



Intelligence Artificielle Symbolique

Guillaume Piolle

guillaume.piolle@centralesupelec.fr

<http://guillaume.piolle.fr/>

CentraleSupélec, campus de Rennes
mineure « Techniques Inspirées du Vivant »

12 novembre 2015



- 1 Introduction
- 2 Fondements logiques
- 3 Agents et systèmes multi-agents
- 4 Exemples d'applications



Turing et l'intelligence artificielle

L'article fondateur

Alan M. Turing, **Computing Machinery and Intelligence**, *Mind* LIX(236), pp. 433-460, 1950.

Pose la question de la capacité des calculateurs à simuler l'intelligence (humaine).

Point de départ de la notion même d'**Intelligence Artificielle**.

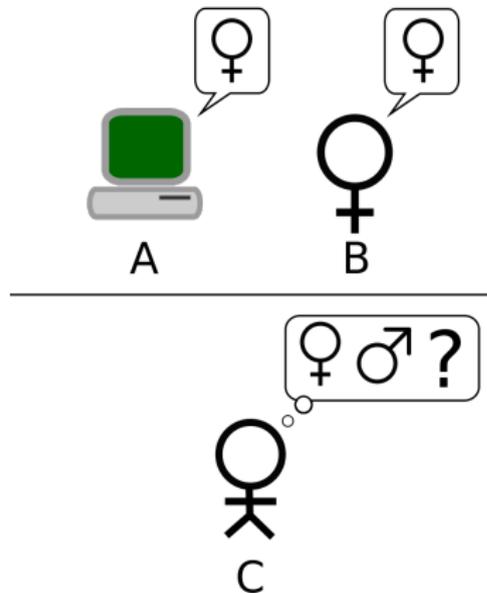




Turing et l'intelligence artificielle

Le Test de Turing

- Un opérateur humain pose des questions à l'aide d'un terminal caractère ;
- Un humain ou une machine répond à ses questions ;
- L'opérateur doit deviner si son interlocuteur est un humain ou une machine ;
- Si une machine arrive à tromper un humain dans plus de la moitié des cas, alors elle peut être qualifiée d'intelligente.





Turing et l'intelligence artificielle

Vidéo : Cleverbot



IA forte et IA faible

L'intelligence artificielle « forte »

- Quête de la conscience réelle et authentique ;
- Expression de sentiments authentiques ;
- Compréhension de son propre raisonnement ;
- Notion de *Ghost in the machine*.



IA forte et IA faible

L'intelligence artificielle « forte »

- Quête de la conscience réelle et authentique ;
- Expression de sentiments authentiques ;
- Compréhension de son propre raisonnement ;
- Notion de *Ghost in the machine*.

Critiques

- Critique de principe de penseurs religieux ou de philosophes pour qui la conscience demeure le propre des êtres vivants ;
- Critique au niveau de la notion de calculabilité ;
- Scénario de la chambre chinoise de Searle ;
- Comparaison avec le chaos déterministe.



IA forte et IA faible

L'intelligence artificielle « faible »

On ne recherche pas l'émergence d'une conscience, mais on s'applique à **simuler** (ou à modéliser, imiter) le raisonnement, l'introspection, l'émotion, etc.

Point de vue beaucoup plus pragmatique, moins « iconoclaste », mais sans doute moins poétique.

Perception de l'intelligence artificielle comme une « étiquette » accolée à certains outils.

The question of whether machines can think is about as relevant as the question of whether submarines can swim.

Edsger W. Dijkstra



IA forte et IA faible

Pertinence de la distinction

Peut-on vraiment distinguer une IA forte d'une IA faible (cf Turing) ?

Comment saurons-nous qu'une conscience réelle a émergé alors que nous ne sommes sans doute pas capable de caractériser notre propre conscience ni de la distinguer d'une conscience « simulée » ?



1 Introduction

2 Fondements logiques

- Formaliser le raisonnement humain
- Logique propositionnelle
- Logique du premier ordre
- Logiques d'ordre supérieur
- Logiques modales et temporelles

3 Agents et systèmes multi-agents

4 Exemples d'applications



Formaliser le raisonnement humain

Aristote (IV^{ème} siècle av. JC)

Cherche à extraire les règles formelles utilisées pour l'argumentation, afin (par exemple) de déduire mécaniquement si un raisonnement est fallacieux ou non.

- Notion de **syllogisme** ;
- Logique de la nécessité et de la possibilité.





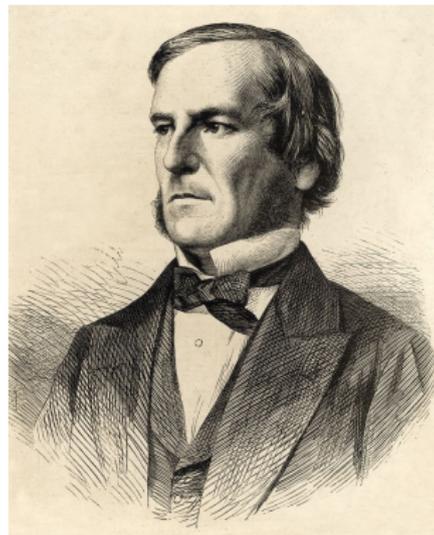
Formaliser le raisonnement humain

Georges Boole (XIX^{ème} siècle)

Tente d'**algébriser** la logique philosophique d'Aristote

- Utilisation d'opérateurs multiplicatif et additif ;
- Intégration du **calcul** de Leibniz dans le raisonnement logique.

Amélioré par William Stanley Jevons pour former l'**algèbre de Boole** : base de l'informatique, mais à l'origine une expérience de philosophie (analytique) !



Entre Aristote et Boole, pas de tentative sérieuse pour formaliser le raisonnement humain.



Logique propositionnelle

Logique propositionnelle = Calcul des propositions
Logique d'ordre zéro

Notion de proposition

Chrysippe de Soles (philosophe stoïcien, III^{ème} siècle av. JC) : une **proposition** est « ce qui est vrai ou faux, ou un état de choses complet qui, pour autant qu'il est lui-même concerné, peut être asserté ».

Notation typique des propositions : p, q



Logique propositionnelle

Construction des opérateurs usuels

- $\varphi \wedge \psi \stackrel{def}{=} \neg(\neg\varphi \vee \neg\psi)$
- $\varphi \rightarrow \psi \stackrel{def}{=} \neg\varphi \vee \psi$
- $\varphi \leftrightarrow \psi \stackrel{def}{=} (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$

Nous avons défini le **langage** (aspect syntaxique) de la logique propositionnelle.

Logique propositionnelle

Aspects sémantiques

- Une **interprétation** est un ensemble de valuations portant sur l'ensemble des variables propositionnelles du système ;

Logique propositionnelle

Aspects sémantiques

- Une **interprétation** est un ensemble de valuations portant sur l'ensemble des variables propositionnelles du système ;
- Si une formule est vraie pour une interprétation donnée, alors l'interprétation est un **modèle** de la formule ;
- Si une formule admet au moins un modèle, alors elle est dite **satisfaisable**. On dit que le modèle \mathcal{M} satisfait la formule φ ($\mathcal{M} \models \varphi$) ;
- Si une formule est vraie pour toute interprétation, alors elle est dite **valide**, c'est une **tautologie** ($\models \varphi$) ;
- Si une formule est fausse pour toute interprétation, alors elle est dite **invalidé**, c'est une **antilogie**.

Logique propositionnelle

Dérivation du théorème de l'identité : $p \rightarrow p$

- ① $(L_1)p \rightarrow ((q \rightarrow p) \rightarrow p)$
- ② $(L_2)(p \rightarrow ((q \rightarrow p) \rightarrow p)) \rightarrow ((p \rightarrow (q \rightarrow p)) \rightarrow (p \rightarrow p))$
- ③ $(1, 2, MP)(p \rightarrow (q \rightarrow p)) \rightarrow (p \rightarrow p)$
- ④ $(L_1, 3, MP)p \rightarrow p$

Logique propositionnelle

Dérivation du théorème de l'identité : $p \rightarrow p$

- ① $(L_1)p \rightarrow ((q \rightarrow p) \rightarrow p)$
- ② $(L_2)(p \rightarrow ((q \rightarrow p) \rightarrow p)) \rightarrow ((p \rightarrow (q \rightarrow p)) \rightarrow (p \rightarrow p))$
- ③ $(1, 2, MP)(p \rightarrow (q \rightarrow p)) \rightarrow (p \rightarrow p)$
- ④ $(L_1, 3, MP)p \rightarrow p$

Un peu laborieux, non ?



Logique propositionnelle

Bonnes propriétés d'un système logique

- **Cohérence** (non-contradiction) : Le système admet au moins un modèle (incohérence : *ex contradictione sequitur quod libet*);
- **Correction** : Tous les théorèmes du système sont valides (si $\vdash \varphi$ alors $\models \varphi$);
- **Complétude** : Tous les énoncés valides sont des théorèmes (si $\models \varphi$ alors $\vdash \varphi$);
- **Décidabilité** ou **calculabilité** : De n'importe quelle formule on peut décider si c'est un théorème ou non : il existe une **procédure de décision**, ou **algorithme**.

Chance !

La logique propositionnelle est cohérente, correcte, (fortement) complète, décidable.



Logique du premier ordre

Beaucoup plus expressive que la logique propositionnelle ! Suffit pour traiter de nombreux concepts.

Propriétés formelles

- Les systèmes de calcul courants sont **corrects** ;
- La logique est **complète** (Gödel 1929).



Logique du premier ordre

Un cas particulier : les clauses de Horn

Cluses (disjonctions) contenant au maximum un littéral (élément atomique) positif :

$$g \vee \neg p_1 \vee \neg p_2 \vee \neg p_3 \vee \neg p_4$$

Cas particulier relativement courant et très pratique en logique propositionnelle ou du premier ordre, notamment grâce à l'algorithme de SLD-résolution (polynômial pour les clauses propositionnelles).



Logique du premier ordre

Exemple de programme Prolog

```

pere(marc, helena).
mere(helena, david).
pere(benjamin, herve).
mere(helena, herve).

parent(A, B) :- pere(A, B) ; mere(A, B).

grandpere(A, B) :- parent(C, B),
                  pere(A, C).

sibling (A, B) :- parent(C, A),
                 parent(C, B).
  
```

Exemple de requêtes

```

?- grandpere(X, herve).
X = marc.
?- sibling (david, herve).
true.
  
```



Logiques d'ordre supérieur

Logique du second ordre

Généralement la seule logique d'ordre supérieur considéré.

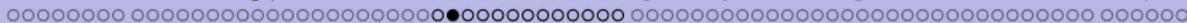
Principe : introduction de **variables relationnelles**, pouvant représenter des symboles de fonctions ou de prédicats, et sur lesquels il est possible de faire de la quantification.

Exemple : $\forall P \exists x \exists y P(x, y)$

Logique *monadique* du second ordre

On se limite à des variables relationnelles d'arité 1 (ensembles).

Beaucoup plus expressive que la logique du premier ordre, mais quasiment impossible à automatiser (pas de système de calcul correct et complet).



Notion de modalité

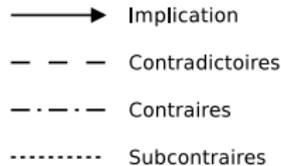
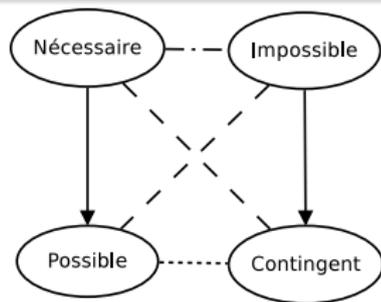
On utilise \Box (*box*) et \Diamond (*diamond*) comme symboles génériques pour les modalités **universelle** et **existentielle**, respectivement.

Dualité entre universel et existentiel : $\Diamond\varphi \stackrel{def}{=} \neg\Box\neg\varphi$

(où φ est une formule bien formée de la logique modale).

Exemple : logique aristotélienne (logique aléthique)

- $\Box\varphi$: φ est **nécessaire** ;
- $\Diamond\varphi$: φ est **possible** ;
- $\neg\Box\varphi$: φ est **contingent** ;
- $\neg\Diamond\varphi$: φ est **impossible**.



Sémantique des systèmes de logique modale

Sémantique de Kripke

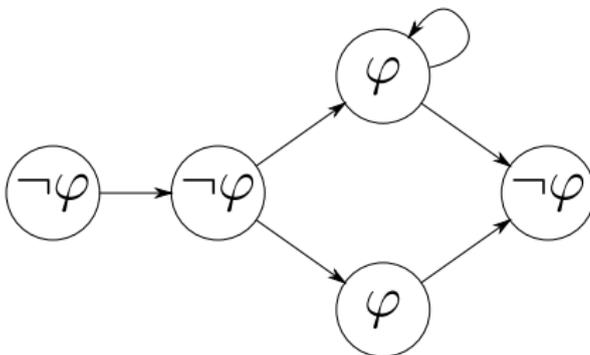
Un exemple de sémantique des « mondes possibles ». Correspond aux systèmes dits normaux. Simple à exprimer, élégante, des théorèmes intéressants, mais pas toujours adaptée.

Éléments d'un modèle \mathcal{M}

- Un ensemble de **mondes** $\mathcal{W} = \{w_i\}$, qui peuvent représenter des états du systèmes ou des exécutions possibles ;
- Une **relation binaire d'accessibilité** \mathcal{R} (arcs d'un graphe orienté), dit quels mondes sont « accessibles » depuis quels autres ;
- Une **fonction de valuation** h , $h(p)$ étant l'ensemble des mondes dans lesquels la proposition atomique p est vraie.

Sémantique des systèmes de logique modale

$\mathcal{M}, w \models \varphi$ se lit « φ est vraie au monde w du modèle \mathcal{M} ». $\Box\varphi$ est vraie ssi φ est vraie dans tous les mondes accessibles, $\Diamond\varphi$ est vraie ssi φ est vraie dans au moins un monde accessible.



$$\mathcal{M}, w_1 \models \Box\Box\varphi$$

$$\mathcal{M}, w_2 \models \Box\varphi$$

$$\mathcal{M}, w_3 \models \Diamond\varphi$$

$$\mathcal{M}, w_4 \models \Box\neg\varphi$$

$$\mathcal{M}, w_5 \models \Box\varphi$$

$$\mathcal{M}, w_5 \models \Box\neg\varphi$$

Sémantique des systèmes de logique modale

Validité, satisfaisabilité, équivalence

- Une formule est **Kripke-valide** (resp. valide dans un modèle) ssi elle est vraie en tout monde de tout modèle de Kripke (resp. en tout monde du modèle) ;
- Une formule est **Kripke-satisfaisable** (resp. satisfaisable dans un modèle) ssi il existe un monde de tout modèle de Kripke (resp. du modèle considéré) où elle est vraie ;
- Deux formules sont **Kripke-équivalentes** (resp. équivalentes dans un modèle) ssi en tout monde de tout modèle de Kripke (resp. en tout monde du modèle considéré), l'une est vraie ssi l'autre est vraie.

Exemple : L'axiome (K) est Kripke-valide.

La notion de validité dans un modèle permet d'exprimer une vérité « contingente », par opposition à la vérité « absolue » de la Kripke-validité.

Sémantique des systèmes de logique modale

Théorème de Sahlqvist (simplifié)

Tout système de logique modale normale fondé sur l'axiome K + des axiomes de Sahlqvist est **correct et fortement complet** pour la classe de modèles de Kripke dont les relations d'accessibilité ont les propriétés correspondant aux formules de Sahlqvist utilisées.

C'est une bonne nouvelle. On peut donc construire une sémantique de Kripke à partir d'une axiomatique donnée (on dispose d'un algorithme pour trouver la propriété sur la relation à partir de la formule de Sahlqvist).

Complexité des systèmes de logique modale

La plupart des logiques modales normales utilisant les axiomes classiques sont décidables, mais la complexité des problèmes SAT associés varie beaucoup.

Complexité de SAT dans quelques systèmes

- K : PSPACE (PSPACE-complet ?)
- Systèmes Λ tels que $K \subseteq \Lambda \subseteq S4$: PSPACE-difficile ;
- PDL (Propositional Dynamic Logic) : EXPTIME-complet ;
- $S4.3$ et toutes ses extensions : NP-complet ;
- Et certains systèmes ont un problème SAT indécidable. . .

L'ajout d'axiomes dans un système peut augmenter *ou* diminuer la complexité.

Note : $S4 = KT4$, $S4.3 = KT4.3$ (axiomes K , T , 4 et $.3$, non précisé ici)

Logiques temporelles linéaires (LTL)

Les mondes représentent des instants du temps, ils sont organisés en une chaîne linéaire, du passé vers le futur. Représentation d'événements inscrits dans le temps, de notions de succession, de précédenance, de passé, de futur.

Modalité *neXt*

$X\varphi$ est vraie ssi φ est vraie à l'instant suivant. Modalité universelle (presque) équivalente à sa modalité existentielle duale. La relation d'accessibilité associée \mathcal{R}_X relie chaque instant au suivant (non réflexive, non transitive).

Modalité *G* (*Going to be true*)

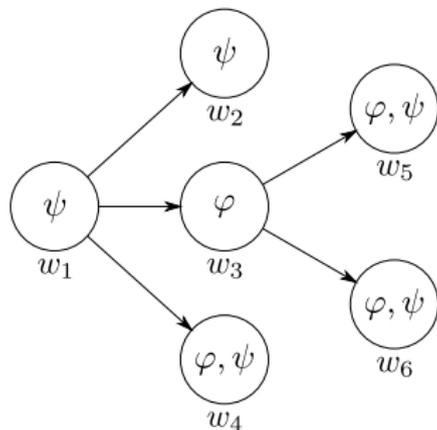
$G\varphi$ est vraie ssi φ sera vraie à tout instant du futur. Modalité universelle dont le dual existentiel est $F\varphi$ (φ sera vraie à un instant du futur).
Relation d'accessibilité \mathcal{R}_G associée : fermeture transitive de \mathcal{R}_X .

Computational Tree Logic (CTL)

Exemple classique de logique temporelle arborescente (*branching time logic*). CTL* est une variante combinant LTL et CTL.

Mêmes modalités qu'en LTL, plus un préfixe modal A (universel, pour tout chemin futur) ou E (existential, il existe un chemin futur).

- $AX\varphi$: pour tous les instants successeurs possibles, φ est vraie ;
- $EX\varphi$: il existe un instant successeur où φ est vraie ;
- $AG\varphi$: φ est vraie à tout instant de tout chemin futur possible ;
- $EG\varphi$: il existe un chemin futur où φ est vraie à tout instant ;
- ...



$$\mathcal{M}, w_1 \models EX\varphi$$

$$\mathcal{M}, w_1 \models EG\psi$$

$$\mathcal{M}, w_1 \models AF\psi$$



- 1 Introduction
- 2 Fondements logiques
- 3 Agents et systèmes multi-agents**
 - Agents réactifs et agents cognitifs
 - Modélisation des croyances et des connaissances
 - Le modèle BDI
 - Systèmes multi-agents
 - Modélisation des normes
 - Modélisation des émotions
 - Mise en œuvre
- 4 Exemples d'applications

Agents réactifs et agents cognitifs

On dispose d'outils logiques et informatiques qui pourraient nous permettre de **modéliser** et de **reproduire** certaines formes de **raisonnement**.

Dans le cadre de la démarche de l'intelligence artificielle, il est courant de vouloir attacher ces facultés de raisonnement à une **entité** bien définie, candidat pour être la machine « intelligente » de Turing.

Agents réactifs et agents cognitifs

Notion d'agent

On appelle **agent** une entité logicielle ou matérielle :

- Capable de **percevoir** son environnement ;
- Capable d'**agir** sur son environnement ;
- Capable de se comporter de manière **autonome**.

Agents réactifs et agents cognitifs

Notion d'agent

On appelle **agent** une entité logicielle ou matérielle :

- Capable de **percevoir** son environnement ;
- Capable d'**agir** sur son environnement ;
- Capable de se comporter de manière **autonome**.

Caractéristiques complémentaires courantes

- Capable d'**interagir** avec d'autres agents ;
- Capable de poursuivre des **objectifs** ;
- Capable d'agir de manière **rationnelle** ;
- Disposant de ressources propres ;
- ...

Agents réactifs et agents cognitifs

Agent réactif

Un **agent réactif** fonctionne en répondant à des stimuli externes.

Un agent réactif n'a pas de but ni de vision à « long terme », il n'a que des règles de fonctionnement qui lui permettent de réagir à des situations données (éventuellement en mettant en œuvre des facultés de raisonnement logique).

Cas typique : fourmis artificielles.



Modélisation des croyances et des connaissances

Premier besoin : représenter la base de connaissances de l'agent, de manière qu'il puisse raisonner dessus avec des outils logiques.

→ base logique (vs. relationnelle, objet, etc.)

Logique épistémique et logique doxastique

La logique **épistémique** est la logique modale de la connaissance.

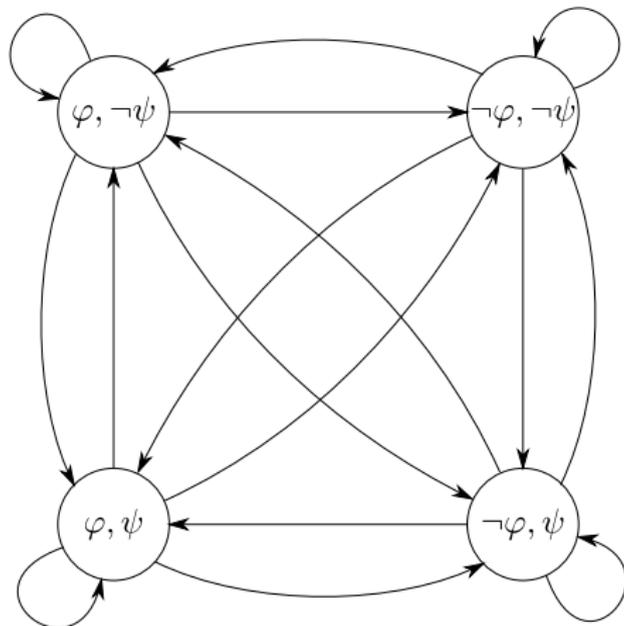
$K_a\varphi$: l'agent *a* connaît l'information φ (il sait que φ est vraie).

La logique **doxastique** est la logique modale de la croyance.

$Bel_a\varphi$: l'agent *a* croit que l'information φ est vraie.

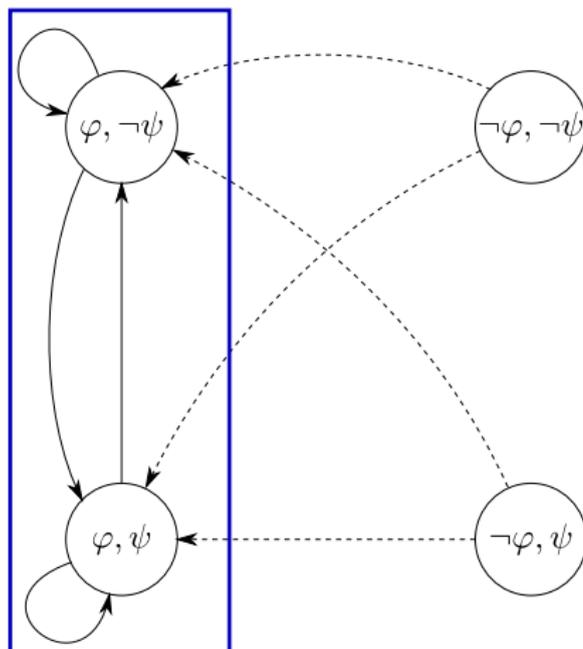


Modélisation des croyances et des connaissances



Modélisation des croyances et des connaissances

$\text{Bel } \varphi$

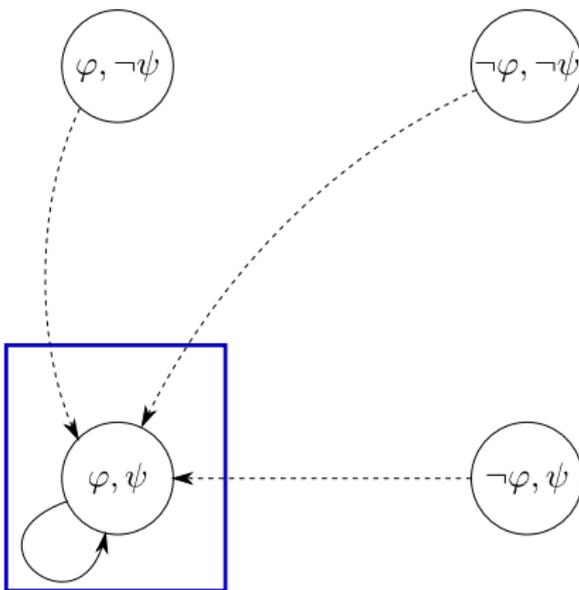




Modélisation des croyances et des connaissances

Bel φ

Bel ψ



Le modèle BDI

Belief - Desire - Intention

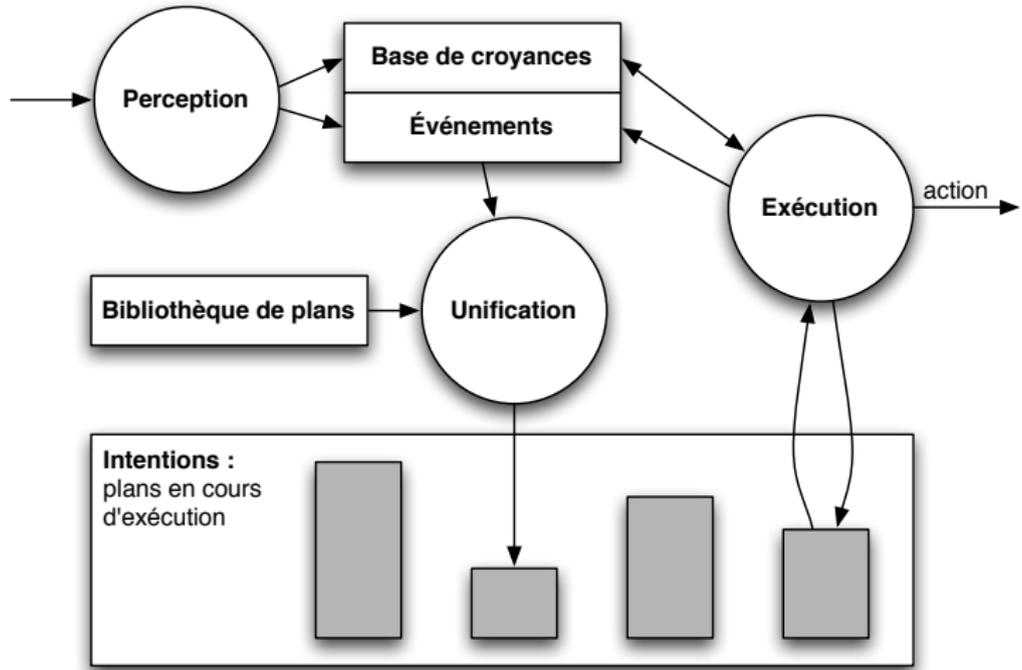
Modèle cognitif le plus répandu pour les agents cognitifs. Il peut être mis en œuvre sous forme procédurale (Bratman 87) ou sous forme logique (Cohen & Levesque 1990).

Éléments du modèle :

- **Croyances** : la représentation du monde par l'agent ;
- **Désirs** ou buts : objectifs désirables pour l'agent ;
- **Intentions** : parmi les buts, ceux que l'agent considère comme possibles à atteindre et sur lesquels il s'est engagé.

Le modèle BDI

Le modèle AgentSpeak





Le modèle BDI

Modèles logique pour agents BDI

Représentation, à l'aide de modalités, des intentions, des croyances, des buts de l'agent, ainsi que de toutes les notions intermédiaires ou nécessaires, sous la forme d'une théorie logique.

On part d'un ensemble minimal de modalités (notamment temporelles) reliées par leurs axiomes, puis construction de modalités « abrégées » pour décrire des concepts plus complexes en fonction des modalités fondamentales du modèle.

Beaucoup de travaux sur la formalisation de la notion d'intention (complexe à cerner).

Le modèle BDI

Exemple de modélisation simple des buts et intentions adapté d'un cours d'Andreas Herzig

Modalités fondamentales

- *Choice*_{*i*} φ : *i* préfère que φ soit vraie ;
- Modalités *F*, *G*, *U*, *S* de la logique temporelle linéaire.

Modalités abrégées

- But : $Goal_i \varphi \stackrel{def}{=} Choice_i F \varphi$;
- But à réaliser : $AGoal_i \varphi \stackrel{def}{=} Choice_i \varphi \wedge Bel_i \neg \varphi$;
- But persistant : $PGoal_i \varphi \stackrel{def}{=} (AGoal_i \varphi) U (Bel_i \varphi)$
- Intention : $Int_i \varphi \stackrel{def}{=} PGoal_i \varphi \wedge Choice_i F \exists \alpha (acteur(\alpha, i) \wedge [\alpha] \varphi)$.

- Nature essentiellement dynamique de l'intention ;
 → Intention d'un **état** : causalité entre α et φ ?



Le modèle BDI : Cohen & Levesque (version simplifiée)

Cohen & Levesque, **Intention is Choice with Commitment**, *Artificial Intelligence*, 42, pp. 231–261, 1990.

Modalités de base

- $Bel_i \varphi, G \varphi, H \varphi$;
- $Goal_i \varphi$: l'agent i a pour but que φ soit vraie ;
- $Happens_i \alpha$: l'agent i va réaliser l'action α à l'instant suivant ;
- $Done_i \alpha$: l'action α vient d'être réalisée par l'agent i ;
- $\varphi? \alpha$: si φ est vraie, alors l'action α est exécutée, sinon échec ;
- $\alpha; \alpha'$: succession des actions α et α' .

On définit également les modalités temporelles $F, P, Later$

($Later \varphi \stackrel{def}{=} \neg \varphi \wedge F \varphi$) et *Before*

($Before(\varphi, \psi) \stackrel{def}{=} G(Happens \psi \rightarrow P Happens \varphi)$).



Systèmes multi-agents

Un **système multi-agents**, ben... C'est quand y'a plusieurs agents.

Tout de même, il faut que les agents interagissent un minimum : partage d'un *Agent Communication Language*.

- Les agents peuvent poursuivre des buts communs, convergents, compatibles, incompatibles, complètement opposés...
- Les agents peuvent partager le même modèle ou former un groupe hétérogène ;
- Les agents peuvent être bienveillants (*benevolent*) ou malveillants (*malevolent, malicious*) ;
- Les agents peuvent avoir des visions différentes de leur environnement et de leurs voisins.

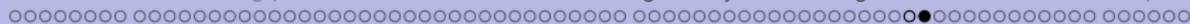
Systèmes multi-agents

Notion d'organisation

Une **organisation** est une communauté d'agents structurée par des règles spécifiant son fonctionnement.

Modèles d'organisation courants

- Agent/Groupe/Rôle (Ferber & Gutknecht) : les règles décrivent les conditions à remplir pour qu'un agent intègre un groupe donné ou joue un rôle particulier dans l'organisation.
- Modèle de type OrBAC : les agents jouent des rôles, les groupes sont en fait des organisations imbriquées.



Systèmes multi-agents

Notion d'organisation

Une **organisation** est une communauté d'agents structurée par des règles spécifiant son fonctionnement.

Modèles d'organisation courants

- Agent/Groupe/Rôle (Ferber & Gutknecht) : les règles décrivent les conditions à remplir pour qu'un agent intègre un groupe donné ou joue un rôle particulier dans l'organisation.
- Modèle de type OrBAC : les agents jouent des rôles, les groupes sont en fait des organisations imbriquées.

Notion d'institution

Une **institution** est une organisation dotée de **normes**, règles de fonctionnement que doivent (ou devraient) suivre les agents.

Modélisation des normes

Classification des normes (Tuomela 1995)

- **r-normes** ou règles : les agents les respectent parce qu'il y a un accord entre eux pour les respecter (lois, réglementations) ;
- **s-normes** ou normes sociales : les agents les respectent parce qu'ils savent que les autres s'y attendent ;
- **m-normes** ou normes morales : les agents les respectent parce que leur conscience individuelle le leur dicte ;
- **p-normes** ou normes de prudence : les agents les respectent parce qu'elles constituent la stratégie la plus rationnelle.

Modélisation des normes

Et dans un système multi-agents ?

Système multi-agents normatif : SMA dans lequel le comportement des agents est régi par des normes.

Technique possible : la **régimentation**. On impose le respect des normes aux agents par voie programmatique.

Modélisation des normes

Et dans un système multi-agents ?

Système multi-agents normatif : SMA dans lequel le comportement des agents est régi par des normes.

Technique possible : la **régimentation**. On impose le respect des normes aux agents par voie programmatique.

- *Quid* de l'autonomie des agents alors ?
- Comment prétendre représenter l'intelligence humaine si l'on exclut les comportements en marge, en infraction ?
- Et s'il est souhaitable pour un agent de violer une norme ?
- Et s'il est nécessaire pour un agent de violer une norme ?

Modélisation des normes

Pourquoi raisonner sur les normes ?

- Adaptation automatique du fonctionnement de l'agent pour satisfaire les normes du système, pour « respecter les règles » ;
- Détection de comportements inadaptés chez soi ou chez les voisins, actions en conséquence (mesures de prudence, sanctions, comportement plus agressif, réorganisation des liens sociaux. . .) ;
- Intérêt du raisonnement logique pour anticiper sur les conséquences des normes, sur les éventuelles incohérences avec les buts poursuivis ou avec d'autres aspects ;
- Comparaison et critique de systèmes normatifs.



Modélisation des normes

Quelques fonctionnalités nécessaires, souhaitables, désirables... pour le formalisme choisi :

- Représentation d'obligations, d'interdictions, de permissions ;
- Représentation de violations, de sanctions ;
- Prise en compte de l'aspect multi-agent ;
- Prise en compte d'autorités normatives multiples ;
- Gestion des conflits normatifs ;
- Intégration du temps dans le système ;
- ...

Modélisation des normes

Les logiques déontiques sont multiples, souvent fondées sur la Logique Déontique Standard (SDL) : axiomatique KD.

$$(K) \quad Ob (\varphi \rightarrow \psi) \rightarrow (Ob \varphi \rightarrow Ob \psi)$$

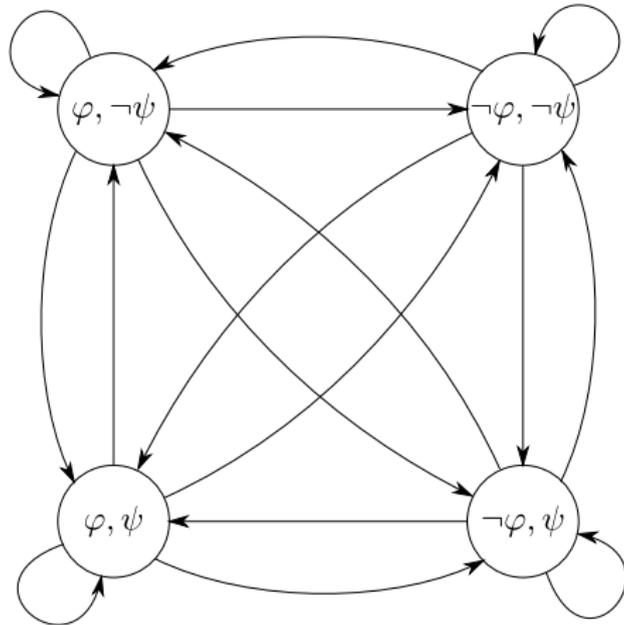
$$(D) \quad Ob \varphi \rightarrow Per \varphi$$

La logique déontique est souvent associée à une logique temporelle pour augmenter l'expressivité des obligations.

Limitations

La logique déontique standard (comme la plupart des autres logiques déontiques) souffre de nombreux **paradoxes**, dont certains très complexes et difficilement évitables.

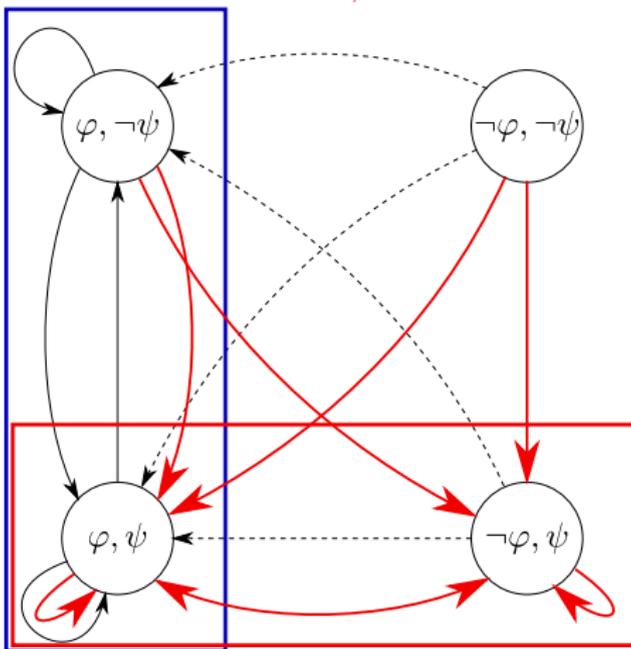
Modélisation des normes



Modélisation des normes

Bel φ

Ob ψ



Modélisation des émotions

Exemple adapté des travaux de Carole Adam, Dominique Longin et Andreas Herzig

Modalités de base

- Croyance : $Bel_i \varphi$;
- Temps : G, F, H, P ;
- Probabilité : $Prob_i \varphi$ (probabilité subjective, « croyance faible ») ;
- Choix : $Choice_i \varphi$ (désirabilité d'une formule) ;
- Idéalité : $Idl_i \varphi$ (modalité déontique) ;
- Action : $Happens_\alpha \varphi, Done_\alpha \varphi, After_\alpha \varphi, Before_\alpha \varphi$.

Divers axiomes font le lien entre les modalités de base.

Modélisation des émotions

Modalité intermédiaire : $Expect_i \varphi \stackrel{def}{=} Prob_i \varphi \wedge \neg Bel_i \varphi$
 $Expect_i \varphi$: l'agent i s'attend à ce que φ soit vraie

Exemple d'émotions formalisées à partir des modalités de base

- **Jolie** d'un agent à cause d'un fait : $Joy_i \varphi \stackrel{def}{=} Bel_i \varphi \wedge Choice_i \varphi$;
- **Espoir** d'un agent : $Hope_i \varphi \stackrel{def}{=} Expect_i \varphi \wedge Choice_i \varphi$;
- **Soulagement** : $Relief_i \varphi \stackrel{def}{=} Bel_i \neg \varphi \wedge Expect_i \varphi \wedge Choice_i \varphi$;
- Satisfaction, crainte, déception, ressentiment, fierté, honte, admiration, etc.

→ démonstration de théorèmes pour montrer la cohérence de la formalisation avec le bon sens et les travaux en psychologie.

Mise en œuvre

Wooldridge identifie trois étapes de la conception de SMA dans lesquelles la logique (modale) est utilisée :

- **Spécifications** : on utilise la logique modale pour spécifier le fonctionnement d'un modèle d'agent, qu'il soit ou non implémenté par la suite ;
- **Développement** :
 - On peut considérer le modèle comme des spécifications exécutables (embryonnaire pour la logique modale) ;
 - On peut faire de la génération de code automatique à partir des spécifications (solutions spécifiques uniquement) ;
 - On peut raffiner le modèle théorique jusqu'à le rendre implémentable (nécessité de caractériser les réductions).
- **Vérification** : utilisation de logiques modales (surtout dynamiques et temporelles) pour faire de la déduction ou du *model checking*, afin de démontrer ou prouver des propriétés sur le système.

Mise en œuvre

Il existe des outils dédiés pour faire de la programmation logique modale, comme MOLOG (développé à l'IRIT) ou MPROLOG (université de Varsovie). Leur utilisation est limitée à un mode purement interactif et/ou à des axiomatiques prédéfinies.

Habituellement, on se fonde plutôt sur une implémentation directe des modalités par des prédicats Prolog.

La réduction d'une logique modale à un modèle implémentable passe forcément par une perte de complétude, qu'il faut caractériser.

Pour des logiques un peu exotiques, il peut également être délicat de garantir la correction.

D'une manière générale, l'utilisation des logiques modales pour le développement est limitée, on la réserve souvent aux spécifications (pour leur capacité à transmettre les concepts) et au *model checking* (où elles se révèlent utiles et efficaces).



- 1 Introduction
- 2 Fondements logiques
- 3 Agents et systèmes multi-agents
- 4 Exemples d'applications**
 - Dans le domaine de la santé
 - Dans le domaine des loisirs
 - Dans le domaine de la sécurité informatique
 - Autres domaines

Dans le domaine de la santé

Assistance au diagnostic

Utilisation de systèmes non logiques ou partiellement logiques, systèmes bayésiens, systèmes de règles et arbres de décision
 Domaine d'application privilégié des **systèmes experts**

Application liée : validation et critique de thérapies

Dans le domaine des loisirs

Cinéma

Utilisation de systèmes multi-agents et de modèles de « nuées » pour simuler des foules dans *Le Seigneur des Anneaux*.

Nombreuses références aux agents et à l'IA, évidemment...

Jeux vidéo

Utilisation des principes des SMA ou des agents cognitifs pour gérer des PNJ.

Utilisation de l'*affective computing* pour des jeux (de moins en moins) expérimentaux.

Dans le domaine de la sécurité informatique

Détection d'intrusions

Propositions pour gérer la sécurité d'un réseau informatique à l'aide d'agents mobiles.

Politiques de sécurité

Utilisation de la logique déontique pour rédiger et raisonner sur des politiques de sécurité techniques, conception de systèmes auto-adaptatifs pour la protection des données personnelles.



Autres domaines

Agents conversationnels animés

Ou *embodied conversational agents* : utilisés pour l'amélioration de l'IHM par l'introduction d'un dialogue naturel entre humain et agent artificiel.

Domaine militaire

Utilisation de systèmes multi-agents cognitifs pour gérer des patrouilles de drones autonomes (armée australienne).

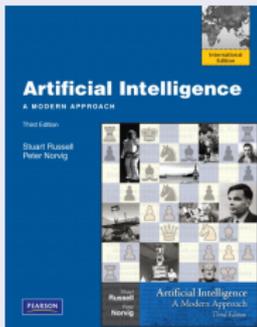
IA symbolique dans le domaine de la planification, de la simulation, de l'entraînement.

Gestion de crise

Utilisation de systèmes multi-agents adaptatifs pour coordonner la gestion de crise à grande échelle.



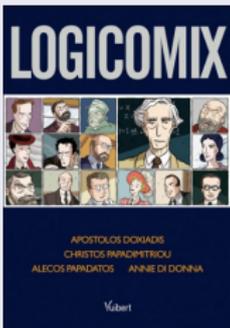
Bibliographie



Stuart Russel & Peter Norvig, *Artificial Intelligence, a modern approach*, 1994-2010

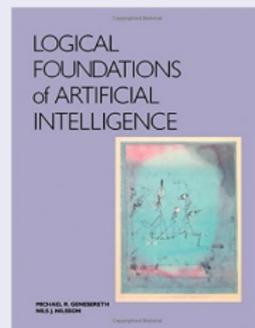


Bibliographie



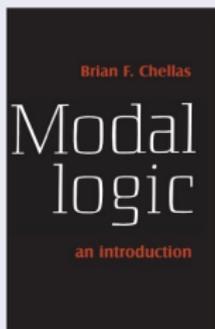
Apóstolos K. Doxiàdis & Christos Papadimitriou,
Logicomix, 2008

Michael R. Genesereth & Nils J. Nilsson,
Logical Foundations of Artificial Intelligence,
1986





Bibliographie

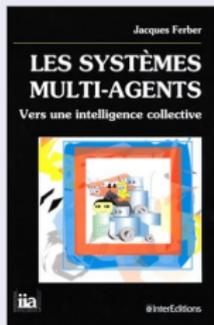


Brian Chellas, *Modal Logic, an Introduction*, 1980

Patrick Blackburn, Maarten de Rijke & Yde Venema, *Modal Logic*, 2001



Bibliographie



Jacques Ferber, *Les systèmes multi-agents : vers une intelligence collective*, 1995

Michael Wooldridge, *Introduction to MultiAgent Systems*, 1998-2009

