



Agents cognitifs

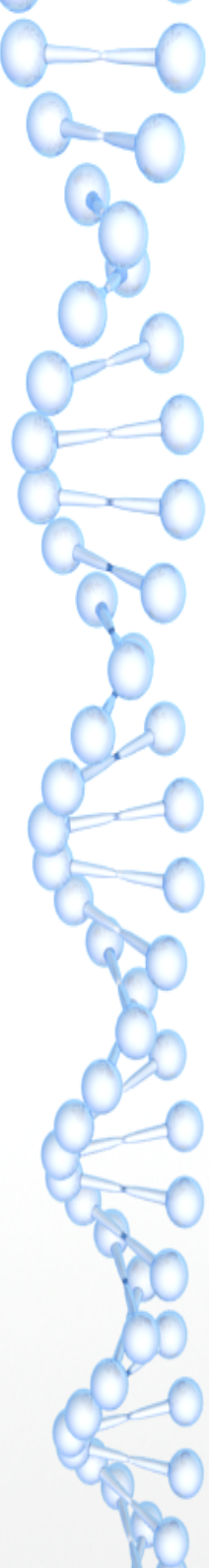
Syma – cursus SCIA

Julien Saunier - julien.saunier@ifsttar.fr



Plan

- Agents cognitifs?
- Le modèle BDI
- Approches hybrides



Agents cognitifs - caractéristiques

- Représentation explicite de l'environnement et des autres agents
- Peut tenir compte de son passé et dispose d'un but explicite – mode "social" d'organisation (planification, engagement)
- Petit nombre d'agents, hétérogènes à gros grain
- Les relations entre agents s'établissent en fonction des collaborations nécessaires à la résolution du problème



Cycle

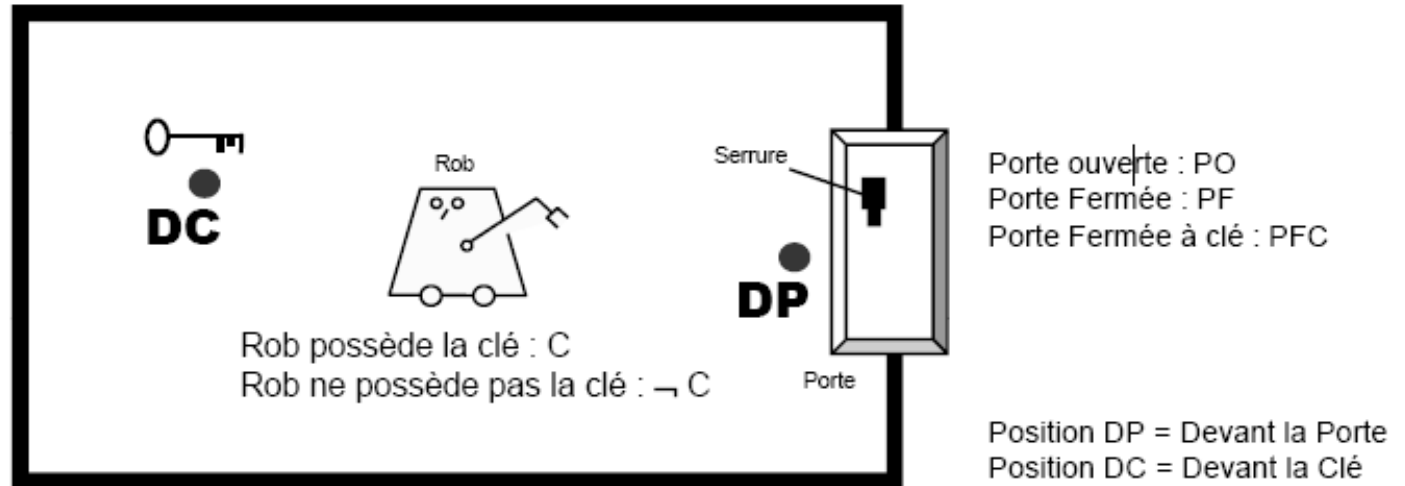
Données manipulées

- Ensemble d'états
- Ensemble de percepts
- Ensemble d'actions

```
states := initialise_state();  
  
while (true) {  
  percepts := see();  
  states := update_states(percepts)  
  action := deliberate(states);  
  execute(action);  
}
```

Agent Délibératif - Exemple

Problème :
« Rob le robot doit sortir
de la pièce ».



- $E = \{ PO, PF, \dots \}$
- $Etat = \{ (but, action) , C, \neg C \}$,
 - action = sortir | chercher-clef | hasard | ramasser
- Option : Si $\langle etat \rangle^*$ et $\langle perception \rangle^*$ alors option $\equiv \langle etat \rangle^*$
- Choix : Si $\langle etat \rangle^*$ et $\langle perception \rangle^*$ alors choix $\equiv \langle etat \rangle$
- Execute : Si $\langle (but, action) \rangle$ alors action



Agent Délibératif

- Logique de représentation des connaissances et croyances
 - Pallier les limites des logiques classiques
 - Syntaxique,
 - Sémantique.
 - Distinguer connaissances et croyances suppose de :
 - Prendre en compte leur différence qualitative
 - Prendre en compte la relativité d'évaluation des croyances.

Le modèle BDI

- Modèle de représentation des connaissances

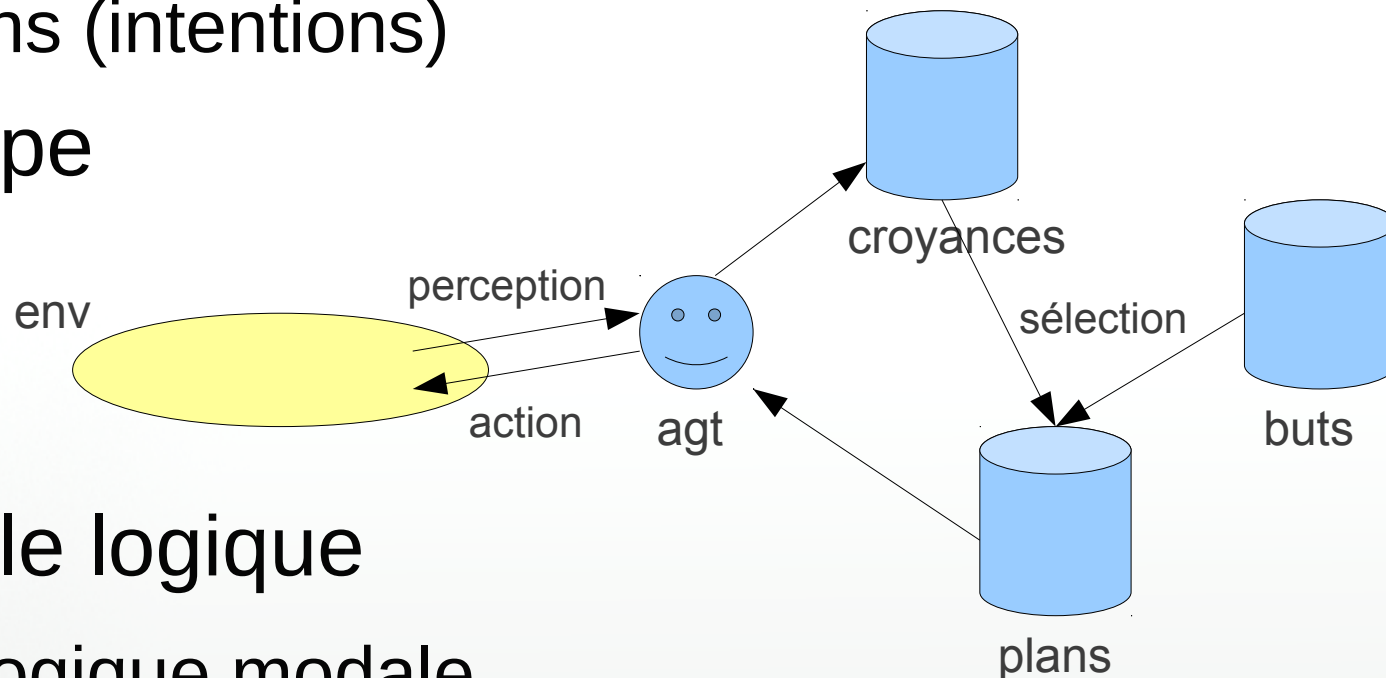
- Agents rationnels

 - Croyances (sur les autres, sur l'environnement)

 - Buts

 - Plans (intentions)

- Principe



- Modèle logique

 - logique modale

Logique classique

• Syntaxe

- $V = \{ a, b, c, \dots \neg, \wedge, \vee, \rightarrow, \equiv, () \}$
- atomes connecteurs
- $L_0 =$ Langage propositionnel $\subset V^*$
 - (i) Tout atome appartient à L_0 .
 - (ii) Si $\varphi, \psi \in L_0$ alors $\neg\varphi, (\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \rightarrow \psi), (\varphi \equiv \psi) \in L_0$.
 - (iii) Il n'existe pas d'autres façons d'obtenir un élément de L_0 .

• Sémantique

- On appelle **interprétation** toute application $I : \{ \text{atomes} \} \rightarrow \{V, F\}$.
Vrai et **Faux** sont **appelées valeurs de vérité**.
- On appelle **tautologie** toute formule φ telle que $I[\varphi] = V$ pour toute interprétation I . On écrit alors : $\models \varphi$
- On note $\varphi \models \psi$ (ψ conséquence logique de φ) si pour toute interprétation $I, I[\varphi] = V$ implique $I[\psi] = V$.

Logique classique (2)

- **V**: ensemble de symboles
- **E**: ensemble d'expressions bien formées $\in V^*$
- **A**: ensemble d'axiomes ($A \subset E$)
- **R**: ensemble de règles de déduction de la forme:
 - $r_i : f_1, f_2, \dots, f_n \vdash g$ avec $f_i, g \in E$
- SF0 Un système formel pour la logique des propositions
 - **V** = { a, b, c, ... \neg , \Rightarrow , (,) }
 - **E** = ensemble des expressions bien formées sur **V***
 - **A** = 3 axiomes :
 - A1: $(A \Rightarrow (B \Rightarrow A))$
 - A2: $((A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C)))$
 - A3: $((\neg A \Rightarrow \neg B) \Rightarrow (B \Rightarrow A))$
 - **R** = une unique règle (Modus Ponens):
 - $A, (A \Rightarrow B) \vdash B$



Exercice

représenter les connaissances suivantes en logique des prédicats et vérifier si elles sont des théorèmes...

1. Jean et Henri ont le même âge, autrement Jean est plus vieux que Henri. Or si Jean et Henri ont le même âge, alors Marie et Jean n'ont pas le même âge. Mais si Jean est plus vieux que Henri, alors Jean est plus vieux que Paul. Donc Marie et Jean n'ont pas le même âge ou Jean est plus vieux que Henri.
2. Si Jean est en prison, alors il n'est pas une nuisance pour sa famille. S'il n'est pas en prison, alors il est en disgrâce. S'il n'est pas en disgrâce, alors il est dans l'armée. S'il boit en plus, il est une nuisance pour sa famille. Donc il ne boit pas ou il est dans l'armée.
3. Si le suspect a commis le vol, alors il avait un complice. Mais le suspect avait un complice seulement si un butin important a été emporté. Or cela n'est pas le cas. Donc le suspect n'a pas commis le vol.
4. Au moins un de Pierre, Jean et Jacques est coupable. Or, si Jean est coupable, Pierre l'est aussi. En outre, si Jacques est coupable, les deux autres le sont aussi. Donc Pierre est coupable.

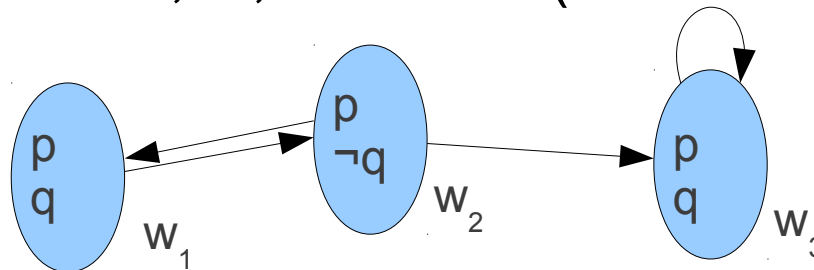
Logiques modales

- Modalité \neq fonction de vérité
- Sémantique de Kripke \rightarrow mondes possibles

Modèle = $\langle W, R, S \rangle$

(seulement S en LPPO)

Ex :



Logique du savoir

- Ajout d'un opérateur K

- Kf = je sais que f

- Théorèmes souhaités

- Modus ponens $(K(\varphi \supset \psi)) \rightarrow (K\varphi \supset K\psi)$

- Ce que je sais est vrai $(K\varphi) \rightarrow \varphi$

- Je sais ce que je sais $(K\varphi) \rightarrow (K K \varphi)$

- Je sais ce que je ne sais pas $(\neg K \varphi) \rightarrow (K \neg K \varphi)$

- Système S5

Logique d'opinion

- Opérateur B

- Bf = je crois que f

- Axiome (D)

- Axiome (T) trop fort

- on peut croire qqch de faux

$$\cancel{(B\varphi) \rightarrow \varphi}$$

- Axiome (D)

$$(B\varphi) \rightarrow \neg B\neg\varphi$$

- on ne peut pas croire une chose et son contraire

- Système KD45 (Halpern & More, 85)

- Axiomes H1, H2, D, S4 et S5



Agent BDI

- Agent BDI [Rao, Georgeff 91], selon ce modèle un agent possède :
 - Des croyances (Belief) sur lui-même et le monde,
 - Des désirs (Desire) éventuellement contradictoires,
 - Un ensemble d'intentions (Intention) non conflictuelles,
 - Un mécanisme de raisonnement pour mettre à jour ses croyances, choisir ses désirs et générer de nouvelles intentions.



Agent BDI (2)

- Logique propositionnelle d'opinion
 - Modalité B, système KD45
 - K : les connaissances d'un agent sont closes par implication
 - $B(p \rightarrow q) \rightarrow (B(p) \rightarrow B(q))$
 - D : les croyances d'un agent sont consistantes
 - $B(p) \rightarrow \neg \perp B(\neg \perp p)$
 - 4 : axiome d'introspection positive.
 - $B(p) \rightarrow B(B(p))$
 - 5 : axiome d'introspection négative
 - $\neg \perp B(p) \rightarrow B(\neg \perp B(p))$



BDI (3)

- Agent → logique d'action
 - Prédicats d'action
 - Opérateurs : Agent(i, a), Done(a, pre), Feasible($a, post$)
 - Logique propositionnelle
 - mais pas de quantificateurs dans les formules
 - Modalités B (croyance), D (but) et I (plan)
- Multi-agent → multi-modalités & opérateurs
 - Modalités B_i → l'agent i croit que ...
 - Notation Agent $_i(a)$

Agent BDI (4)

• Règles :

- Compatibilité but et croyance :
 - $D(\varphi) \rightarrow B(\varphi)$
- Compatibilité But et intention :
 - $I(\varphi) \rightarrow D(\varphi)$
- Intention mène à l'action :
 - $I(\text{faire}(\text{act})) \rightarrow \text{faire}(\text{act})$ avec *act* une action primitive.
- Raisonnement
 - $B(\neg p) \wedge D(p) \wedge B(\text{act} \rightarrow p) \wedge B(\text{possible}(\text{act}))$ alors $I(\text{faire}(\text{act}))$.



Sémantique

- Opérateurs d'actions:

- $\text{Agent}(i,a)$: l'agent i peut faire l'action a
→ décrit les capacités (au sens physique)
- $\text{Done}(a,p)$: l'action a a été faite et avant, p était vraie
→ décrit les préconditions
- $\text{Feasible}(a,p)$: à l'issue de l'action a , p sera vraie
→ décrit les effets d'actions

- Notations

- $\text{Done}(a) = \text{Done}(a,T)$ L'action a a été effectuée
- $\text{Feasible}(a) = \text{Feasible}(a,T)$ L'action peut être effectuée
- $\text{Agenti}(a) = \text{Agent}(i,a)$ homogénéité avec les modalités

- Point de vue global (tout le SMA)

Sémantique (2)

- Modalités

- $B_i f$: l'agent i croit que f est vraie
- $D_i f$: l'agent i a pour but que f soit vraie, un jour...
- $I_i f$: l'agent i a pour intention que f soit vraie, c'est-à-dire qu'il va faire ce qu'il faut pour.
→ Ce sont les buts instanciés (ou plans)

- Exemples:

- $B(\text{onTable}(\text{cube_A}))$
- $B(\neg \text{onTable}(\text{cube_A}))$
- $I(\text{onTable}(\text{cube_B}))$
- $B(\text{does}(\text{take}(\Psi)) \wedge \text{onTable}(\Psi) \rightarrow \neg \text{onTable}(\Psi))$

BDI (5)

b : beliefs, g : desires, i : intentions, eq : event queue

$(b, g, i) := \text{initialize}();$

repeat

$options := \text{option_generator}(eq, b, g, i);$

$selected := \text{deliberate}(options, b, g, i);$

$i := selected \cup i;$

$\text{execute}(i);$

$eq := \text{see}();$

$b := \text{update_beliefs}(b, eq);$

$(g, i) := \text{drop_successful_attitudes}(b, g, i);$

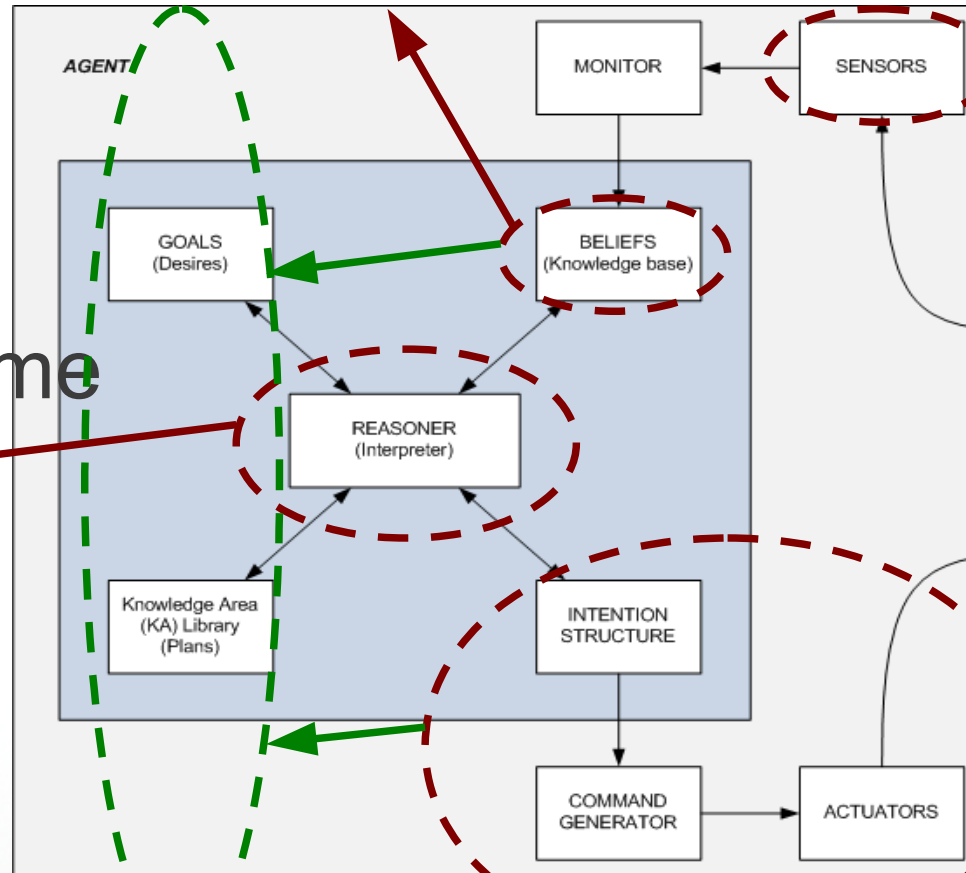
$(g, i) := \text{drop_impossible_attitudes}(b, g, i);$

forever

Exemple d'architecture BDI: PRS

Représentation
interne du monde

Perceptions



Un programme
"intelligent"

buts et plans
(hérité de STRIPS)

Actions



Applications

- Premier modèle formel d'agent rationnel

Langage AgentSpeak – PF Jason

- Planification multi-agent

... en utilisant l'inférence logique

- Ajouts de modalités

→ Communication agent-humain

→ Agents émotionnels

etc.

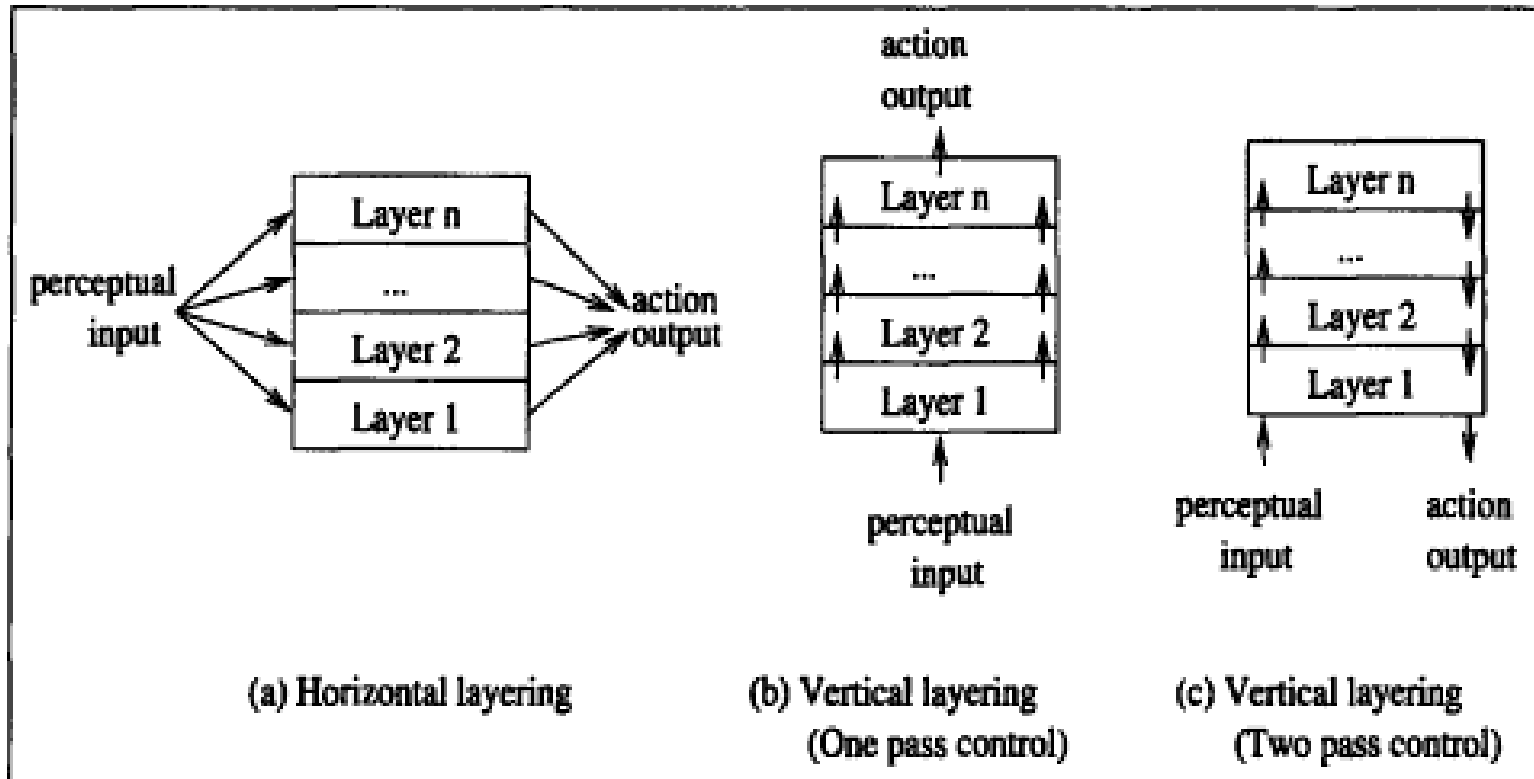
+ Modèle choisi par FIPA pour décrire la sémantique des interactions



Architectures hybrides

- Les architectures réactives sont trop simples pour traiter des problèmes complexes
- Les architectures cognitives sont trop lourdes à développer/exécuter pour une réaction rapide
- Le principe d'une architecture hybride est de combiner des couches réactives et cognitives

Architectures hybrides: typologie



J. P. Müller, M. Pischel, and M. Thiel. Modelling reactive behaviour in vertically layered agent architectures. In M. Wooldridge and N. R. Jennings, editors, *Intelligent Agents: Theories, Architectures, and Languages (LNAI Volume 890)*, pages 261-276. Springer-Verlag: Berlin, Germany, January 1995.

Architecture hybride

Exemple: TouringMachine

