Abstract agent architectures

4ICT2: Information Management

Saturnino Luz

mailto:luzs@cs.tcd.ie

Trinity College, Department of Computer Science

October 25, 2006

Outline

- Recapitulation
  - Properties one would like to model
- Mathematical notation
- A generic model
- A purely reactive model
- A state-transition model
- Modelling concrete architectures
- References: [Weiss, 1999, sections 1.3 and 1.4]

Recapitulating...

- Agent properties:
  - Autonomy: acting independently
  - Situatedness: sensing and modifying the environment
  - Flexibility:
    * re-activity to changes
    * pro-activity in bringing about environmental conditions under which the agent's goals can be achieved, and
    * sociability: communicating with other agents, collaborating, and sometimes competing

P.E.A.S descriptions

Agent Type	Performance measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, minimize costs	Patient, hospital	Questions, tests, treatments	Symptoms, findings, patient's answers
Satellite image analysis system	Correct categorization	Satellite link	Print a categorization of scene	Pixels of varying intensity, color
Part-picking robot	% parts in correct bins	Conveyor belt with parts	Pick up parts and sort into bins	Pixels of varying intensity
Interactive English tutor				

From [Russell and Norvig, 2003]

P.E.A.S descriptions

Agent Type	Performance measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, minimize costs	Patient, hospital	Questions, tests, treatments	Symptoms, findings, patient's answers
Satellite image analysis system	Correct categorization	Satellite link	Print a categorization of scene	Pixels of varying intensity, color
Part-picking robot	% parts in correct bins	Conveyor belt with parts	Pick up parts and sort into bins	Pixels of varying intensity
Interactive English tutor	Maximize student's score on test			

From [Russell and Norvig, 2003]

P.E.A.S descriptions

Agent Type	Performance measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, minimize costs	Patient, hospital	Questions, tests, treatments	Symptoms, findings, patient's answers
Satellite image analysis system	Correct categorization	Satellite link	Print a categorization of scene	Pixels of varying intensity, color
Part-picking robot	% parts in correct bins	Conveyor belt with parts	Pick up parts and sort into bins	Pixels of varying intensity
Interactive English tutor	Maximize student's score on test	Set of students; testing agency		

From [Russell and Norvig, 2003]

P.E.A.S descriptions

Agent Type	Performance measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, minimize costs	Patient, hospital	Questions, tests, treatments	Symptoms, findings, patient's answers
Satellite image analysis system	Correct categorization	Satellite link	Print a categorization of scene	Pixels of varying intensity, color
Part-picking robot	% parts in correct bins	Conveyor belt with parts	Pick up parts and sort into bins	Pixels of varying intensity
Interactive English tutor	Maximize student's score on test	Set of students; testing agency	Print exercises, suggestions, corrections	

From [Russell and Norvig, 2003]

P.E.A.S descriptions

Agent Type	Performance measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, minimize costs	Patient, hospital	Questions, tests, treatments	Symptoms, findings, patient's answers
Satellite image analysis system	Correct categorization	Satellite link	Print a categorization of scene	Pixels of varying intensity, color
Part-picking robot	% parts in correct bins	Conveyor belt with parts	Pick up parts and sort into bins	Pixels of varying intensity
Interactive English tutor	Maximize student's score on test	Set of students; testing agency	Print exercises, suggestions, corrections	Keyboard input

From [Russell and Norvig, 2003]

or, alternatively...

- P.A.G.E.: Percepts, Actions, Goals and Environment

Agent Type	Percepts	Actions	Goals	Environment
Medical diagnosis system	Symptoms, findings, patient's answers	Questions, tests, treatments	Healthy patient, minimize costs	Patient, hospital
Satellite image analysis system	Pixels of varying intensity, color	Print a categorization of scene	Correct categorization	Images from orbiting satellite
Part-picking robot	Pixels of varying intensity	Pick up parts and sort into bins	Place parts in correct bins	Conveyor belt with parts
Interactive English tutor	Typed words	Print exercises, suggestions, corrections	Maximize student's score on test	Set of students

As in (Russell and Norvig, 1995)

Environment properties

- Fully vs. partially observable: whether agent's can obtain complete and accurate information about the environment
- deterministic vs. stochastic: whether each action is guaranteed to produce a single effect
- episodic vs. sequential: whether agent's next action depends only on the current state of the environment (episodic), or on assessment of past environment states (sequential)
- static vs. dynamic: whether the environment changes independently of the agent's actions
- discrete vs. continuous: whether the possible actions and percepts on an environment are finite (discrete environment) or not (continuous environment)
- single vs. multiple agents

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock						
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving						
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully					
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving						
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic				
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving						
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential			
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving						
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi		
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving						
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving						
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving						
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially					
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic				
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential			
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic		
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis						
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially					
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic				
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential			
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic		
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	single
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor						
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	single
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor	partially					
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	single
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor	partially	stochastic				
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	single
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor	partially	stochastic	sequential			
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	single
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor	partially	stochastic	sequential	dynamic		
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	single
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor	partially	stochastic	sequential	dynamic	discrete	
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

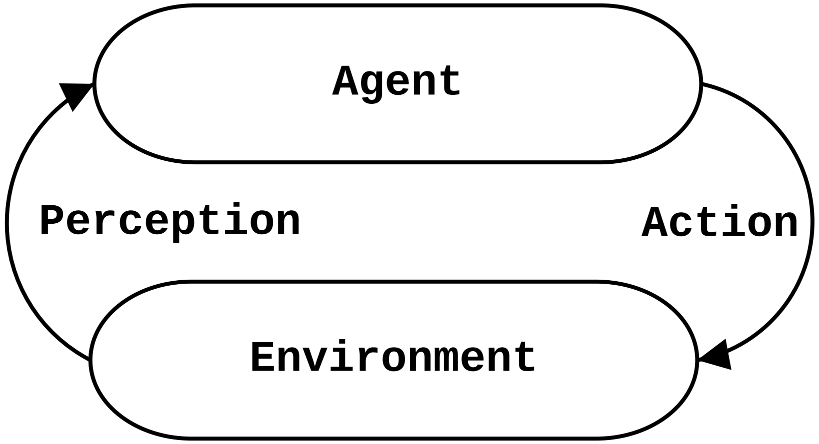
Types of environments

Environment	Observable	Deterministic	Episodic	Static	Discrete	Agents
Crossword puzzle	fully	yes	sequential	static	yes	single
Chess w/ clock	fully	strategic	sequential	semi	yes	multi
Poker	partially	strategic	sequential	static	discrete	multi
Backgammon	fully	stochastic	sequential	static	discrete	multi
Car driving	partially	stochastic	sequential	dynamic	continuous	multi
Medical diagnosis	partially	stochastic	sequential	dynamic	continuous	single
Image analysis	fully	deterministic	episodic	semi	continuous	single
Robot arm	partially	stochastic	episodic	dynamic	continuous	single
English tutor	partially	stochastic	sequential	dynamic	discrete	multi
Plant controller	partially	stochastic	sequential	dynamic	continuous	single

From [Russell and Norvig, 2003]

Abstract agent architectures

- Why?
- What we would like to describe:
  - agent
  - environment



  - their interactions

Notation

- A few tools from discrete maths and logic:
  - $A, S, ...$: sets
  - $\wp(S)$: the powerset of S
  - S^* : all sequences of elements (i.e. ordered subsets) of S
  - $\Rightarrow, \wedge, \vee, \neg$: material implication, conjunction, disjunction and negation
  - Quantifiers: $\forall$ and $\exists$

Defining an architecture

- A standard architecture is a 4-tuple:

$$Arch_s =< S, A, action, env >$$

where

  - $S = \{s_1, s_2, ..., s_n\}$ is the set of all (possible) environment states, and
  - $A = \{a_1, a_2, ..., a_n\}$ is the set of all actions an agent is capable of performing
  - $action$ is a function describing the agents behaviour, and
  - env is a function describing the "behaviour" of the environment

"Agenthood"

- An agent's behaviour will be characterised by the following function:

$$action : S^* \longrightarrow A$$

- Does having S^* as the domain of $action$ make the function most naturally suited to modelling episodic or sequential environments? Why?

"Agenthood"

- An agent's behaviour will be characterised by the following function:

$$action : S^* \to A$$

- Does having S^* as the domain of $action$ make the function most naturally suited to modelling episodic or sequential environments? Why?

- Requirement captured: the current action may depend on the interaction history (i.e. the sequence of environment states)

Environment dynamics

- Changes in the environments will be characterised by the following function:

$$env : S \times A \rightarrow \wp(S)$$

- Intuition: $env(s_j, a_k) = S'$ performing an action a_k on an environment whose state is s_j results in a number of scenarios (S')

- In the general case, what type of environment does env model?

Environment dynamics

- Changes in the environments will be characterised by the following function:

$$env : S \times A \longrightarrow \wp(S)$$

- Intuition: $env(s_j, a_k) = S'$ performing an action a_k on an environment whose state is s_j results in a number of scenarios (S')

- In the general case, what type of environment does env model?

  Answer: env models a non-deterministic environment.

Environment dynamics

- Changes in the environments will be characterised by the following function:

$$env : S \times A \rightarrow \wp(S)$$

- Intuition: $env(s_j, a_k) = S'$ performing an action a_k on an environment whose state is s_j results in a number of scenarios (S')

- In the general case, what type of environment does env model?

  Answer: env models a non-deterministic environment.

- If $|S'| = 1$, then the environment is deterministic.

Interaction history

- The agent-environment interaction will be characterized as follows:

$$h : s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} \dots \xrightarrow{a_{u-1}} s_u \xrightarrow{a_u} \dots$$

- h is a possible history of the agent in the environment iff:

$$\forall u \in \mathbb{N}, a_u = action(< s_0, \dots, s_u >) \tag{1}$$

and

$$\forall u > 0 \in \mathbb{N}, s_u \in env(s_{u-1}, a_{u-1}) \tag{2}$$

Characteristic behaviour

- Characteristic behaviour is defined as a set of interaction histories

$$hist = \{h_0, h_1, ..., h_n\}$$

where

- each h_i is a possible history for the agent in the environment

Invariant properties

- We say that ϕ is an invariant property of an agent architecture *iff*

$$\boxed{\text{For all histories } h \in hist \text{ and states } s \in S \text{ there is a property } \phi \text{ such that } s \models \phi}$$

- Note: $s \models \phi$ would translate as $\phi \in s$, in architectures where reasoning isn't performed
  - (e.g. reactive architectures)

Behavioural equivalence

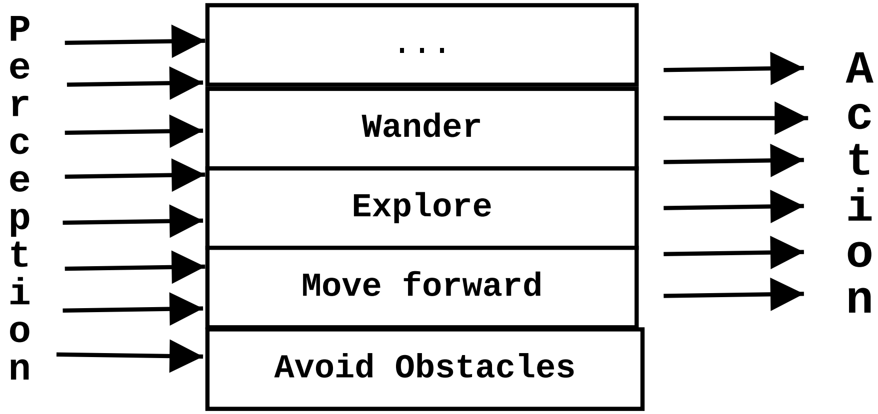
- Equivalence of behaviours is defined as follows (in an abstract architecture)

> An agent $ag1$ is regarded as equivalent to agent $ag2$ with respect to environment env_i iff $hist(ag1, env_i) = hist(ag2, env_i)$

- When the condition above holds for all environments env_i, then we simply say that $ag1$ and $ag2$ have **equivalent behaviour**

Modelling reactive agents

- Recapitulation: reactive architectures
  - production rules
  - a scheme for defining priorities in the application of rules (e.g. subsumption)



  - no reasoning (theorem proving or planning) involved

Abstract reactive agents

- Purely reactive agents can be modelled by assuming

$$action : S \to A$$

- Everything else remains as in the general (standard abstract architecture) case:

$$Arch_r = < S, A, action, env >$$

- Abstract reactive agents operate essentially on an episodic view of environments (i.e. they are memoryless agents). See slide 11

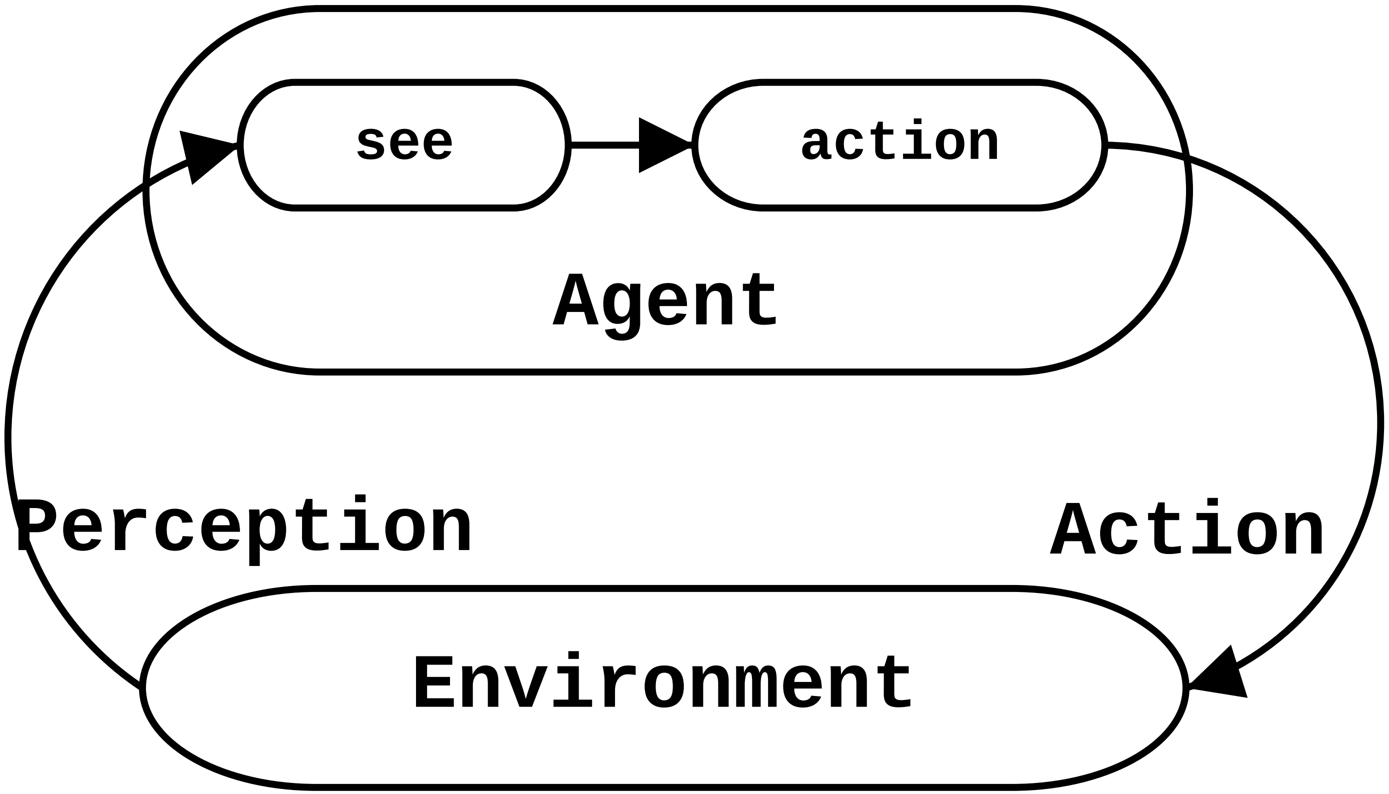
Purely reactive vs. standard agents

- Proposition 1:

> Purely reactive agents form a proper subclass of standard agents. That is, for any given environment description S, and action repertoire A:
>
> (i) every purely reactive agent is behaviourally equivalent standard agent, and
>
> (ii) the reverse does not hold.

Modelling Perception

- Refining the agent's decision function



- Types and sources of perception

Percepts and actions

- Facts perceived by an agent will be represented as set

$$P = \{p_0, p_1, ..., p_n\}$$

- The decision function becomes

$$action : P^* \rightarrow A$$

- which is then linked to environment states via the perception fumction

$$see : S \rightarrow P$$

Properties of perception

- If $s_i \neq s_j$ but $see(s_i) = see(s_j)$, we say that s_i and s_j are indistinguishable

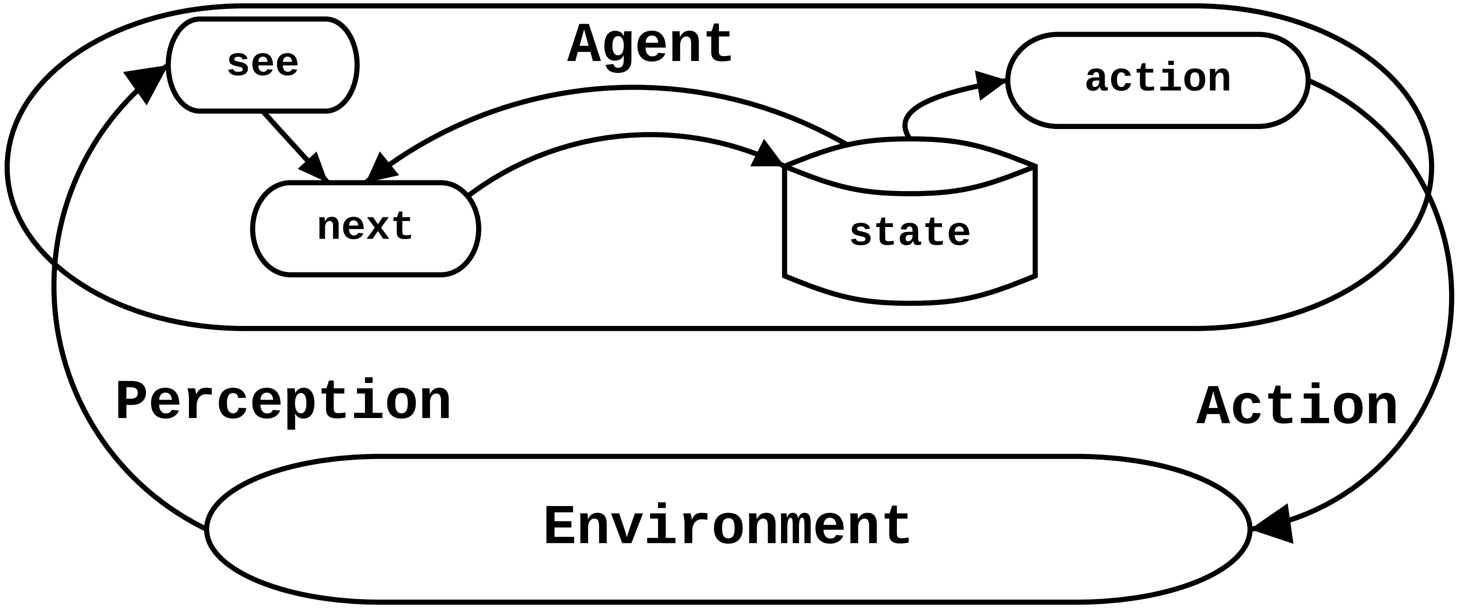
- Define $\equiv$, an equivalence relation over S by saying $s \equiv s'$ iff $see(s) = see(s')$

- If $| \equiv | = |S|$, then we say that the agent is perceptually ominiscient

- On the other hand, if $| \equiv | = 1$, then the agents perceptual ability is nil

Refining the interaction history

- Representation used so far: history as a sequence of environment states



- The next step: represent history (in the agent architecture) as environment changes as perceived by the agent

State-based decision function

- The state-based architecture will be represented as

$$Arch = < S, A, P, I, action, env, see, next >$$

  where

- I is the set of all internal states of an agent,
- $see : S \rightarrow P$,
- $action : I \rightarrow A$, and
- $next : I \times P \rightarrow I$

Properties of State-based architectures

- Proposition 2:

> State-based architectures are equivalent to standard architectures with respect to the behaviours they are able to represent.

Exercise

- Prove Proposition 2.

Further information

- [Weiss, 1999]: overview; this presentation mostly followed the material found there;
- [Russell and Norvig, 2003]:
  $\boxed{\text{agent} = \text{architecture} + \text{program}}$
- [Genesereth and Nilsson, 1987]: foundations of a theory of rational agency

References

Michael Genesereth and Nils Nilsson. Logical Foundations of Artificial Intelligence. Morgan Kaufmann, Los Altos, CA, 1987.

Stuart J. Russell and Peter Norvig. Artificial Intelligence. A Modern Approach. Prentice-Hall, Englewood Cliffs, 2nd edition, 2003.

Gerhard Weiss. Multiagent Systems. MIT Press, Cambridge, 1999.