

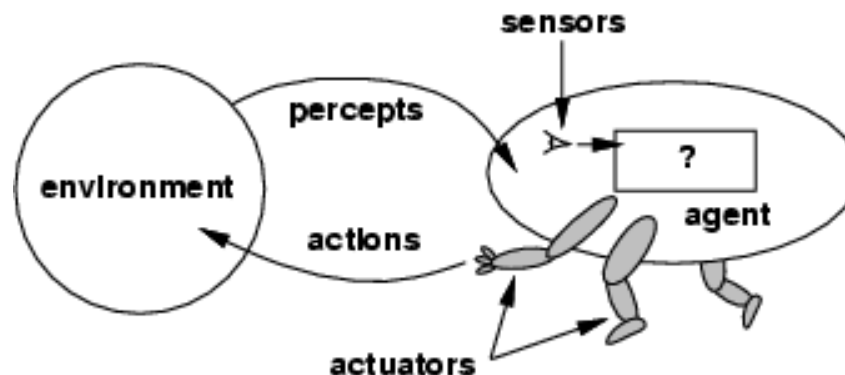
***Системи, основани на знания - зимен семестър,
2020/2021 учебна година***

***Тема 1, част 2:
Интелигентни агенти***

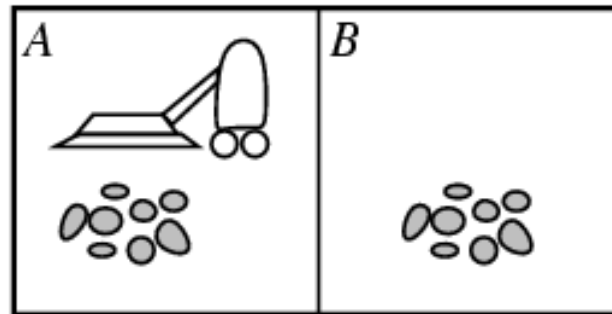
Дефиниция и действие на интелигентните агенти

- **Агент**: нещо, което **възприема** средата (чрез **рецептори** или **сензори**) и **въздейства** върху нея (чрез **ефектори**). В най-широк смисъл този термин обхваща хора, роботи и програми.
- **Интелигентен (разумен, рационален) агент**: агент с целесъобразно (според определени критерии) поведение, изграждано въз основа на възприетите до текущия момент сведения, вградените знания на агента и възможностите му да действа.

- **Изображение (*mapping*):** съответствие между възприятията и очакваните в отговор действия, чрез което се определя поведението на агента. Може да е реализирано чрез таблица (която в много случаи би трябвало да е безкрайна), функция или стратегия за самообучение.



Пример: Vacuum-cleaner world



- ✓ Percepts: location and contents, e.g., [A,Dirty]
- ✓ Actions: *Left*, *Right*, *Suck*, *NoOp*

Основни характеристики на интелигентните агенти

- **автономност**: свойството на агента да определя поведението си въз основа на натрупания опит (за сметка на вградените знания) и да действа по своя инициатива
- **адаптируемост**: способност на агента
 - да се самообучава в действие
 - да реагира на изменения в средата и да се приспособява към тези от тях, които са трайни
 - да общува с други агенти чрез някаква знакова система

- **колаборативност:** способност за сътрудничество – работа в **многоагентни системи**, включваща разделяне на задачите, обединение на резултатите и обмен на знания и опит
- **мобилност:** самонасочено придвижване на агент от един носител към друг в мрежова среда

Оценка на ефективността на поведението (performance measure): формулировката ѝ е съществен елемент на началната спецификация при проектирането на даден интелигентен агент.

PEAS: Performance measure, Environment, Actuators, Sensors
(елементи на спецификацията на всеки конкретен интелигентен агент)

Пример 1

- ✓ Agent: Medical diagnosis system
- ✓ Performance measure: Healthy patient, minimize costs, lawsuits
- ✓ Environment: Patient, hospital, staff
- ✓ Actuators: Screen display (questions, tests, diagnoses, treatments, referrals)
- ✓ Sensors: Keyboard (entry of symptoms, findings, patient's answers)

Пример 2

- ✓ Agent: Interactive English tutor
- ✓ Performance measure: Maximize student's score on test
- ✓ Environment: Set of students
- ✓ Actuators: Screen display (exercises, suggestions, corrections)
- ✓ Sensors: Keyboard

Строеж на интелигентните агенти

агент = архитектура + програма

- **програма на агента:** функция, реализираща изображението от възприятия към действия. В общия случай програмата разполага с памет, в която съхранява част от историята на получените възприятия и предприетите действия, за да може отговорът на всяко въздействие да не бъде функция само на това въздействие. Тялото на програмата представлява цикъл, всяка стъпка на който включва възприятие (отразявано в паметта), избиране и осъществяване на действие (също отразявано в паметта). Задачата на изкуствения интелект е именно създаването на такива програми.

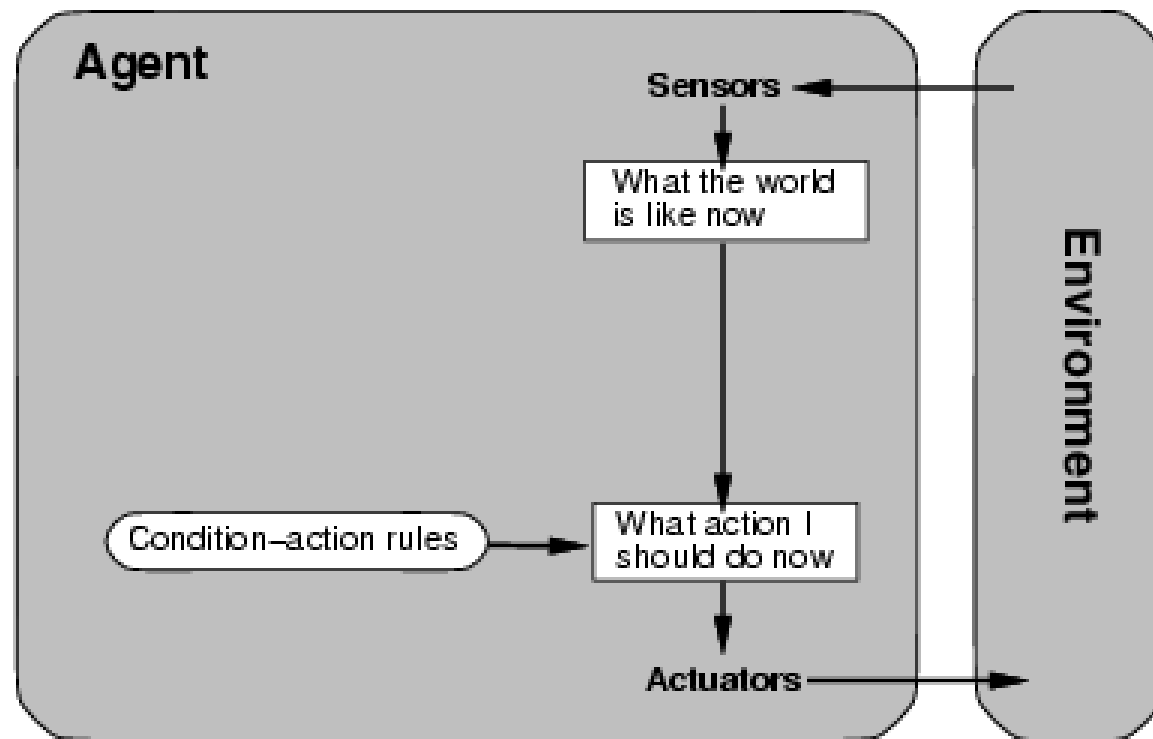
- **архитектура**: това, върху което работи програмата – обикновен компютър или специализирано устройство. Може да включва платформено независимо програмно осигуряване, с помощта на което агентът да бъде програмиран на високо ниво.

За реализацията на програмата на агента могат да се използват различни, често нетрадиционни **техники**: разсъждения, основани на знания (най-често правила), статистически анализ, размита логика, невронни мрежи, еволюционно (генетично) програмиране и др.

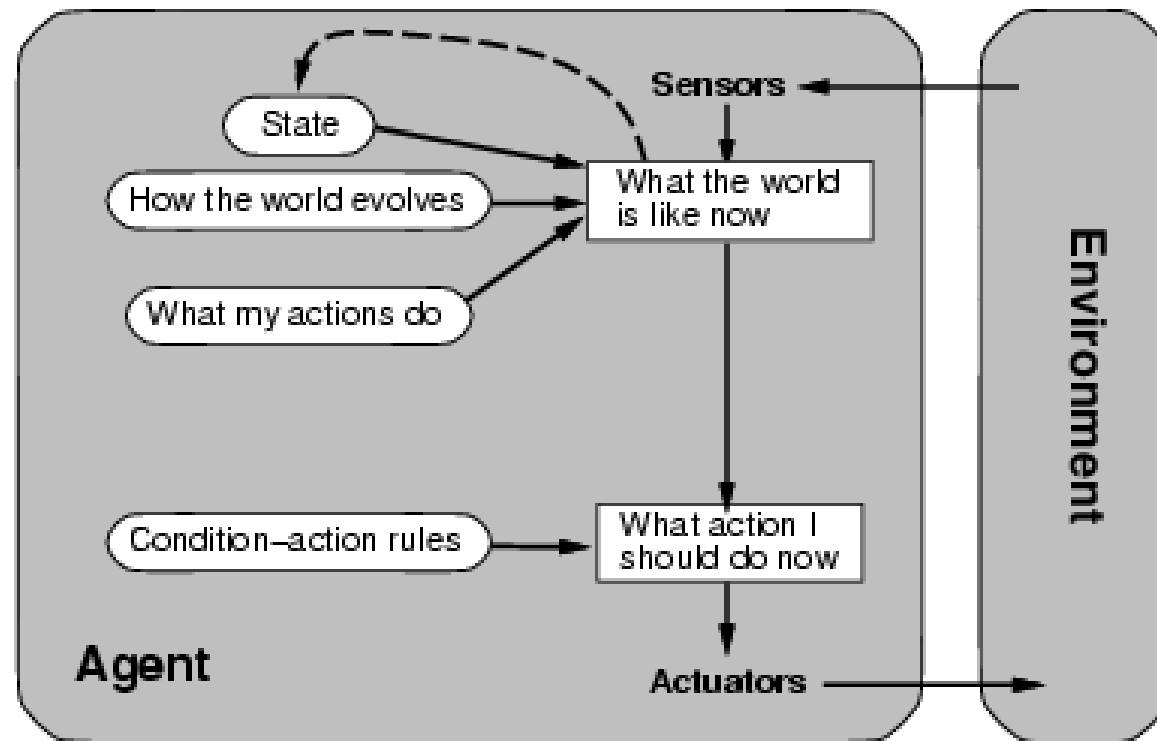
Типове агентни програми

Могат да бъдат обособени четири основни типа агентни програми, подредени по-долу по степента им на общност:

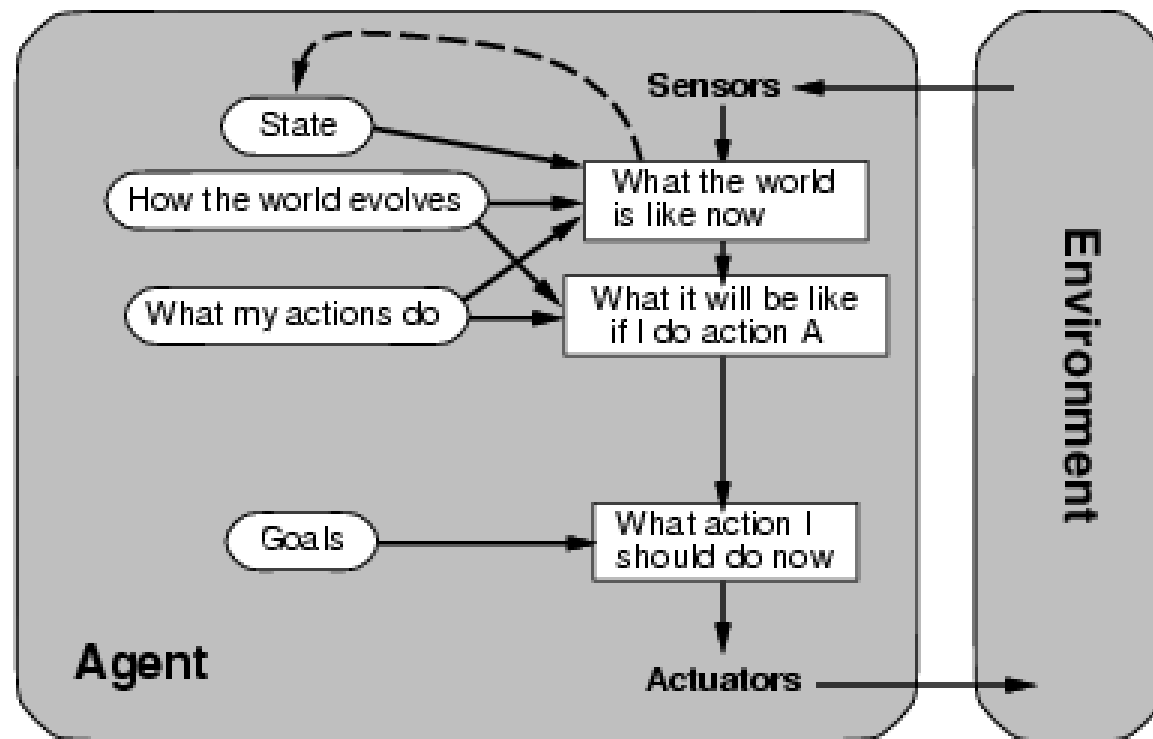
- **прости рефлекторни агенти:** действията се определят от правила, от които агентът избира това, което съответства най-точно на ситуацията. Програмите могат да бъдат много ефективни, но приложимостта им е малка.



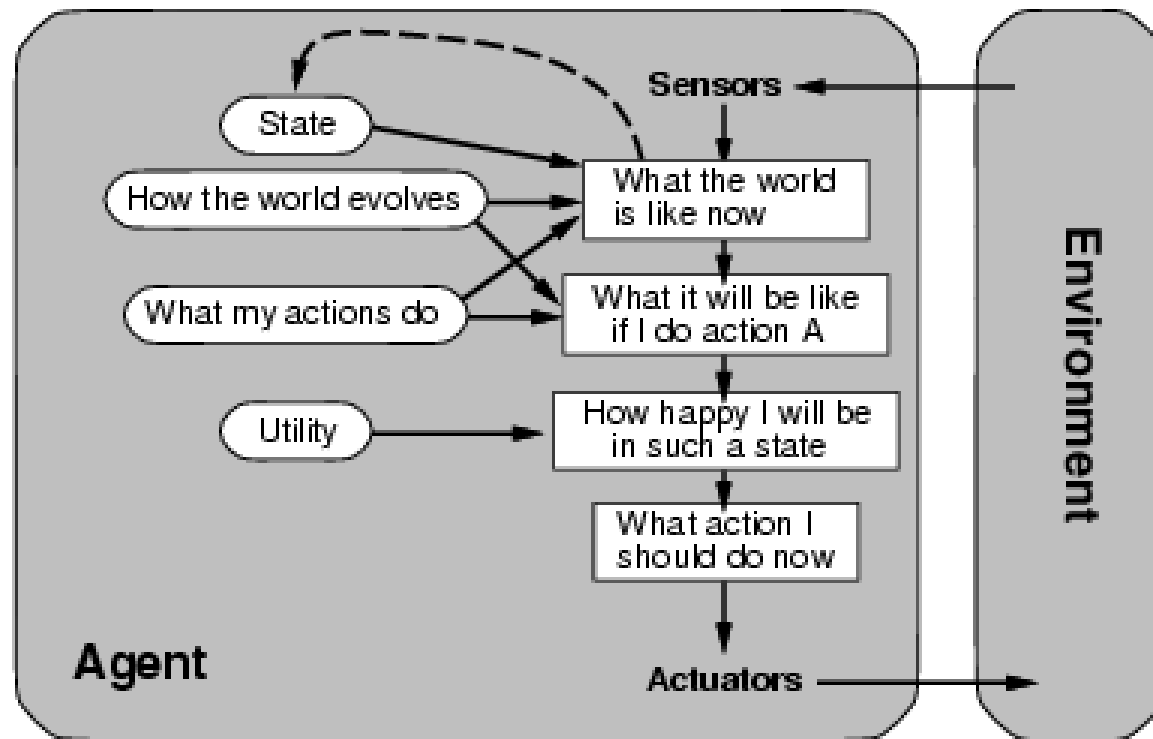
- **рефлекторни агенти, следящи света (основани на модел на света):** в паметта се поддържа информация за състоянието на света, независимо му развитие и ефектите от въздействието на агента върху него.



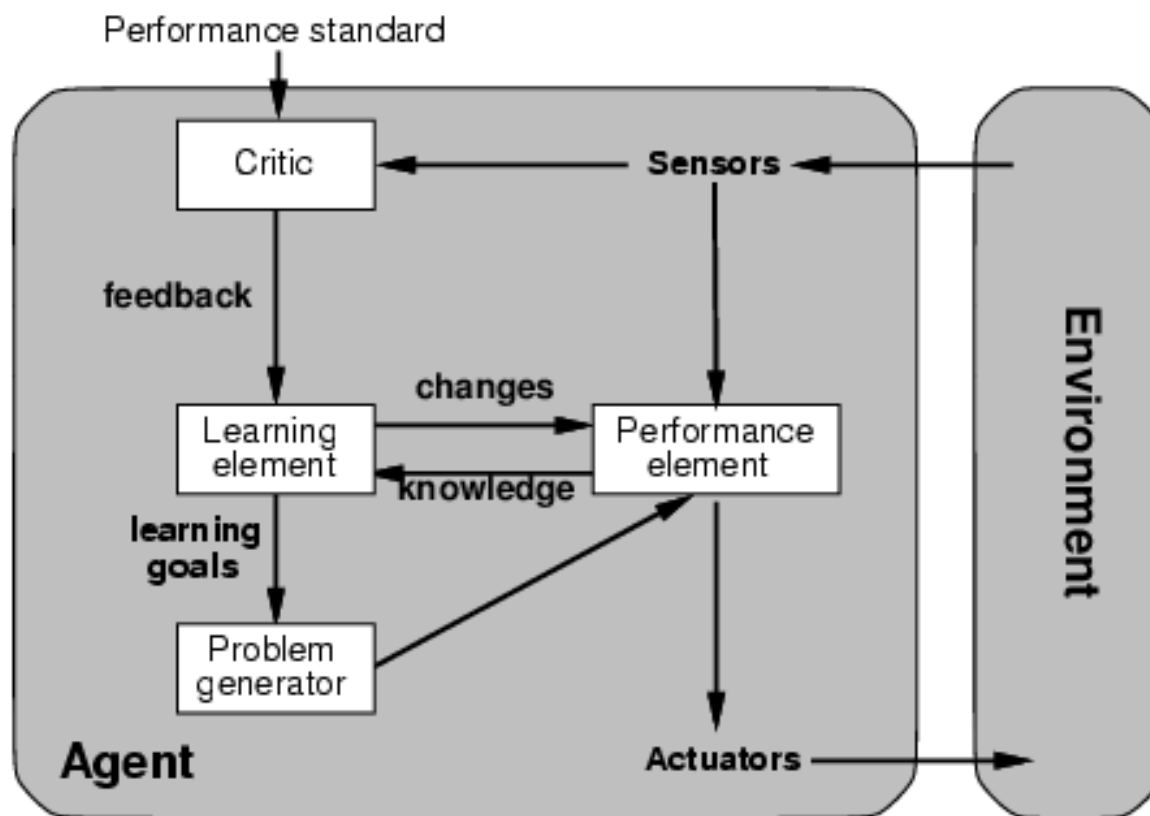
- **агенти, основани на цели:** действията се избират въз основа на това, което се знае за очакваните им резултати, и целите на агента. Методите са търсене и планиране. Поведението им е по-малко ефективно, отколкото на рефлекторните агенти, но е по-гъвкаво.



- **агенти, основани на полезност:** състоянията на света могат да се сравняват по произволно множество критерии. Полезността е функция на състоянието, отразяваща степента на задоволеност на агента. Включва разумно разрешаване на ситуации, които са проблематични за целевите агенти: *конфликтни цели*, всяка от които пречи на постигането на останалите, и *несигурни цели*, различни по важност и вероятност.



Важна характеристика на широк клас интелигентни агенти е способността им да се *самообучават*:



Среди

Характеристики на средите, в които действат агентите:

- **(не)достъпност:** агентът (не) може да възприема цялостно състоянието на средата по всяко време;
- **(не)детерминираност:** агентът (не) може точно да определи следващото състояние на средата въз основа на текущото ѝ състояние и своите действия;

- **(не)епизодичност:** историята (не) е разделена на епизоди, във всеки от които агентът получава възприятие и след това предприема действие, като ставащото във всеки епизод не зависи от станалото в предишните;
- **статичност/динамичност:** средата (не) може да се промени, докато агентът планира действията си. Ако средата не се променя с времето, но се променя оценката за ефективността на агента, средата е **полудинамична**;
- **дискретност/непрекъснатост:** има/няма ограничен брой ясно определени и разграничени възприятия и действия.

Моделиране на средата: средата може да се моделира чрез **симулатор**, който предава на агентите техните възприятия, получава от тях действията им, обновява средата и евентуално пресмята оценки за ефективността на агентите. Желателно е симулаторът да може да работи с различни среди от определена фамилия.