

BÁO CÁO CUỐI KỲ MÔN MÔ HÌNH NGẪU NHIÊN VÀ ỨNG DỤNG

Tác giả: nguicontraiphonui

Website: <http://nguicontraiphonui.blogspot.com>

Soạn thảo văn bản L^AT_EX bởi công cụ MikTeX & TeXmaker

Ngày 30 tháng 12 năm 2017

Mục lục

1	Giới thiệu	2
2	Phương pháp	3
2.1	Quá trình quyết định Markov	3
2.1.1	Quá trình Markov	3
2.1.2	Quá trình Markov có phần thưởng	3
2.1.3	Quá trình quyết định Markov	3
2.2	Thuật toán Q-learning	3
2.3	Thuật toán Policy-learning	3
3	Ứng dụng trong thực tế	4
4	Kết quả	5
5	Tổng kết và bàn luận	6
	Tài liệu tham khảo	7

Chương 1

Giới thiệu

Trong ngành khoa học máy tính, học tăng cường (tiếng Anh: reinforcement learning) là một lĩnh vực con của học máy, nghiên cứu cách thức một tác tử (agent) trong một môi trường nên chọn thực hiện các hành động nào để cực đại hóa một khoản thưởng (reward) nào đó về lâu dài. Các thuật toán học tăng cường cố gắng tìm một chiến lược ánh xạ các trạng thái của thế giới tới các hành động mà agent nên chọn trong các trạng thái đó.

Môi trường thường được biểu diễn dưới dạng một quá trình quyết định Markov trạng thái hữu hạn (Markov decision process - MDP), và các thuật toán học tăng cường cho ngữ cảnh này có liên quan nhiều đến các kỹ thuật quy hoạch động. Các xác suất chuyển trạng thái và các xác suất thu lợi trong MDP thường là ngẫu nhiên nhưng lại tĩnh trong quá trình của bài toán (stationary over the course of the problem).

Khác với học có giám sát, trong học tăng cường không có các cặp dữ liệu vào/kết quả đúng, các hành động gần tối ưu cũng không được đánh giá đúng sai một cách tường minh. Hơn nữa, ở đây hoạt động trực tuyến (on-line performance) được quan tâm, trong đó có việc tìm kiếm một sự cân bằng giữa khám phá (lãnh thổ chưa lập bản đồ) và khai thác (tri thức hiện có). Trong học tăng cường, sự được và mất giữa khám phá và khai thác đã được nghiên cứu chủ yếu qua bài toán multi-armed bandit.

Chương 2

Phương pháp

Chương đầu có nội dung như sau

2.1 Quá trình quyết định Markov

Viết vài lời về phần đầu, bao gồm những phần nhỏ như sau

2.1.1 Quá trình Markov

viết vào đây....

2.1.2 Quá trình Markov có phần thưởng

viết đây....

2.1.3 Quá trình quyết định Markov

2.2 Thuật toán Q-learning

2.3 Thuật toán Policy-learning

Viết chút vào đây...

Ví dụ 1. *Ví dụ minh họa....*

Ví dụ 2. *Ví dụ minh họa tiếp....*

Chương 3

Ứng dụng trong thực tế

Chương 4

Kết quả

Chương 5

Tổng kết và bàn luận

Tài liệu tham khảo

- [1] Họ và tên tác giả, năm, *Tên sách* NXB.
- [2] Họ và tên tác giả, năm, *Tên sách* NXB.
- [3] Họ và tên tác giả, năm, *Tên sách* NXB.
- [4] Họ và tên tác giả, năm, *Tên sách* NXB.
- [5] Họ và tên tác giả, năm, *Tên sách* NXB.