THÔNG TIN LUẬN ÁN

Tên luận án:	Nghiên cứu giải thuật học củng cố trong điều khiển thích nghi bền vững cho hệ phi tuyến
Chuyên ngành:	Tự động hóa
Mã số chuyên ngành:	62.52.60.01
Họ và tên NCS:	Nguyễn Tấn Lũy
Người hướng dẫn khoa học:	TS. Nguyễn Thiện Thành TS. Hoàng Minh Trí
Cơ sở đào tạo:	Trường Đại học Bách Khoa, Đại học Quốc Gia, TP. Hồ Chí Minh

Tóm tắt luận án

Bài toán điều khiển tối ưu cho hệ phi tuyến bị ràng buộc trực tiếp bởi nghiệm của phương trình Hamilton-Jacobi-Bellman (HJB) và bài toán điều khiển tối ưu bền vững bị ràng buộc trực tiếp bởi nghiệm của phương trình Hamilton-Jacobi-Isaacs (HJI). Đây là các phương trình vi phân phi tuyến không có nghiệm giải tích. Từ đó, bài toán xấp xỉ nghiệm HJB và HJI off-line hoặc online được đặt ra. Học củng cố (Reinforcement Learning (RL)) bắt nguồn từ qui hoạch động (Dynamic Programming (DP)), phát triển thành qui hoạch động thích nghi (Adaptive Dynamic Programming (ADP)) trở thành một trong những phương pháp hữu hiệu dùng để xấp xỉ các nghiệm HJB và HJI. Luận án này nghiên cứu giải thuật học củng cố điều khiển thích nghi bền vững hệ phi tuyến, trong đó qui hoạch động thích nghi online (Online Adaptive Dynamic Programming (OADP)) và qui hoạch động thích nghi bền vững online (Online Robust Adaptive Dynamic Programming (ORADP)) là hai giải thuật chính được phân tích và thiết kế. Giải thuật OADP dùng để xấp xỉ nghiệm HJB cho hệ thống phi tuyến với mô hình xác định, sau đó được phát triển thành giải thuật ORADP để xấp xỉ nghiệm HJI cho hệ phi tuyến hoàn toàn không có thông tin về động học nội (internal dynamics).

Những đóng góp mới của luận án

A. Lý thuyết:

1. Cấu trúc ADP chuẩn với hai xấp xỉ hàm được sử dụng để chuyển đổi thành cấu trúc điều khiển phi tuyến với duy nhất một xấp xỉ hàm cho giải thuật OADP để tránh độ phức tạp tính toán và lãng phí tài nguyên nhằm đẩy nhanh tốc độ hội tụ.

2. Luật cập nhật mới cho tham số cho xấp xỉ hàm và các giải thuật điều khiển mới được thiết kế. Trong giải thuật OADP, các luật cập nhật tham số được đồng bộ hóa trong một bước lặp nhằm tăng tốc độ hội tụ.

3. Luật điều khiển ổn định ban đầu để khởi động giải thuật OADP là không cần thiết. Từ đó, thủ tục thiết kế trở nên linh hoạt hơn.

4. Giải thuật OADP đảm bảo rằng hàm chi phí được tối thiểu, tham số xấp xỉ hàm và luật điều khiển hội tụ về giá trị cận tối ưu trong khi toàn bộ trạng thái của hệ kín và sai số xấp xỉ bị chặn theo tiêu chuẩn UUB (Uniform Ultimate Bounded).

5. Cấu trúc ADP chuẩn với ba xấp xỉ hàm được sử dụng để chuyển đổi thành cấu trúc điều khiển với duy nhất một xấp xỉ hàm cho giải thuật ORADP để tránh độ phức tạp tính toán và lãng phí tài nguyên nhằm đẩy nhanh tốc độ hội tụ. Giải thuật áp dụng cho hệ phi tuyến không biết trước thông tin về động học nội.

6. Luật cập nhật mới cho tham số cho xấp xỉ hàm và các giải thuật điều khiển mới được thiết kế. Trong giải thuật ORADP, các luật cập nhật tham số được đồng bộ hóa trong một bước lặp nhằm tăng tốc độ hội tụ.

7. Giải thuật ORADP không sử dụng thông tin về động học nội hệ thống, từ đó thủ tục nhận dạng hệ thống được bỏ qua.

8. Luật điều khiển ổn định ban đầu để khởi động giải thuật ORADP là không cần thiết. Từ đó, thủ tục thiết kế trở nên linh hoạt hơn.

9. Giải thuật ORADP đảm bảo rằng hàm chi phí được tối thiểu, tham số xấp xỉ hàm

và luật điều khiển hội tụ về giá trị cận tối ưu trong khi toàn bộ trạng thái của hệ kín và sai số xấp xỉ bị chặn theo tiêu chuẩn UUB (Uniform Ultimate Bounded).

B. Thực tiễn:

Để kiểm tra khả năng ứng dụng của giải thuật ORADP, mô phỏng số và thực nghiệm cho robot di động dạng xe (Wheeled Mobile Robot (WMR)) được tiến hành. So sánh với các giải thuật điều khiển thích nghi khác, giải thuật ORADP điều khiển WMR có một số ưu điểm mới:

1. Việc chia tách bộ điều khiển động học (kinematic) và động lực học (dynamic) sử dụng phổ biến trong điều khiển thích nghi cho WMR là không cần thiết. Từ đó, tránh phụ thuộc vào kinh nghiệm của người thiết kế trong việc lựa chọn các tham số cho bộ điều khiển động học.

2. Không đòi hỏi nhận dạng trực tiếp hoặc gián tiếp thành phần động học không chắc chắn, không cấu trúc trong mô hình robot.

3. Hàm chỉ tiêu chất lượng có liên quan đến sai số bám cả về động học, động lực học lẫn năng lượng điều khiển được tối thiểu.

Giải thuật ORADP tiếp tục được sử dụng để thiết kế mở rộng cho bài toán điều khiển hợp tác nhiều hệ phi tuyến MIMO không sử dụng thông tin về động học nội hệ thống. Ban đầu, lý thuyết đồ thị được sử dụng để thiết lập cấu hình truyền thông phân tán cho nhiều hệ phi tuyến hợp tác. Sau đó, giải thuật ORADP được thiết kế mở rộng thành giải thuật điều khiển hợp tác thích nghi bền vững. Kết quả điều khiển đồng bộ hóa hệ thống robot bầy đàn từ mô phỏng cho thấy tính hiệu quả của giải thuật ORADP mở rộng.

Hướng phát triển

1. Mở rộng giải thuật ORADP để thiết kế giải thuật không sử dụng thông tin các ma trận ngõ vào.

2. Mở rộng giải thuật ORADP để điều khiển hệ phi tuyến tổng quát hơn, ví dụ như

hệ phi tuyến không biết trước cấu trúc.

3. Mở rộng giải thuật ORADP để phát triển lý thuyết điều khiển hệ phi tuyến hồi tiếp ngõ ra.

4. Mở rộng giải thuật ORADP trong học củng cố phân cấp, để tăng tốc độ hội tụ.

5. Thực nghiệm giải thuật ORADP trên mô hình thực hệ thống robot bầy đàn hợp tác, hoặc trên các đối tượng bầy đàn sinh học khác.

6. Nghiên cứu tích hợp công nghệ sinh học vào giải thuật ORADP để nâng cao khả năng ứng dụng.

Người hướng dẫn khoa học	**Nghiên cứu sinh**
TS. Nguyễn Thiện Thành TS. Hoàng Minh Trí	**Nguyễn Tấn Lũy**