XÂY DỰNG VÀ THỬ NGHIỆM MODULE DỊCH TÊN RIÊNG TRONG DỊCH MÁY THỐNG KÊ

Nông Trung Nghĩa

*Khóa QH-2007-I/CC, nghành Công nghệ Thông tin – Khoa học máy tính*

**Tóm tắt khóa luận tốt nghiệp:**

Dịch tự động, dịch máy, là một vấn đề khó trong xử lý ngôn ngữ tự nhiên. Cùng với lượng thông tin trao đổi giữa các nền ngôn ngữ ngày càng lớn trên thế giới hiện nay, dịch máy ngày càng tỏ ra quan trọng. Chính vì thế có rất nhiều các nhà khoa học đã và đang nghiên cứu về dịch máy. Dịch máy thống kê là một trong những phương pháp dịch máy phổ biến và đạt được nhiều thành tựu nhất, đặc biệt là dịch máy thống kê dựa trên cụm từ. Bài toán dịch tên riêng luôn là một bài toán được ưu tiên giải quyết trong dịch máy thống kê nói riêng và dịch máy nói chung. Việc nhận đúng tên riêng từ câu dịch nguồn là một vấn đề khá quan trọng, nó không những làm tăng chất lượng dịch mà còn làm tăng tính thực tế của hệ dịch máy thống kê. Trong khóa luận này, chúng tôi đề cập tới một vài phương pháp giải quyết bài toán dịch tên riêng và thử nghiệm các phương pháp đó.

*Từ khóa:* dịch tên riêng, mô hình dịch tên riêng đơn ngữ, mô hình dịch tên riêng song ngữ.