



هوش مصنوعی

حسین کارشناس

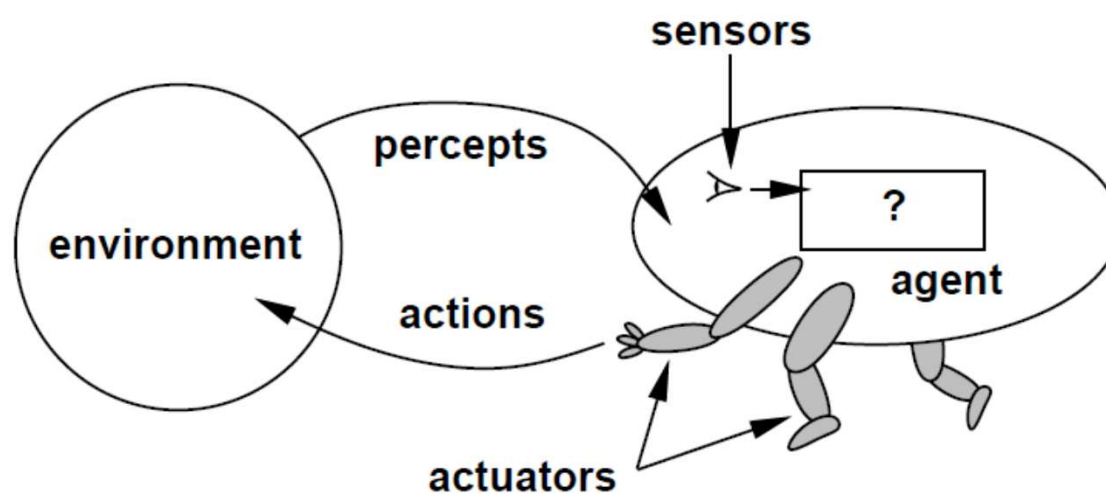
دانشکده ریاضی

ترم اول ۹۴ - ۹۳

عوامل های هوشمند (intelligent)

عامل‌ها

- مفهوم عقلانیت (rationality) در تعریف هوشمندی
- عامل‌های هوشمند
- عامل: هر چیزی که محیط خود را توسط حسگرها درک می‌کند و توسط محرک‌ها در آن کنش انجام می‌دهد



عامل‌ها

• نمونه‌ها

• انسان

• حسگرها: چشم‌ها، گوش‌ها، سایر قوای حسی

• محرک‌ها: دست‌ها، پاها، تارهای صوتی و ...

• ربات

• حسگرها: دوربین، حسگرهای مادون قرمز

• محرک‌ها: انواع موتورها

• نرم‌افزار

• حسگرها: ضربات کلید، محتویات فایل، بسته‌های شبکه

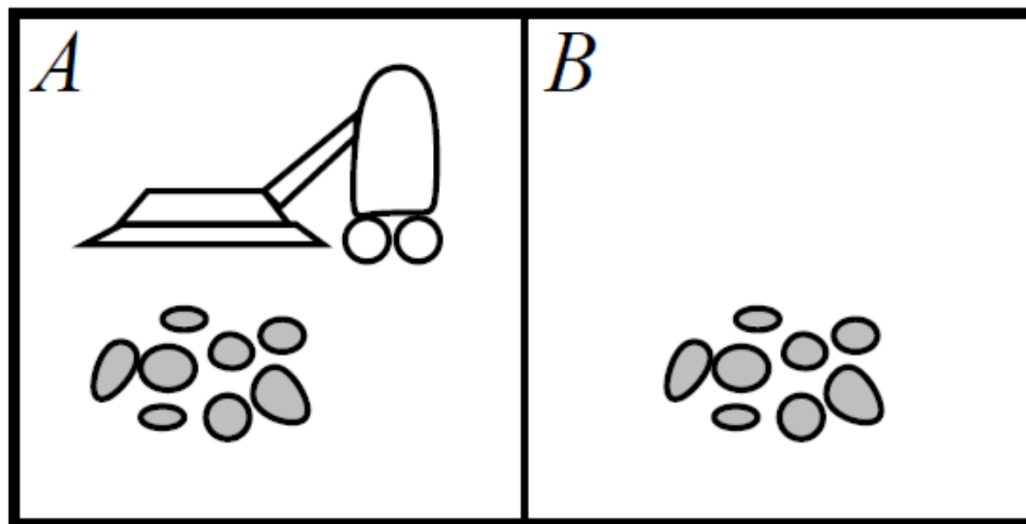
• محرک‌ها: نمایاندن در صفحه نمایش، نوشتن فایل‌ها و فرستادن بسته‌های شبکه

عامل‌ها – مفاهیم

- دنباله‌ی ادراکی (percept sequence)
- کنش‌ها فقط بر اساس دنباله ادراکی ملاحظه شده تا لحظه کنونی و نه مشاهدات آینده
- تابع عامل (agent function): هر دنباله ادراکی را به یک کنش می‌نگارد
$$f: P^* \rightarrow A$$
- نمایش جدولی (tabulated): یک نمایش بیرونی از نحوه عملکرد عامل
- توصیف انتزاعی
- برنامه عامل (agent program): یک پیاده‌سازی واقعی از عملکرد عامل
که در یک سیستم فیزیکی اجرا می‌شود

عاملها

• مثال: دنیای جاروبرقی



• دریافتها: [مکان، وجود آلودگی]

• کنشها: {حرکت به چپ، حرکت به راست، مکش}

عاملها

- دنیای جاروبرقی – نمایش جدولی تابع عامل

Percept sequence	Action
<i>[A, Clean]</i>	<i>Right</i>
<i>[A, Dirty]</i>	<i>Suck</i>
<i>[B, Clean]</i>	<i>Left</i>
<i>[B, Dirty]</i>	<i>Suck</i>
<i>[A, Clean], [A, Clean]</i>	<i>Right</i>
<i>[A, Clean], [A, Dirty]</i>	<i>Suck</i>
<i>⋮</i>	<i>⋮</i>
<i>[A, Clean], [A, Clean], [A, Clean]</i>	<i>Right</i>
<i>[A, Clean], [A, Clean], [A, Dirty]</i>	<i>Suck</i>
<i>⋮</i>	<i>⋮</i>

عامل‌ها

- دنیای جاروبرقی – یک نمونه برنامه عامل

```
function REFLEX-VACUUM-AGENT( [location,status] ) returns an action
```

```
  if status = Dirty then return Suck
```

```
  else if location = A then return Right
```

```
  else if location = B then return Left
```


عوامل‌های معقول

- عاملی معقول است که **کار درست** را انجام دهد
- انجام دنباله‌ای از کنش‌ها که به نتیجه مطلوب (desirable) منجر شود
- نیاز به یک معیار کارایی (performance measure)
- به ارزیابی یک دنباله از حالت‌های محیط می‌پردازد
- نگاه به هدف در مقابل نگاه به مسیر دستیابی به هدف
- **تعریف:** عاملی معقول است که برای هر **دنباله‌ی ادراکی** ممکن **کنشی** را انتخاب کند که انتظار می‌رود منجر به **میزان کارایی** بیشینه بر اساس شواهد بدست آمده از دنباله‌ی ادراکی و هرگونه **دانش درونی** عامل شود.

عوامل‌های معقول

- مثال در دنیای جاروبرقی
 - دریافت‌ها: [مکان، وجود آلودگی]
 - کنش‌ها: {حرکت به چپ، حرکت به راست، مکش}
 - معیار کارایی: یک امتیاز به ازای هر مربع تمیز در یک واحد زمانی
 - دانش درونی یا پیش‌فرض: ساختار محیط (شامل دو مربع)، مکش مربع را تمیز می‌کند، مربع‌ها پس از تمیز کردن کثیف نمی‌شوند، اگر عامل در مربع راست (چپ) باشد حرکت به چپ (راست) عامل را به مربع چپ (راست) برده و در غیر اینصورت حرکت نمی‌کند
- تابع عامل معرفی شده یک عامل معقول را نشان می‌دهد

عوامل‌های معقول

• اشتباهات

• عامل معقول \neq دانای کامل (omniscient)

• تأثیر کنش‌ها همیشه مطابق انتظار نیست

• تفاوت کارایی مورد انتظار و کارایی واقعی

• عقلانیت \neq سهل‌انگاری

• گردآوری اطلاعات و اکتشاف

• یادگیری یا دانش پیش فرض

• خودمختاری

• قابلیت وفق‌پذیری

• به سمت استقلال از دانش پیش فرض



محیط کاری (Task Environment)

محیط کاری

- پیش نیاز طراحی عامل های معقول: شناسایی مسأله
- توصیف PEAS از یک مسأله
- کارایی (Performance)
- محیط (Environment)
- محرک ها (Actuators)
- حسگرها (Sensors)
- مثال: راننده خودکار تاکسی

Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi driver	Safe, fast, legal, comfortable trip, maximize profits	Roads, other traffic, pedestrians, customers	Steering, accelerator, brake, signal, horn, display	Cameras, sonar, speedometer, GPS, odometer, accelerometer, engine sensors, keyboard

محیط کاری - ویژگی‌ها

- دسته‌بندی محیط‌ها (مسائل) بر اساس ویژگی‌ها
- مشاهده‌پذیری کامل یا جزئی
- آیا حسگرها تمامی جنبه‌های محیط که مرتبط با عملکرد عامل هستند را در هر لحظه از زمان رصد می‌کنند؟
- مشاهده‌پذیری کامل ← عدم نیاز به **حالت درونی** (حافظه)
- تک‌عاملی یا چندعاملی
- آیا عوامل دیگری نیز در محیط فعالیت دارند؟
- نحوه شناسایی عوامل دیگر: موجودیت‌هایی که رفتار عامل در عملکرد و معیار کارایی آنها تأثیر گذار است

محیط کاری - ویژگی‌ها

- تک‌عاملی یا چندعاملی - ادامه
- محیط‌های رقابتی (competitive) یا تعاونی (cooperative)
- نیاز به برقراری ارتباط به عنوان یک عملکرد عقلانی
- قطعی یا تصادفی
- آیا حالت بعدی محیط را می‌توان تنها از روی حالت فعلی و عملکرد عامل بدست آورد؟
- عدم قطعیت (uncertainty)
- احتمالات
- قسمت به قسمت (episodic) یا متوالی
- آیا حالات محیط قابل تقسیم به واحدهای مستقل از هم در بستر زمان است؟
- تأثیرات بلندمدت برای عملکردهای کنونی در محیط‌های متوالی

محیط کاری - ویژگی‌ها

- ایستا یا پویا
- آیا محیط در فاصله بین ادراک و کنش (تصمیم‌گیری) عامل می‌تواند تغییر کند؟
- محیط‌های نیمه‌پویا: اگر گذشت زمان فقط در معیار کارایی محیط تأثیر گذار باشد (مثال: بازی شطرنج با ساعت)
- گسسته یا پیوسته
- حالت‌ها، زمان، ادراک‌ها و کنش‌ها
- شناخته شده یا ناشناخته
- بیشتر به دانش عامل یا طراح از محیط برمی‌گردد
- شناخته یا ناشناخته بودن با مشاهده‌پذیری تفاوت دارد
- نیاز به یادگیری برای تصمیم‌گیری درست

محیط کاری - ویژگی‌ها

• مثال

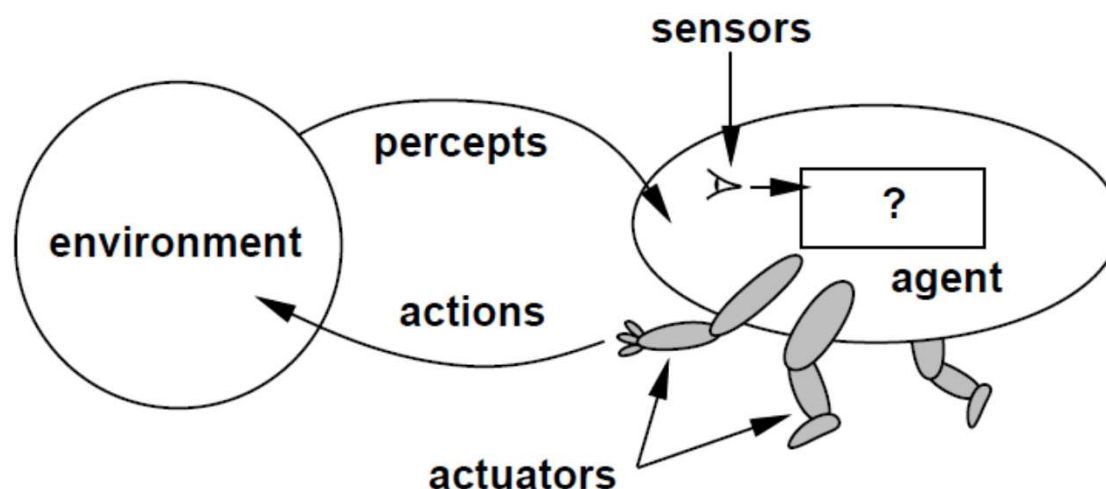
Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword puzzle	Fully	Single	Deterministic	Sequential	Static	Discrete
Chess with a clock	Fully	Multi	Deterministic	Sequential	Semi	Discrete
Poker	Partially	Multi	Stochastic	Sequential	Static	Discrete
Backgammon	Fully	Multi	Stochastic	Sequential	Static	Discrete
Taxi driving	Partially	Multi	Stochastic	Sequential	Dynamic	Continuous
Medical diagnosis	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Image analysis	Fully	Single	Deterministic	Episodic	Semi	Continuous
Part-picking robot	Partially	Single	Stochastic	Episodic	Dynamic	Continuous
Refinery controller	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Interactive English tutor	Partially	Multi	Stochastic	Sequential	Dynamic	Discrete

• وجود پیاده‌سازی برخی محیط‌ها در وبسایت کتاب

ساختار عامل‌ها

ساختار عامل‌ها

- حرکت از نگاه بیرونی (رفتار) به نگاه درونی (نحوه پیاده‌سازی عملکرد)
- عامل = برنامه + معماری
- معماری شامل حسگرها و محرک‌ها و در واقع ارتباط دهنده برنامه با محیط است
- شالوده (skeleton) در نظر گرفته شده



ساختار عامل‌ها

- یک طراحی ساده: مبتنی بر تابع عملکرد عامل

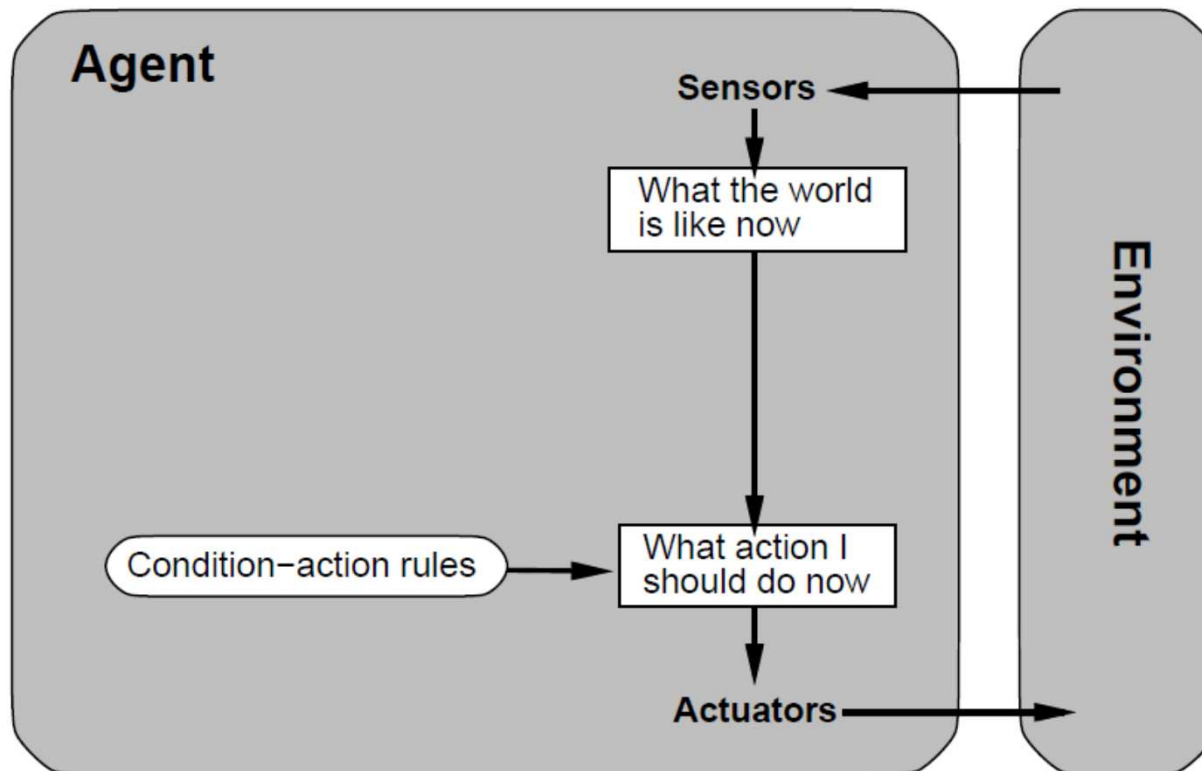
```
function TABLE-DRIVEN-AGENT(percept) returns an action
  persistent: percepts, a sequence, initially empty
               table, a table of actions, indexed by percept sequences, initially fully specified

  append percept to the end of percepts
  action ← LOOKUP(percepts, table)
  return action
```

- پیاده‌سازی ساده
- نیاز به داشتن جدول کامل عملکرد عامل برای هر دنباله ادراکی
- $\sum_{t=1}^T |Q|^t$ حالت ممکن
- برای دوره زمانی T وقتی Q ادراک‌های ممکن را نشان دهد

انواع عامل‌ها

- عامل‌های واکنشی ساده (simple reflex)
- تصمیم‌گیری فقط بر اساس ادراک فعلی
- استفاده از قوانین اگر – آنگاه (شرط – کنش)



انواع عامل‌ها

- عامل‌های واکنشی ساده

```
function SIMPLE-REFLEX-AGENT(percept) returns an action
  persistent: rules, a set of condition–action rules

  state ← INTERPRET-INPUT(percept)
  rule ← RULE-MATCH(state, rules)
  action ← rule.ACTION
  return action
```

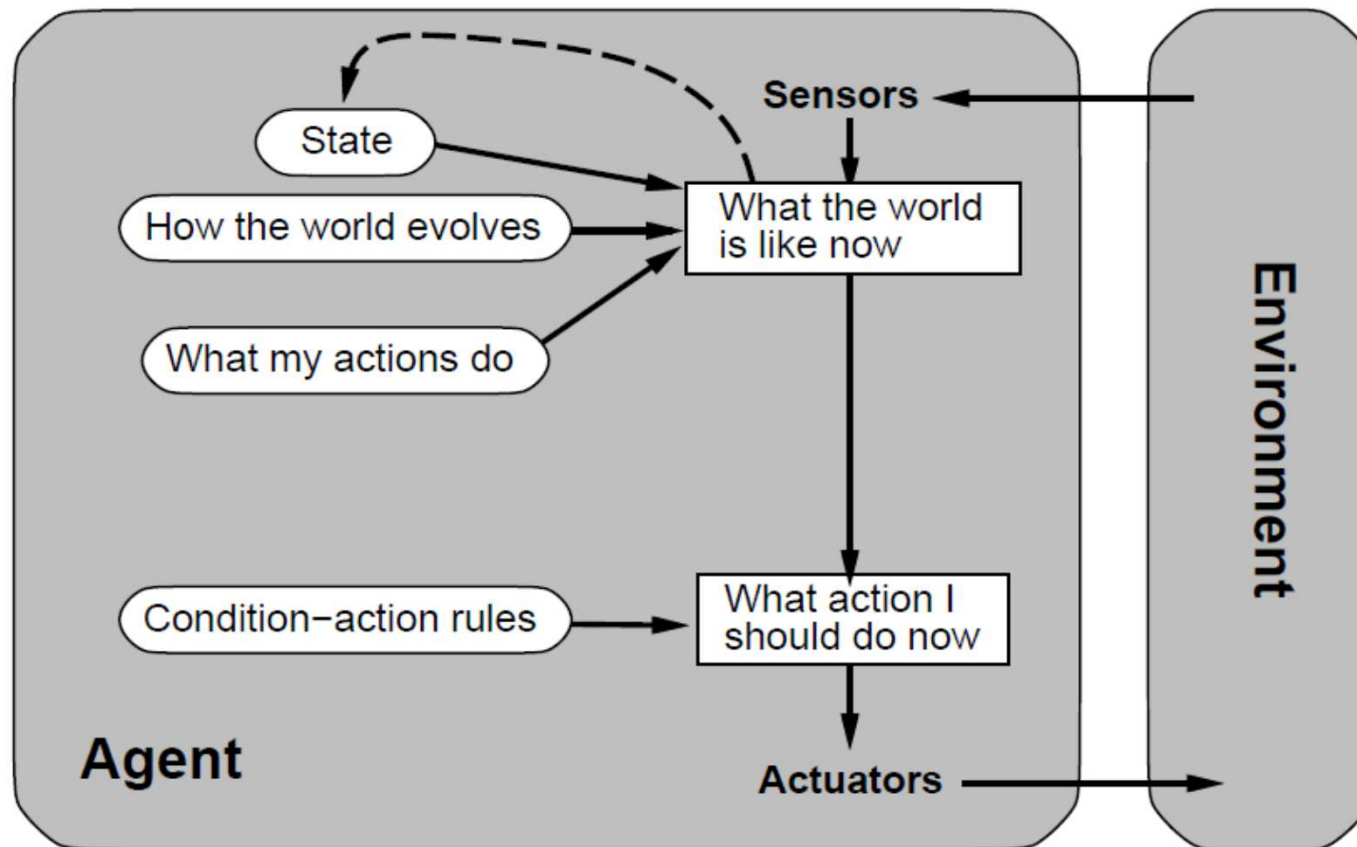
- ممکن است با عدم امکان تصمیم‌گیری مواجه شوند
- مناسب برای محیط‌های کاملاً مشاهده‌پذیر
- استفاده از شانس می‌تواند در برخی موارد راه‌گشا باشد

انواع عامل‌ها

- عامل‌های واکنشی مبتنی بر مدل (model-based reflex)
- استفاده از یک حالت درونی: به خاطر سپردن بخشی از محیط که در حال حاضر توسط عامل دیده نمی‌شود
- مدل: دانش عامل در مورد محیط
- دو نوع دانش درونی
 - نحوه تغییر محیط مستقل از عامل
 - قوانین ناظر به محیط
- نحوه تأثیر کنش‌های عامل بر محیط

انواع عامل‌ها

- عامل‌های واکنشی مبتنی بر مدل (model-based reflex)



انواع عامل‌ها

- عامل‌های مبتنی بر مدل

```
function MODEL-BASED-REFLEX-AGENT(percept) returns an action
  persistent: state, the agent's current conception of the world state
               model, a description of how the next state depends on current state and action
               rules, a set of condition–action rules
               action, the most recent action, initially none

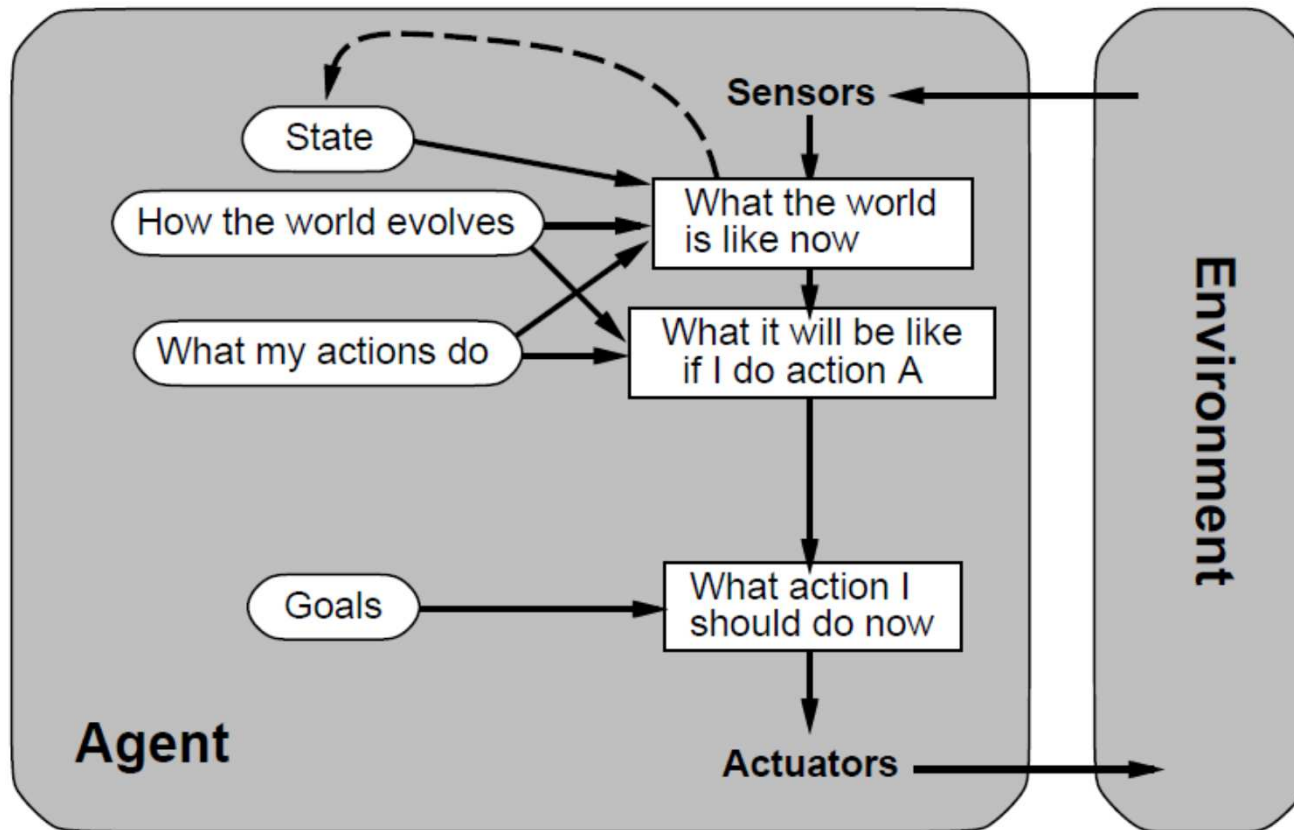
  state ← UPDATE-STATE(state, action, percept, model)
  rule ← RULE-MATCH(state, rules)
  action ← rule.ACTION
  return action
```

انواع عامل‌ها

- عامل‌های مبتنی بر هدف (goal-based)
- اطلاعات در مورد حالت فعلی برای تصمیم‌گیری کافی نیست
- نیاز به اطلاعاتی در مورد حالات مطلوب (هدف)
- نحوه رسیدن به حالت مطلوب از حالت فعلی همیشه مشخص نیست
- استفاده از جستجو و برنامه‌ریزی در انتخاب کنش‌ها
 - در نظر گرفتن آینده
 - تبعات کنش‌ها در حالات مختلف
- انعطاف‌پذیری بیشتر نسبت به عامل‌های واکنشی
 - تطابق با محیط‌های مختلف
 - سرعت عمل پایین‌تر

انواع عامل‌ها

- عامل‌های مبتنی بر هدف (goal-based)

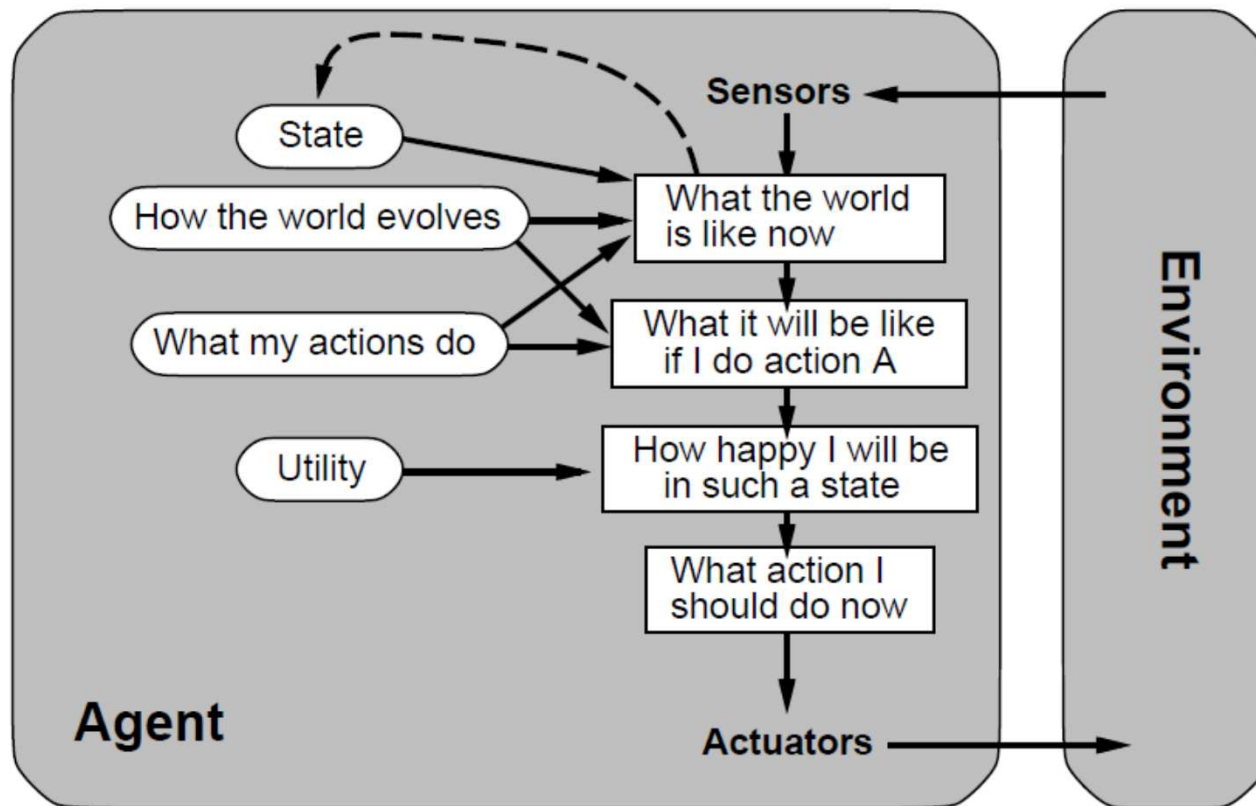


انواع عامل‌ها

- عامل‌های مبتنی بر سودمندی (utility-based)
- استفاده از میزان سودمندی حالت‌های مختلف بجای هدف بودن یا نبودن آنها
 - وجود چندین هدف غیرقطعی
 - هدف‌های متناقض
- تابع سودمندی یک بومی‌سازی (internalization) از معیار کارایی در عامل
 - منظور سودمندی مورد انتظار است
- سادگی بیشتر در طراحی عامل‌های معقول
 - در محیط‌های غیرقطعی
 - نزدیکی بیشتر به تعریف عامل معقول

انواع عامل‌ها

- عامل‌های مبتنی بر سودمندی (utility-based)

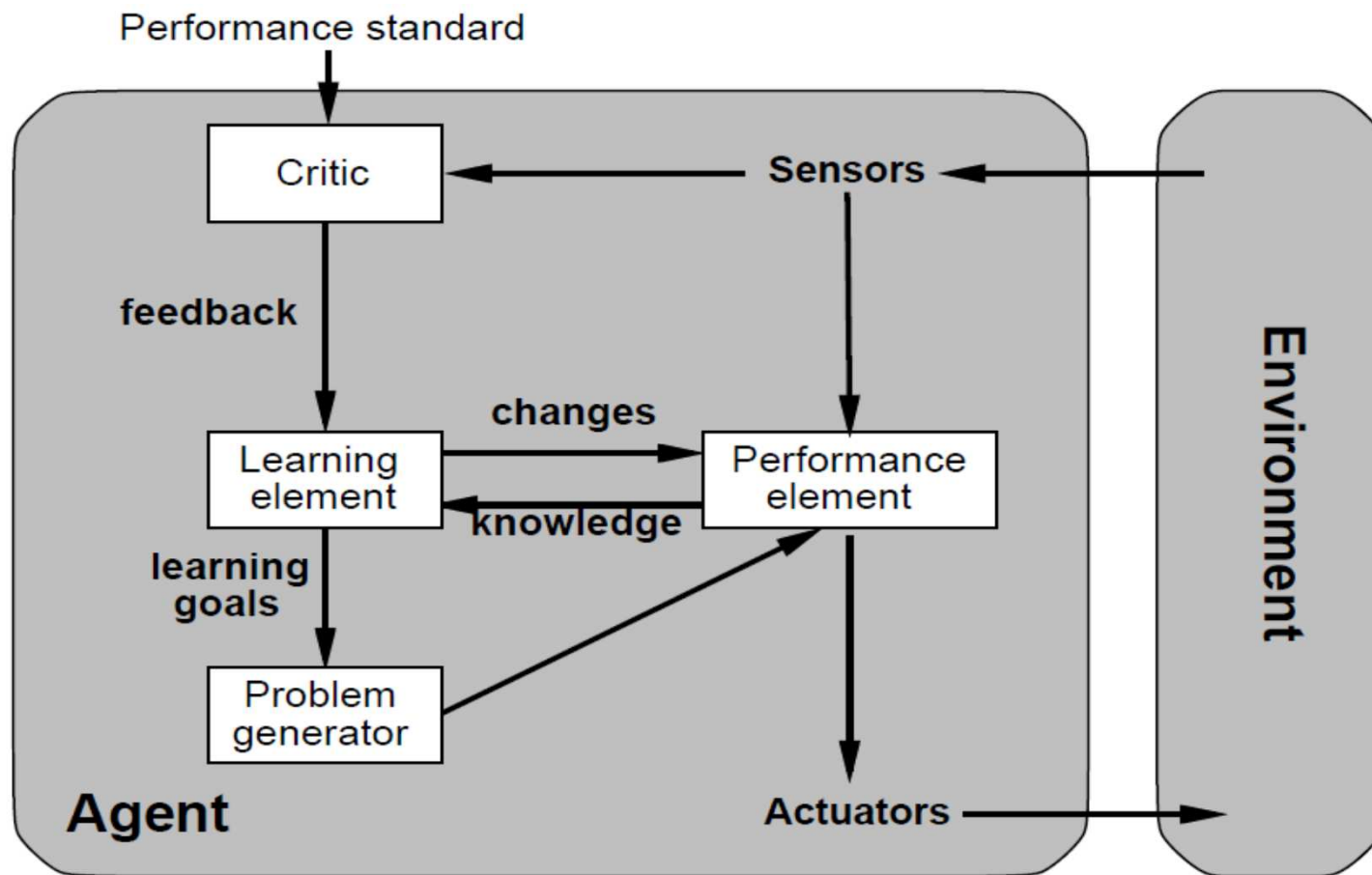


انواع عامل‌ها

- عامل‌های یادگیر
- ایده‌های اولیه از Alan Turing
- یادگیری: فرآیند تغییر مؤلفه‌های عامل برای وفق دادن آنها بر حسب اطلاعات بازخوردی
 - طراحی سریعتر
 - رقابت‌پذیری بیشتر عامل‌ها در محیط‌های ناشناخته
- چهار مؤلفه اصلی
 - واحد کارایی: مسئول انتخاب کنش‌ها – مکانیسم داخلی عامل
 - واحد یادگیری: مسئول بهسازی
 - نظارت: بررسی عملکرد عامل از دید بیرونی
 - مولد مسأله: مسئول پیشنهاد کنش‌های جدید (مثال: جستجو)

انواع عامل‌ها

- عامل‌های یادگیر



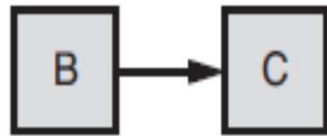
نمایش دانش در عامل‌ها

• سطوح مختلف پیچیدگی نمایش دانش

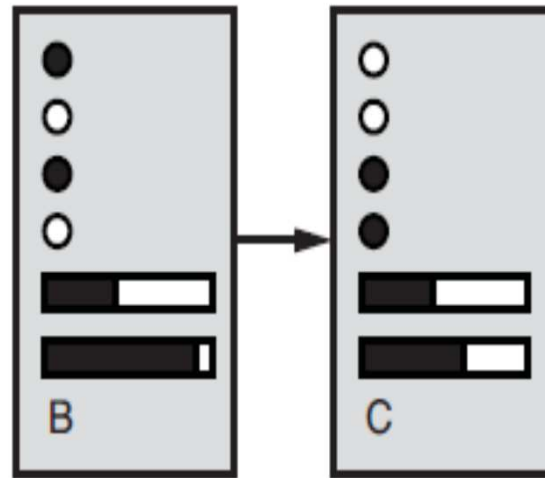
- (1) نمایش تجزیه ناپذیر: دانش حالت‌های محیط به شکل واحدهای ریز
 - (2) نمایش فاکتوربندی شده: دانش حالت‌های محیط به شکل مجموعه‌ای از صفات و متغیرهای دارای مقدار
 - امکان وجود صفات مشترک در حالت‌های متفاوت
 - امکان نمایش عدم قطعیت
 - (3) نمایش ساختار یافته: بیان ارتباطات بین اشیاء در دانش حالت‌های محیط
- افزایش توان بیان در زبان‌های رساتر (expressive)
 - امکان بیان موجزتر
 - یادگیری و استدلال پیچیده‌تر

نمایش دانش در عامل‌ها

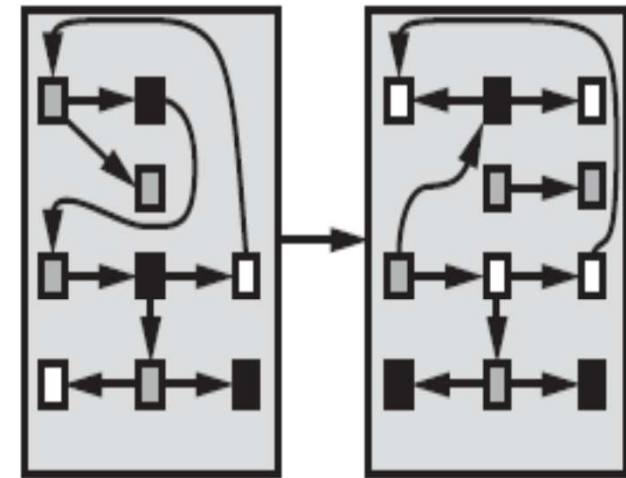
• سطوح مختلف نمایش دانش



(a) Atomic



(b) Factored



(b) Structured

جمع‌بندی

- عامل‌ها
- تعاریف مربوطه: دنباله ادراکی، تابع عملکرد، برنامه عامل
- تعریف رسمی عامل معقول
- توصیف PEAS برای محیط کاری
- ویژگی‌های محیط کاری
- طراحی‌های پایه برای عامل‌ها: واکنشی ساده، واکنشی مبتنی بر مدل، مبتنی بر هدف، مبتنی بر سودمندی، عامل‌های یادگیر
- سطوح نمایش دانش در عامل‌ها