*//Nga dịch*

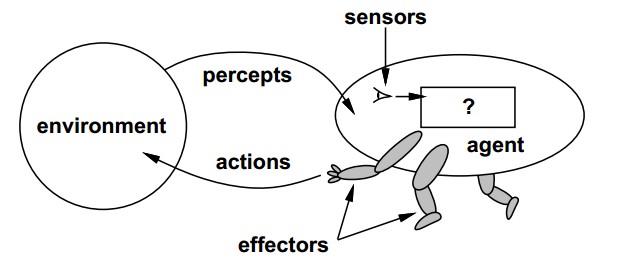
**2.1 Giới thiệu**

Một agent là bất cứ điều gì có thể được cảm nhận thông qua effectors. Một agent của con người luôn có mắt, tai, và các cơ quan khác để cảm nhận, và bàn tay, chân, miệng, và các bộ phận cơ thể khác cho những hành động. Một agent tự động có thể là camera dò tìm hồng ngoại cho các cảm biến và động cơ khác nhau để tạo hành động tương ứng. Một phần mềm Agent mã hóa chuỗi bit và hành động của mình. Mô hình tổng quát của Agent được biểu diễn trong hình 2.1

Mục tiêu của chúng tôi trong cuốn sách này là để thiết kế các tác nhân làm tốt công việc tác động lên môi trường của họ. Trước tiên chúng tôi sẽ nói một chút về ý nghĩa của một công việc tốt. Sau đó, chúng tôi sẽ nói về những thiết kế khác nhau cho những agent thành công ở những câu hỏi trong hình 2.1. Chúng tôi thảo luận về một số trong những nguyên tắc chung được sử dụng trong các thiết kế của các agent trong suốt cuốn sách, những nguyên tắc chủ yếu cần biết trong agent. Cuối cùng, chúng tôi sẽ trình bày cách cài đặt agent trên 1 môi trường và mô tả ý nghĩa ở những môi trường khác

**2.2 Một agent hoạt động thế nào**

Một agent phù hợp luôn làm đúng. Rõ ràng điều này luôn tốt hơn khi làm sai, nhưng điều đó có nghĩa là gì? Khi một xấp xỉ đầu tiên, chỉ có một hành động đó là duy nhất chính là nguyên nhân agent hoạt động chính xác.



Hình 2.1

Chúng tôi sử dụng cách nào để tối ưu hiệu năng - Tiêu chuẩn nào để xác định một agent hoạt động đúng. Rõ ràng, không có một biện pháp cố định thích hợp (cho) mọi agent.

Chúng tôi đã có thể hỏi agent về một quan điểm chủ quan rằng thế nào là hạnh phúc với những gì họ có nhưng một số agent không có câu trả lời hoặc là họ tự đánh lừa bản thân (những agent con người nói riêng – họ không muốn mọi thứ sau khi thất bại). Do đó, chúng tôi sẽ nhấn mạnh vào một biện pháp thực hiện được áp đặt bởi một số chính sách. Nói cách khác, chúng tôi thiết lập một tiêu chuẩn: thế nào là thành công trong một môi trường và sử dụng nó để đo lường hiệu quả hoạt động của các agent.

Ví dụ một agent có chức năng hút bụi thì hiệu năng của nó là chi phí để làm sạch bụi bẩn trong vòng 8h. Nếu sử dụng hiệu năng cao hơn thì sẽ ảnh hưởng tới số lượng điện tiêu thụ và tiếng ồn tạo ra cũng lớn.sử dụng một biện pháp có hiệu suất khác có thể cung cấp một cách hiệu quả để giải quyết vấn đề.

Đánh giá hiệu năng là một việc quan trọng. Nếu chúng ta cân nhắc có bao nhiêu lượng bụi bẩn agent sẽ làm sạch trong giờ đầu tiên của 1 ngày, chúng ta sẽ khen thưởng cho những ai làm nhanh (ngay cả khi làm ít việc hoặc là sau này không làm việc) và trừng phạt những người làm việc chậm trễ. Do đó nếu muốn đánh giá hiệu suất làm việc về lâu dài thì cần đánh giá hiệu năng của họ trong 8 tiếng/ngày hoặc là trong suốt cuộc đời.

Chúng ta cần cẩn thận để phân biệt giữa hợp lý và sự thông suốt. Một agent thông suốt luôn biết được kết quả của hành động của mình và hành động cho phù hợp. Nhưng một agent thông suốt là điều không tưởng trong thực tế. Hãy xem xét ví dụ sau đây:

Khi tôi đi bộ trên đường Champs Elyse vào một ngày và tôi nhìn thấy một người bạn cũ trên đường phố, lúc đó giao thông trên đường phố thông suốt nên tôi đã sang đường. Khi đó tại địa điểm cách 33.000 feet, một cánh cửa của 1 chiếc máy bay rơi xuống, như thế tôi có nên sang đường? Tôi không thích được nghe người ta nói “thật ngốc nghếch khi cố gắng vượt qua đường”. Thay vào đó, điều này chỉ ra rằng tính hợp lý là có liên quan với sự thành công mong đợi cho những gì đã được nhận thức. Băng qua đường là hợp lý vì hầu hết sẽ thành công, và không có cách nào tôi có thể đoán trước cửa rơi xuống. Lưu ý rằng một agent được trang bị radar phát hiện cửa rơi hoặc một cái lồng thép đủ mạnh để kéo chúng lại nhưng không có điều gì là hợp lý cả.

Thay vào đó điều này chỉ ra rằng tính hợp lý có liên quan đến những mong đợi thành công được định hình trong nhận thức. Hành động băng qua đường sẽ là hợp lý vì hầu hết sẽ không gặp vấn đề trong khoảng thời gian đó và cũng không thể đoán trước cánh cửa có rơi xuống không. Một agent nào khác được trang bị radar phát hiện cửa rơi hoặc có một cái lồng thép lớn kéo cái cửa đi, tuy nhiên sẽ không thể có những chuyện như vậy.

Nói cách khác chúng ta không thể đổ lỗi cho những agent gây ra vấn đề gì đó mà nó không hiểu hoặc những hành động thất bại khi nó không đủ khả năng thực hiện. Nhưng khi hạ thấp các yêu cầu không chỉ là một câu hỏi thẳng thắn tới agent. Vấn đề là chúng ta xác định một agent thông minh nên làm những điều thực sự là đúng , không thể thiết kế được một agent để thực hiện được điều này trừ khi nâng cao hiệu suất. Tóm lại những gì là hợp lý ở trong mọi thời điểm phụ thuộc vào 4 điều:

Tính toán hiệu suất cần bỏ ra để đạt được thành công

Tất cả mọi thứ mà agent đang có cho đến thời điểm này. Chúng tôi gọi đó là những trải nghiệm cảm giác trong quá khứ.

Agent biết những gì về môi trường mà nó sẽ tương tác

Hành động agent cần áp dụng

Điều này dẫn đến một định nghĩa về những agent lý tưởng:

Trong mỗi trường hợp khác nhau, agent lý tưởng làm bất kì hành động nào phát huy tối đa hiệu năng của nó trên cơ sở những trải nghiệm trong quá khứ và bất cứ những gì được hình thành trong cơ sở tri thức của nó.

Chúng ta cần xem xét cẩn thận định nghĩa này. Thoạt nhìn, nó xuất hiện cho phép agent làm những điều thật thông minh. Ví dụ agent chưa tìm ra cách nào đó trước khi băng qua một con đường lưu lượng giao thông đang quá đông rồi những trải nghiệm trong quá khứ không nói cho nó biết rằng đang có một xe tải lớn đang lao đến với tốc độ cao. Định nghĩa này dường như là đúng khi chọn hành động băng qua đường. Trong thực tế sẽ có 2 điều sai: Đầu tiên là không hợp lý khi qua đường – nguy cơ gặp tai nạn khi không quan sát là quá lớn. Thứ hai một agent thông minh sẽ suy nghĩ tìm giải pháp trước khi bước qua đường bởi khi đó sẽ vận dụng hiệu suất tối đa của agent đã được tích lũy trong quá khứ trước khi hành động. Thu thập thông tin hữu ích liên quan đến vấn đề trước khi hành động là yếu tố quan trọng và sẽ được trình bày chi tiết trong chương 16

Những quan điểm, suy nghĩ của agent được hiểu là công cụ phân tích hệ thống, không phải là đặc tính tuyệt đối chia thế giới thành hai phần là agent và không phải agent. Hãy xem xét một chiếc đồng hồ, nó là một đối tượng vô tri vô giác hoặc có thể được sử dụng như một agent đơn giản. Nó giống với agent, nó thường có những hành động như dịch chuyển kim hoặc là hiển thị số. Nó là một loại agent cấp thấp không có lịch sử trải nghiệm, đồng hồ không chịu tác động của bên ngoài nên nó sẽ không bị ảnh hưởng

Vâng. Điều này không hoàn toàn đúng . Nếu người ta mang nó từ California tới Australia, nó sẽ chạy đúng khi quay nó lại 6 giờ. Chúng tôi sẽ không phá hỏng nó bởi vì chúng tôi hiểu rõ rằng nó hoạt động đúng.

//Len dịch

Ánh xạ lý tưởng từ nhận thức sang hành động.

Vì hành vi của đối tượng chỉ phụ thuộc vào chuỗi nhận thức của nó, chúng ta có thể mô tả bất cứ đối tượng cụ thể nào bằng cách tạo ra một bảng các hành động mà đối tượng thực hiện tương ứng với từng chuỗi nhận thức cụ thể (với phần lớn các đối tượng, các chuỗi này thường dài - thậm chí vô hạn, nếu không đặt trước giới hạn). Bảng này được gọi là bảng ánh xạ từ các chuỗi nhận thức đến các hành động. Về nguyên tắc, chúng ta có thể tìm ra các ánh xạ một cách chuẩn xác bằng cách thử tất cả các trường hợp có thể xảy ra, và ghi lại các hành động đáp ứng tương ứng. Nếu đối tượng sử dụng một số tính toán ngẫu nhiên, chúng ta có thể thử một số lần để tính toán đáp ứng trung bình. Ánh xạ lý tưởng là ánh xạ mô tả đối tượng lý tưởng. Xác định được các hành động mà đối tượng cần thực hiện để đáp ứng với các chuỗi nhận thử sẽ cho phép thiết kế một đối tượng lý tưởng.

Điều này không có nghĩa là chúng ta phải tạo ra một bảng tường minh với các hàng tương ứng với tất cả các chuỗi nhận thức. Có thể xác định mô tả của ánh xạ mà không cần dùng phương pháp vét cạn. Ví dụ một đối tượng đơn giản: hàm tính căn bậc 2 trong máy tính. Chuỗi nhận thức cho đối tượng này là các sự kiện ấn phím để nhập số, và hành động tương ứng là hiển thị các số trên màn hình. Ánh xạ lý tưởng trong trường hợp này là khi nhận thức là 1 số dương x, thì hành động là hiển thị số dương z sao cho z^2=x với độ chính xác, 15 số sau thập phân. Mô tả lý tưởng này không yêu cầu người thiết kế phải xây dựng bảng căn bậc 2, và cũng bản thân hàm căn bậc 2 cũng không sử dụng bảng này để thực hiện. Hình 2.2 thể hiện ánh xạ lý tưởng, và một chương trình đơn giản thực hiện ánh xạ này, dùng phương pháp Newwton.

Ví dụ về căn bậc 2 minh hoạ mối quan hệ giữa ánh xạ lý tưởng và một đối tượng lý tưởng trong một tác vụ rất hạn chế. Như đã thấy, bảng thì rất lớn, nhưng chương trình lại rất ngắn. Điều này cho thấy có thể thiết kế một đối tượng gọn nhẹ để thực hiện ánh xạ lý tưởng cho các tình huống tổng quát hơn: đối tượng có thể giải quyết vô số các tác vụ trong vô số môi trường.

*//Thúy dịch*

**Sự tự trị (Autonomy)**

Có một điều nữa cần quan tâm trong việc định nghĩa một Agent lý tưởng : trong phần “Sự hiểu biết có sẵn (built - in knowledge)”. Nếu các hành vi của agent được bố trí hoàn toàn trong “Sự hiểu biết có sẵn” , ví dụ rằng nó không cần chú ý tới các nhận biết (percept) của nó thì chúng ta nói rằng các agent thiếu sự tự trị. Ví dụ , nếu nhà sản xuất đồng hồ có thể thấy trước rằng chủ nhân của chiếc đồng hồ sẽ đến Australia vào một vài ngày đẹp trời nào đó, thì một cơ chế có thể được cài đặt sẵn để điều chỉnh các kim một cách tự động 6h cho đúng thời gian. Tất nhiên điều này có thể thực hiện thành công , nhưng sự thông minh dường như thuộc về người thiết kế đồng hồ hơn là chính chiếc đồng hồ

Cách hành động của agent có thể được bố trí vào trong cả kinh nghiệm của chính nó và “sự hiểu biết có sẵn” trong khi xây dựng cấu trúc agent, agent đó hoạt động trong một môi trường cụ thể . Một hệ thống là “tự trị - autonomous” tới một mức độ nhất định đó là hệ thống mà cách hành động của nó được quyết định bởi chính kinh nghiệm của nó. Mặc dù có thể là quá khó khăn để đạt được sự tự trị hoàn toàn : khi agent có ít hoặc không có kinh nghiệm, nó phải hành động ngẫu nhiên trừ khi người thiết kế đưa ra một vài sự giúp đỡ. Vì vậy khi mà sự tiến hóacung cấp cho động vật với đầy đủ các phản xạ tự nhiên thì chúng có thể tồn tại đủ để tự học tập, nó có lý để cung cấp một agent thông minh nhân tạo với một vài sự hiểu biết ban đầu giốn như khả năng để học tập

Sự tự trị không chỉ phù hợp với trực giác của chúng ta mà nó còn là một ví dụ của thực tiễn cơ sở khoa học. Một agent hoạt động trên cơ sở của các giả định có sẵn sẽ chỉ hoạt động thành công khi hiểu rõ các giả định , và như vậy thiếu đi sự linh hoạt . Giả sử cho ví dụ con bọ hung. Sau khi tìm ra tổ và đẻ trứng, nó đem về các viên phân từ một đống phân gần đó để bịt lối vào, nếu viên phân bị rớt trên đường đi, con bọ hung sẽ tiếp tục làm và bịt lối vào tổ với những viên phân khác theo kịch bản, mà không bao giờ để ý đến những viên phân đó đã đánh rơi . Sự tiến hóa đã gắn vào giả định của con bọ hung , và khi nó bị vi phạm , kết quả của hành vi không thành công (vì đã không nhặt lại các viên phân bị rơi). Một agent thông minh thực sự phải có khả năng hành động thành công trong các môi trường rộng lớn khác nhau đã cho đủ thời gian thích nghi

*//Phương dịch*

**2.3. Kiến trúc của Agents thông minh**

Đến đây, chúng tôi đã nói về các agent bằng cách mô tả hành vi của nó - hành động (action) được thực hiện sau khi có mọi percepts sequence. Bây giờ, chúng tôi nói về cách làm việc bên trong của nó. Công việc của AI là để thiết kế các chương trình agent (agent program): một hàm thực hiện ánh xạ từ percepts của agent đến các hành động. Chúng tôi giả định rằng chương trình này sẽ chạy trên một số loại thiết bị máy tính, sẽ được gọi kiến trúc (architecture). Các kiến trúc có thể là một máy tính đơn giản, hoặc nó có thể bao gồm phần cứng chuyên dùng cho các nhiệm vụ nhất định, như xử lý hình ảnh hoặc lọc âm thanh đầu vào. Nó có thể bao gồm phần mềm mà cung cấp một mức độ cách ly giữa máy tính (raw computer) và chương trình agent, để chúng ta có thể lập trình ở một cấp độ cao hơn. Nhìn chung, kiến trúc làm cho các percept từ các cảm biến sẵn sáng đến chương trình, chạy chương trình, và đẩy những lựa chọn các hoạt động của chương trình tới những phần tử khi chúng được phát sinh. Mối quan hệ giữa các agent, kiến trúc, và chương trình có thể tóm tắt như sau:

*agent = kiến trúc + chương trình*

Cuốn sách này phần lớn bàn về thiết kế các chương trình agent, các chương 24 và 25 đề cập trực tiếp tới vấn đề kiến ​​trúc. Trước khi chúng tôi thiết kế một chương trình agent, chúng tôi cần phải có một ý tưởng tương đối tốt về khả năng của percepts và các hành động, cái đích nào hoặc hiệu suất đo được của agent giả định phải đạt được, cần môi trường như thế nào để nó có thể hoạt động. Điều này dẫn tới 1 sự đa dạng lớn. Hình 2.3 cho thấy rằng những phần tử cơ bản cho sự lựa chọn kiểu của agent.

Điều đó có thể đem đến sự ngạc nhiên cho một số độc giả khi trong danh sách chúng tôi bao gồm các kiểu agent của một số chương trình mà có vẻ hoạt động hoàn toàn trong môi trường nhân tạo, được xác định bởi nhập liệu từ bàn phím và các dữ liệu xuất ra màn hình. "Chắc chắn" người ta có thể nói, "đây không phải là môi trường thực tế, đúng không?” Trong thực tế, điều quan trọng không phải là sự khác biệt giữa môi trường “thực” và “nhân tạo”, nhưng sự phức tạp của các mối quan hệ giữa hành vi của các agent, percept sequence tạo ra bởi môi trường, và những mục đích mà agent giả định đạt tới. Một vài môi trường “thực” lại khá đơn giản. Ví dụ, một robot được thiết kế để xem xét, kiểm tra những thứ mà chúng đi qua trên một dây chuyền sản xuất có thể làm cho việc sử dụng đơn giản hơn giả thiết: ánh sáng là luôn luôn chỉ cần như vậy, thứ duy nhất trên băng truyền sẽ là những phần của một cái gì đó nhất định, và chỉ có hai hành động - chấp nhận thứ đó hoặc đánh dấu nó như là một phế phẩm.

Ngược lại, một số phần mềm agent (hoặc những phần mềm robot hoặc softbots) tồn tại phong phú, không có miền giới hạn. Hãy tưởng tượng một softbot được thiết kế để bay một chuyến bay giả lập cho 747. Mô phỏng này rất chi tiết, môi trường phức tạp, và các phần mềm agent phải chọn từ một loạt các hành động trong thời gian thực. Hoặc hãy tưởng tượng một softbot thiết kế để quét các nguồn tin tức trực tuyến và hiển thị các thứ thú vị cho các khách hàng của mình. Để làm tốt, nó sẽ cần một số khả năng xử lý ngôn ngữ tự nhiên, nó sẽ cần phải học những gì khách hàng cho là thú vị, và phải tự động thay đổi kế hoạch của mình khi cần, ví dụ, kết nối đến 1 nguồn dữ liệu bị lỗi hoặc có 1 nguồn tin mới được trực tuyến.

Một số môi trường làm mờ sự khác biệt giữa "thực" và "nhân tạo”. Trong môi trường ALIVE (Maes et al, 1994), các phần mềm agent đưa ra percepts là ảnh chụp từ camera một căn phòng nơi một người đi về được số hóa. Agent xử lý ảnh từ camera và chọn một hành động. Môi trường này cũng hiển thị hình ảnh của camera trên một màn hình lớn, ở đó con người có thể xem, và chồng vào hình ảnh đồ họa máy tính trả lại bởi các phần mềm agent. Một trong những hình ảnh là một con chó phim hoạt hình, đã được lập trình để di chuyển về phía con người (trừ khi con người chỉ sang hướng khác dể con chó đi) và bắt tay hoặc nhảy lên háo hức trong khi làm một số cử chỉ với con người.

Môi trường nhân tạo nổi tiếng nhất là Turing Test, trong đó toàn bộ các điểm thực tế và agent nhân tạo khá tương đồng, nhưng môi trường thì có đủ thách thức và rất khó cho một phần mềm agent làm tốt được như con người. Phần 2.4 mô tả chi tiết hơn những nhân tố làm cho một số môi trường đòi hỏi cao hơn.

Note: thuật ngữ

* Percepts: Nó cơ bản là cùng một khái niệm như là một cảm nhận trong [tâm lý](http://translate.googleusercontent.com/translate_c?depth=1&hl=vi&ie=UTF8&prev=_t&rurl=translate.google.com.vn&sl=en&tl=vi&u=http://en.wikipedia.org/wiki/Psychology&usg=ALkJrhgdp_Em4gYr3-3kOzFRC47t3LLdsw) , ngoại trừ việc nó đang được cảm nhận không phải do [não](http://translate.googleusercontent.com/translate_c?depth=1&hl=vi&ie=UTF8&prev=_t&rurl=translate.google.com.vn&sl=en&tl=vi&u=http://en.wikipedia.org/wiki/Brain&usg=ALkJrhi5CAZB2LBG-IK6aeTPzRnHf1mLXw) nhưng bởi các agent.

Đơn giản hơn, đó là tất cả mọi thứ mà agent thông minh nhận thức trong môi trường