

Aprentatge de Comportaments

APRENTATGE EN AGENTS I SISTEMES
MULTIAGENTS

Programa doctorat en IA

LSI-UPC

Aprentatge de comportaments: Índex

- Aprendre que?
- Aprendre a partir de que?
- Aprendre com?
 - Aprentatge per reforç

Representacions de comportaments

- Comportaments reactius, no deliberatius (es l'únic que es pot aprendre)
- Comportaments com a mapping entre percepcions i accions
 - Look-up tables
 - Neural networks
 - Rules $s \rightarrow a$
 - Autòmats finits

Lock-up tables

- Primer model simple de mapping

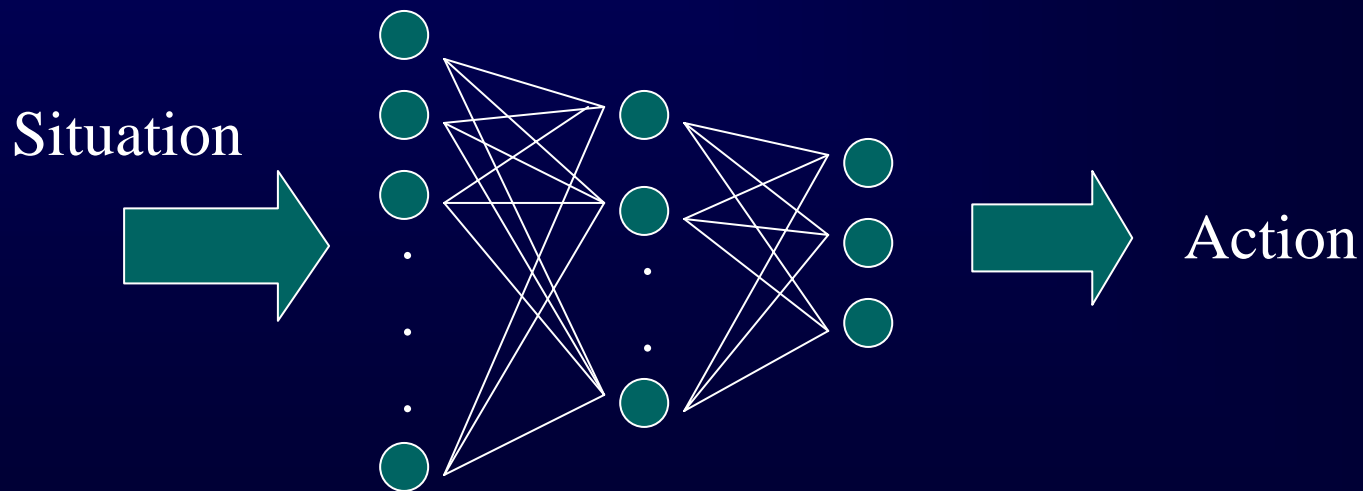
s1	a1
s2	a2
s3	a3
s4	a4
sN	aN

Lock-up tables

- Avantatges:
 - Simples. Accés ràpid, no calen pre-càlculs de capacitat
- Inconvenients
 - No generalitzen de forma natural
 - Moltes entrades possibles inclòs per casos de joguina

Neural Networks

- Mapping situació-acció representat amb xarxes neurals



Neural Networks

- Avantatges
 - Generalitzen
 - Compactes
- Inconvenients
 - Els de les NN: cros-talk, forgetting, ...

Regles

- Mapping com a llista de decisió de regles:

Llista de regles

$$1. \quad p1 \wedge p3 \wedge \neg p7 \quad \longrightarrow \quad a3$$

$$2. \quad (p5 \wedge \neg p3) \vee \neg p7 \quad \longrightarrow \quad a1$$

.

.

.

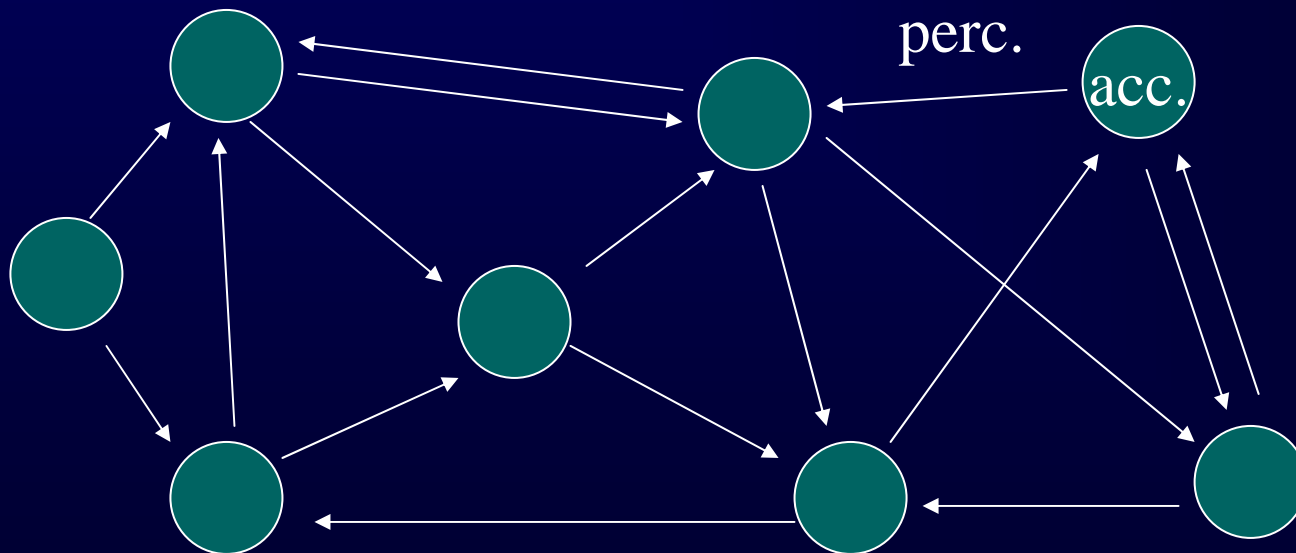
$$3. \quad \neg p9 \vee \neg p6 \quad \longrightarrow \quad a6$$

Regles

- Avantatges
 - Explicites
 - Generalitzen

Autòmates finits

- Cada fletxa es una percepció, cada node es estat i acció a realitzar



Autòmats finits

- Avantatges
 - Forma compacte de representar comportaments, de forma explícita
 - Facilitat de representar problemes POMDPs
- Problemes
 - Sols poden representar comportaments representables com llenguatges regulars

Aprendre a partir de que?

- Dues possibilitats
 - Supervisat
 - Per reforç
- No hi ha aprenentatge estrictament “no-supervisat” en aprendre comportaments (comportaments són dirigits cap a objectius per definició)

Aprenentatge supervisat

- Agent mestre realitza comportament motor
 - Exemple: Robot i pintura
- Agent mestre realitza comportament pel seu compte. L'agent aprenent interpreta el comportament realitzat pel mestre i l'intenta imitar
 - Exemple: Seguiment d'un robot mòbil

Problemes de l'aprenentatge supervisat

- Cal tenir un mestre
 - Cal dissenyar el comportament
 - Cal programar-lo
- Problemes en disseny del comportament. No sempre solucions òptimes (grounding problem)
- Normalment no val la pena aprendre d'aquesta manera

Aprenentatge per reforç

Senyal de reforç avalua el comportament de l'agent indicant-li si es correcte o no

Aprenentatge per reforç

- Avantatges
 - No cal conèixer el comportament a aprendre “a priori”
 - Es fàcil construir avaluadors del comportament
 - No grounding problem
- Problemes
 - Computacionalment sembla més complex