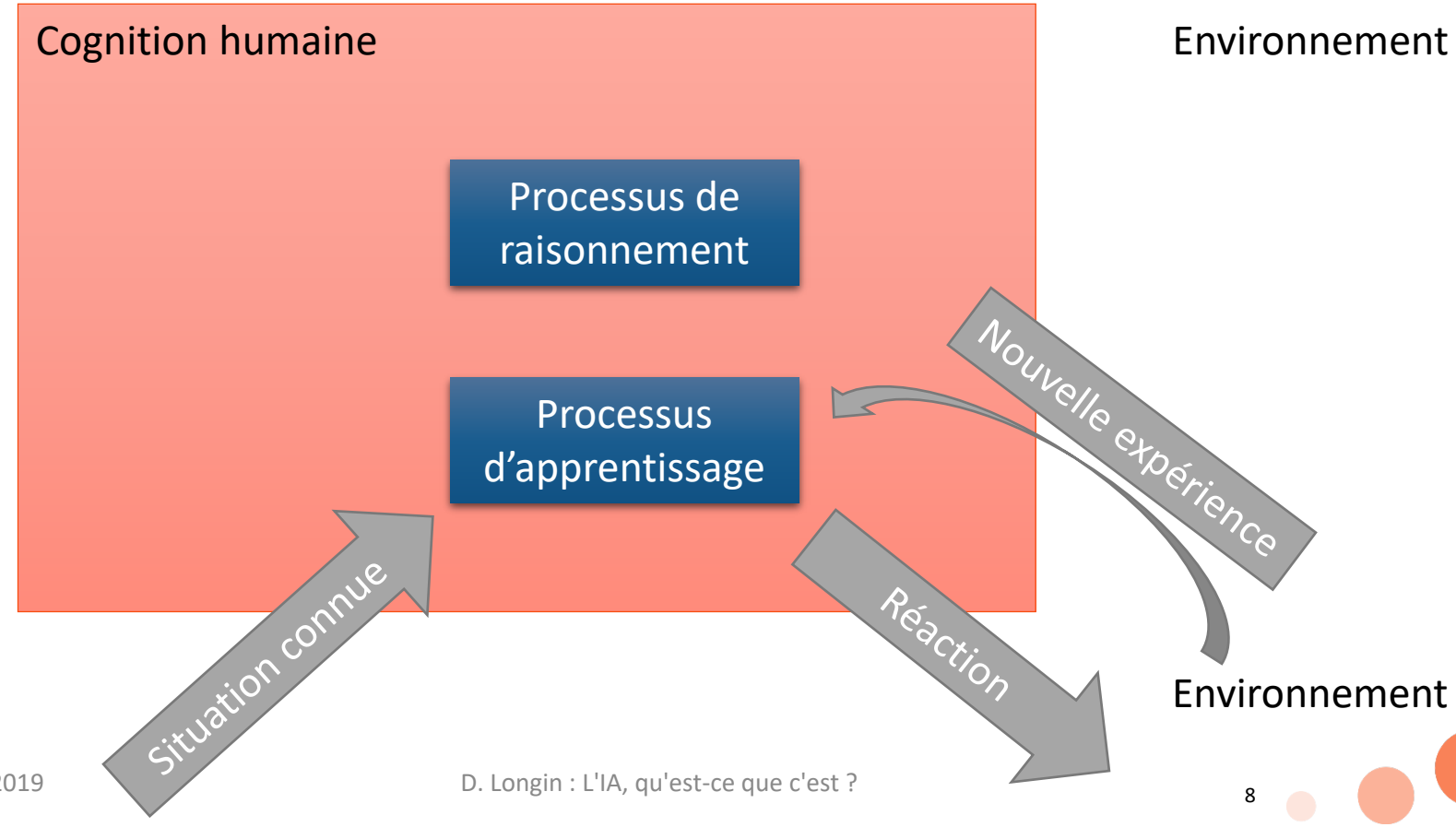
- Vision
  - Issue des Sciences Humaines (psycho, philo, socio…)
  - Et les Sciences du Vivant (neurosciences…)

⇨L'informatique s'inspire pleinement de ces domaines (théoriquement et expérimentalement)

L'intelligence artificielle (IA) : une partie de l'informatique…

… mais pas QUE !

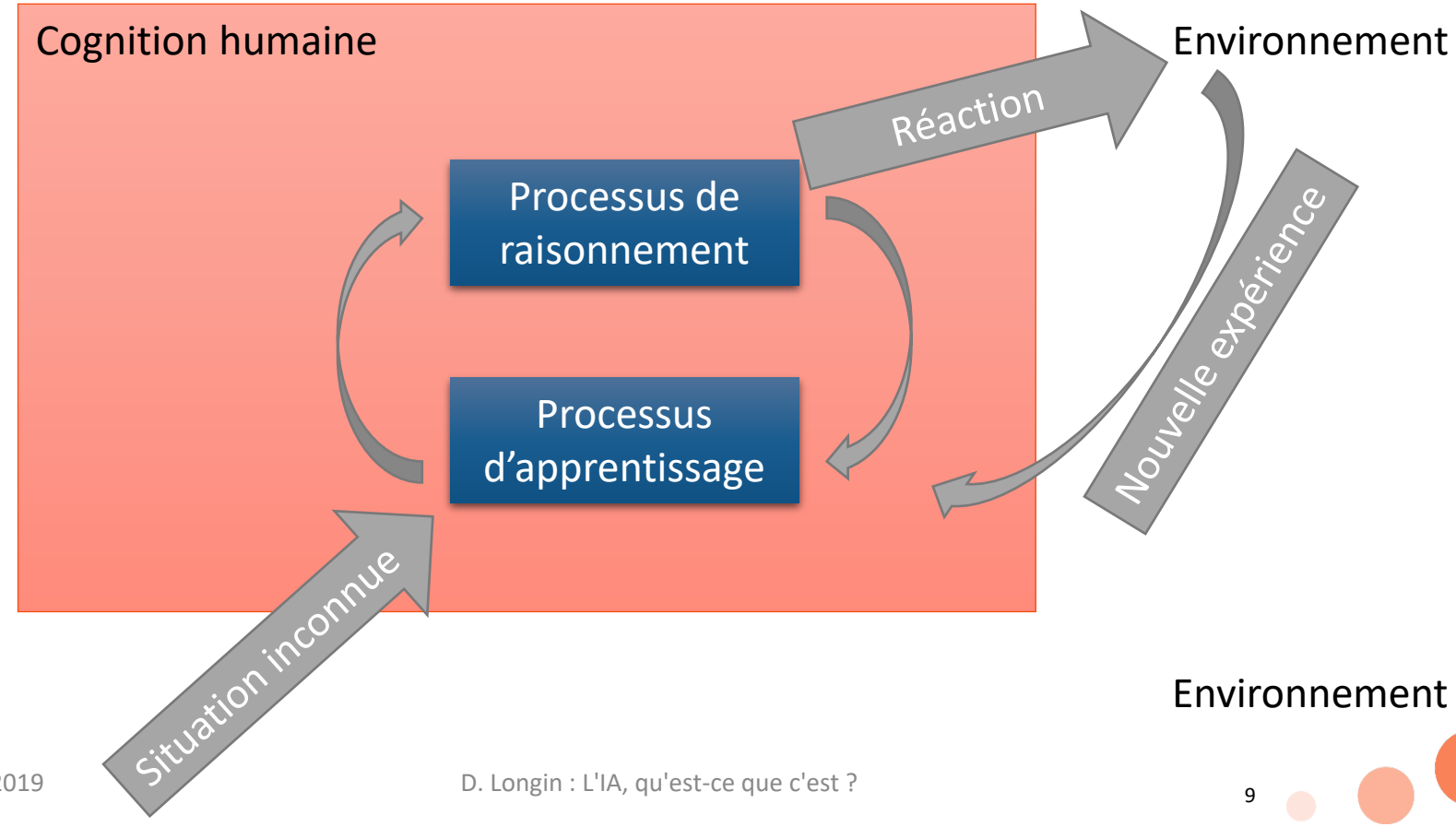
- Pour Sciences Humaines :



L'intelligence artificielle (IA) : une partie de l'informatique…

… mais pas QUE !

- Pour Sciences Humaines :



Le raisonnement

Représentation des connaissances
Méthodes d'inférences

Le principe

- **Analyser un domaine donné** à partir de réflexions ou d'expériences
  - statistiques connues (combien d'individu ont telle propriété…)
  - Résultats empiriques sur des êtres humains (émotions, décisions….)
  - Analyse philosophique …. (nature des concepts, propriétés, …)
- **Extraire les concepts** permettant de capturer ce domaine
- **Caractériser les propriétés** de ces concepts et leurs relations
- **Élaborer un processus de raisonnement** sur ces connaissances

Le principe

- **Analyser un domaine donné** à partir de réflexions ou d'expériences
  - statistiques connues (combien d'individu ont telle propriété…)
  - Résultats empiriques sur des êtres humains (émotions, décisions….)
  - Analyse philosophique …. (nature des concepts, propriétés, …)
- **Extraire les concepts** permettant de capturer ce domaine
- **Caractériser les propriétés** de ces concepts et leurs relations
- **Élaborer un processus de raisonnement** sur ces connaissances

Mathématiques

Linguistique

Philosophie

L'informatique a besoin des autres sciences !
Approche inter/pluridisciplinaire

Neurosciences

Psychologie

…

Économie

Exemple de la croyance

- Croyance *versus* connaissance
- Propriétés :
  - Si je crois p alors p est vrai dans le monde réel
  - Si je crois p alors je ne crois pas son contraire $\neg p$
  - Si je crois p alors je crois que je le crois
  - …
- Liens avec d'autre concepts ?
  - Désirs, buts, intention, obligations, permissions…
  - Ex. : Je crois qu'il veut manger
    (croyance, désir, action)

Exemple de la croyance

- Croyance *versus* connaissance
- Propriétés :
  - Si je crois p alors p est vrai dans le monde réel
  - Si je crois p alors je ne crois pas son contraire $\neg p$
  - Si je crois p alors je crois que je le crois
  - …
- Liens avec d'autre concepts ?
  - Désirs, buts, intention, obligations, permissions…
  - Ex. : Je crois qu'il veut manger (croyance, désir, action)

$$Bel_i p \rightarrow p$$

$$Bel_i p \rightarrow \neg Bel_i \neg p$$

$$Bel_i p \rightarrow Bel_i Bel_i p$$

$$Bel_i Intend_j \langle eat^{-1} \rangle$$

Les différents types de raisonnement

- On distingue deux types de raisonnement basés sur trois règles d'inférence différentes:
  - **Raisonnement déductif**
    - S'il n'utilise **QUE** la **règle de déduction** $\dfrac{\varphi \quad \varphi \rightarrow \psi}{\psi}$

DE φ *et* φ *implique* ψ
ON INFERE ψ

  - **Raisonnements hypothétiques**
    - S'il utilise au moins la **règle d'abduction** $\dfrac{\psi \quad \varphi \rightarrow \psi}{\varphi}$
    - Ou la **règle d'induction** $\dfrac{\varphi \quad \psi}{\varphi \rightarrow \psi}$

DE ψ *et* φ *implique* ψ
ON INFERE φ

DE φ et ψ
ON INFERE φ *implique* ψ

L’alcool et les mineurs

- Si le barman sert de l’alcool à l’individu X, alors X est majeur
- Que peut-on déduire de cette règle quand :
  - Le barman sert un Jack Daniel à Tom
  - Fattou est mineure
  - Le barman sert un coca à Mehdi
  - Louise est majeure

L'alcool et les mineurs

- Si le barman sert de l'alcool à l'individu X, alors X est majeur

- Que peut-on déduire de cette règle quand :

  - Le barman sert un Jack Daniel à Tom
    - ↳Tom est majeur
  - Fattou est mineure
    - ↳Le barman ne lui sert pas d'alcool
  - Le barman sert un coca à Mehdi
    - ↳ ??? (on ne peut rien dire sur son âge)
  - Louise est majeure
    - ↳ ??? (on ne peut rien dire sur ce qu'elle boit)

Règle du *modus ponens*

Règle du *modus tollens*

Exemple d'application

Évacuation des personnes en situation d'urgence

- **Modéliser l'incendie d'une boite de nuit (100 morts)** pour
  - Empêcher que ça se reproduise
  - Minimiser les pertes (humaines, matérielles)
  - Faciliter l'intervention des secours…
- **Modèles classiques de l'évacuation**
  - Basés sur la physique (fluides, gaz)
  - Comportements atypiques modélisés par de l'aléatoire
    ⇨ N'explique pas pourquoi certains s'arrêtent/retournent en arrière/attendent…
- **Modèles à base d'IA :**
  - Chaque individu est représenté (connaissances, actions possibles)
  - Son **comportement** découle de ses **actions**
  - Ses **actions** découlent de ses **connaissances** et de sa **personnalité**

Exemple d'application

Évacuation des personnes en situation d'urgence

- But des modèles à base d'IA :
  - **Redonner** à chaque personne son libre arbitre en fonction de certains **critères comportementaux** :
    - Des émotions
    - Des liens sociaux au seins de groupes d'individus
    - Des décisions autonome en fonction du contexte

⇨(RIEN n'est « programmé à l'avance », CHAQUE simulation est différente)

- Démonstration

L’apprentissage automatique

Les systèmes à base d'apprentissage

- 2 phases distinctes :
  - L'entrainement du modèle
    - Les règles ne sont plus des règles explicites de raisonnement
    - Mais le résultat de l'agrégation de l'analyse d'exemples rencontrés (des milliers, des millions, des milliards, des milliards de milliards)
    - Exemple : un série d'images de **cucurbitacées** et de **visages humains**
  - L'utilisation du modèle (phase d'inférence)
    - Un cas est soumis (par exemple une image)
    - Et le système doit identifier ce qu'il y a dessus (une courge, ou un visage ?)

Les systèmes à base d'apprentissage

- 2 phases distinctes :
  - L'entrainement du modèle
    - Les règles ne sont plus des règles explicites de raisonnement
    - Mais le résultat de l'agrégation de l'analyse d'exemples rencontrés (des milliers, des millions, des milliards, des milliards de milliards)
    - Exemple : un série d'images de **cucurbitacées** et de **visages humains**
  - L'utilisation du modèle (phase d'inférence)
    - Un cas est soumis (par exemple une image)
    - Et le système doit identifier ce qu'il y a dessus (une courge, ou un visage ?)
- 2 grandes familles :
  - Apprentissage statistique
  - Apprentissage profond (*deep learning*)

L’apprentissage statistique

- Apprentissage supervisé
  - À chaque image soumise au système durant la **phase d’apprentissage**
  - Le concepteur du système indique à la machine ce que c’est (courge ou visage humain par exemple)
- Apprentissage non supervisé
  - À chaque image soumise durant la phase d’apprentissage
  - Le système la classe dans une catégorie ou une autre selon la similarité des autres images déjà vues
    ⇨ tâche de classification (*clusterisation*)
- Apprentissage par renforcement
  - Ne fait **pas** appel à des exemples dans la phase d’apprentissage
  - Le système fait des actions sur son environnement et reçoit une récompense (ou non) en fonction du résultat
  - Exemple : AlphaGo
    (action = 1 coup, environnement = plateau de jeu, récompense = victoire ou défaite)

L’apprentissage profond

(*deep learning*)

- Systèmes construits à partir d’éléments de calcul de base : les « neurones »
  - Les neurones sont connectés entre eux
  - Ces liaisons peuvent avoir plus ou moins d’influence (ajustée durant la phase d’apprentissage)
- Neurone : fonction qui fusionne les informations reçues (des nombres flottants) en un résultat (un autre nombre) qu’il diffuse ensuite à tous les neurones auxquels il est connecté
- Avantage : particulièrement efficaces sur les signaux (vidéo, audio…)
- Inconvénient : la complexité du système empêche de comprendre **pourquoi** le système a donné une réponse plutôt qu’une autre

Conclusion

Pourquoi faire appel à l'intelligence artificielle ?

- Expérimentations sur l'utilisation de systèmes automatiques montrent toutes le même résultat :
  - Plus une machine donne l'illusion de se comporter comme un être humain
  - Plus l'usager se sent en confiance et moins l'utilisation de cette machine lui demande d'effort
- L'humain est doté de capacités exceptionnelles : c'est un modèle infini d'inspiration

Comment réconcilier raisonnement et apprentissage ?

- Actuellement deux types d'approche séparés
- L'approche raisonnement est
  - Très générique
  - Fort pouvoir explicatif sur ses résultats
  - Mais elle a conduit à beaucoup de travaux théoriques peu appliqués
- L'approche apprentissage
  - Fonctionne bien sur des problèmes bien identifiés
  - Mais uniquement si on dispose de (beaucoup de) données
  - Ne permet pas d'expliquer (justifier) les décisions du système (problématique sur les systèmes devant être certifiés, tels les avions par exemple)
- Enjeux de l'Institut Interdisciplinaire en IA (3IA) toulousain ANITI

Les craintes liées à l'utilisation de l'IA

- Essentiellement, de type éthique
  - Si (quand) les robots donnent « l'illusion de la vie », devons-nous les considérer socialement comme des être humains ?
    - Auront ils le droit de vote ?
    - Devons nous être polis avec eux ?
    - Devons nous tenir compte de leurs émotions ?
    - …
  - Comment rendre compte de comportements « typiquement humains » dans lequel le comportement n'est pas toujours « comptable/rationnel » (exemple du train, des véhicules autonomes…)
  - *etc.*