



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0053357

(51)^{2020.01} G06F 16/24

(13) B

(21) 1-2021-03527

(22) 14/06/2021

(30) 1-2020-05846 13/10/2020 VN

(45) 25/11/2025 452

(43) 27/09/2021 402A

(73) CÔNG TY CỔ PHẦN CÔNG NGHỆ OLLI (VN)

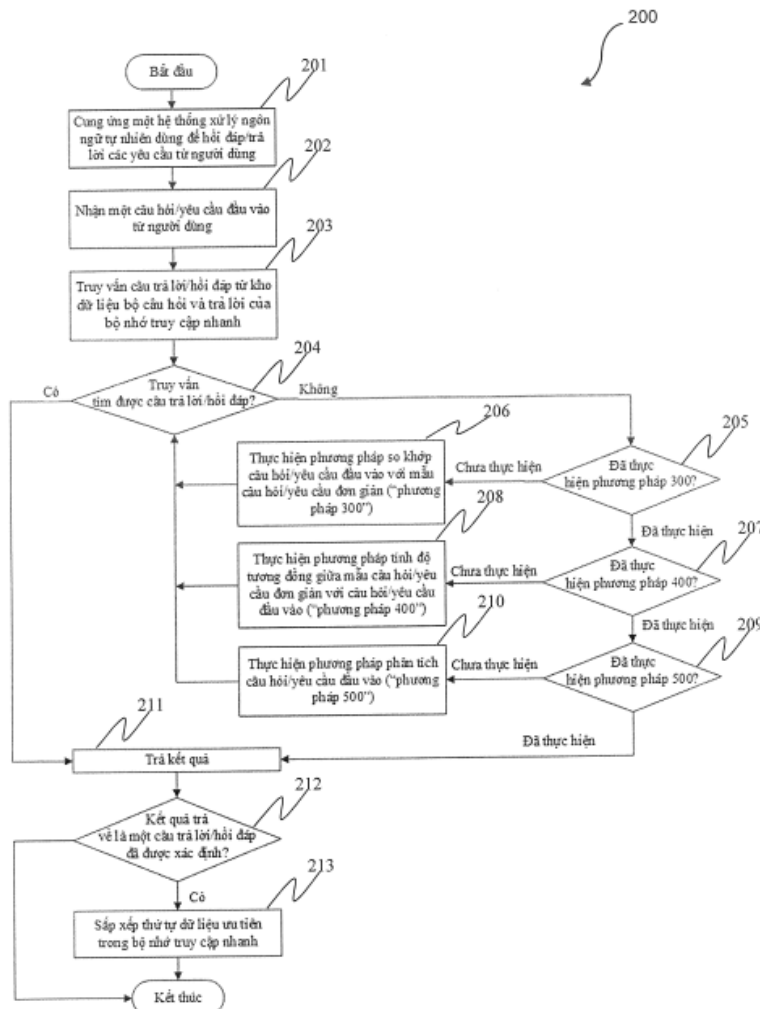
96 Nguyễn Thị Minh Khai, phường 6, quận 3, thành phố Hồ Chí Minh

(72) Bùi Bách Việt (VN); Tạ Thanh Hải (VN); Đỗ Đức Hào (VN).

(54) HỆ THỐNG VÀ PHƯƠNG PHÁP XỬ LÝ NGÔN NGỮ TỰ NHIÊN DÙNG ĐỂ
HỒI ĐÁP CÁC YÊU CẦU TỪ NGƯỜI DÙNG

(21) 1-2021-03527

(57) Một hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp các yêu cầu từ người dùng bao gồm: ít nhất một thiết bị điện tử thông minh, một hệ thống thiết bị điện toán, tất cả được kết nối vào mạng dựa trên đám mây thông qua kênh truyền dẫn internet. Thiết bị điện tử thông minh dùng để nhận câu yêu cầu đầu vào là các lệnh/các yêu cầu từ người dùng; hệ thống thiết bị điện toán dùng để hỏi đáp các câu yêu cầu đầu vào từ người dùng thông qua thiết bị điện tử thông minh. Ngoài ra, sáng chế còn cung cấp một phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp các yêu cầu từ người dùng bao gồm các bước: i) nhận một câu yêu cầu đầu vào từ người dùng thông qua thiết bị điện tử thông minh; ii) truy vấn câu hỏi từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh; iii) thực hiện phương pháp so khớp câu yêu cầu đầu vào với mẫu câu yêu cầu đơn giản; iv) thực hiện phương pháp tính độ tương đồng giữa mẫu câu yêu cầu đơn giản với câu yêu cầu đầu vào; v) thực hiện phương pháp phân tích câu yêu cầu đầu vào; vi) trả kết quả truy vấn/hỏi đáp yêu cầu đến người dùng và/hoặc thiết bị điện tử thông minh.



HÌNH 2

Lĩnh vực kỹ thuật được đề cập

Sáng chế liên quan đến kỹ thuật xử lý ngôn ngữ tự nhiên, một phân nhánh trong công nghệ trí thông minh nhân tạo. Cụ thể hơn, sáng chế đề cập đến hệ thống và phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp các yêu cầu từ người dùng thông qua các thiết bị điện tử thông minh.

Tình trạng kỹ thuật của sáng chế

Hiện nay, phương pháp và hệ thống xử lý ngôn ngữ tự nhiên được xây dựng theo hai hướng tiếp cận bao gồm so khớp và phân tích văn bản. Cụ thể như sau:

Hướng tiếp cận so khớp thường dùng trong trường hợp các câu nói của người dùng là đơn giản, ngắn gọn. Các áp dụng thường thấy của hướng tiếp cận này là điều khiển thông minh. Ví dụ như người ta điều khiển máy, rô bốt hay các thiết bị tự hành. Các câu lệnh chỉ bao gồm "lên", "xuống", "trái", "phải", "tới", "lui". Tuy quá trình xử lý rất đơn giản, nhưng về cơ bản là hệ thống xử lý được ý của người dùng, qua đó thi hành đúng mệnh lệnh.

Hướng tiếp cận thứ hai dùng các kỹ thuật phân tích văn bản, qua đó xử lý nhiều hơn về cấu trúc của câu nói của người dùng. Hướng tiếp cận này có thể áp dụng trên những câu dài, cấu trúc phức tạp. Tuy nhiên, các phương pháp trong hướng này lại chỉ giải quyết được một nhóm các câu có những nét tương đồng nhất định, không xử lý được một văn bản tổng quát.

Nói cách khác, mỗi phương pháp trong hướng này chỉ có thể xử lý tốt với các câu hỏi về pháp luật, hoặc về giải trí, hoặc về thể thao, hoặc chỉ phù hợp với việc xử lý câu mệnh lệnh, hoặc xử lý câu hỏi,... Trong một phạm vi đủ hẹp, cho dù câu yêu cầu có phức tạp thì hướng này cũng có khả năng giải quyết được. Nhìn chung, phân tích văn bản là một bài toán lớn và hiện chưa có một phương pháp xử lý trọn vẹn.

Trong thực tế sử dụng, có những câu yêu cầu mà cả hai phương pháp trên đều không xử lý được, từ đó dẫn đến câu trả lời bị sai. Một trong những nhóm thường bị xử lý sai nhiều nhất là các biến thể của các câu đơn giản. Phương pháp so khớp chắc chắn không thể áp dụng, còn các kỹ thuật phân tích văn bản cũng không thể áp dụng vì các câu biến thể này rất đa dạng về cấu trúc và có nội dung khó đoán. Ngoài ra, tốc độ xử lý cũng là một vấn đề lớn bởi các hệ thống đối đáp có rất nhiều người sử dụng.

Trong sáng chế mang số đơn US7624007B2 nộp tại Hoa Kỳ, các tác giả đã tiết lộ một hệ thống và phương pháp truy vấn dựa trên câu từ người dùng được phân tích bằng công cụ ngôn ngữ tự nhiên để xác định câu trả lời thích hợp từ cơ sở dữ liệu điện tử. Hệ thống và phương pháp của sáng chế được ứng dụng cho các công cụ tìm kiếm dựa trên Internet, cũng như các hệ thống nhận dạng giọng nói phân tán như hệ thống máy khách-máy chủ.

Trong sáng chế mang số đơn US20160188292A1 nộp tại Hoa Kỳ, các tác giả đã tiết lộ các cơ chế được cung cấp để trả lời một câu hỏi đầu vào. Một câu hỏi đầu vào cần trả lời từ một nguồn được nhận và xử lý cho một hoặc nhiều tính năng trích xuất của câu hỏi đầu vào. Một hoặc nhiều tính năng được trích xuất được so sánh với các tính năng được lưu trong bộ nhớ cache được lưu trữ trong một hoặc nhiều mục của bộ nhớ cache câu hỏi và câu trả lời (QA). Việc xác định liệu có mục nhập phù hợp trong một hoặc nhiều mục nhập

của bộ đệm QA dựa trên kết quả so sánh và nếu có, thông tin câu trả lời của ứng viên được truy xuất từ mục nhập phù hợp. Thông tin câu trả lời của ứng viên được truy xuất được trả về nguồn của câu hỏi đầu vào dưới dạng thông tin câu trả lời của ứng viên để trả lời câu hỏi đầu vào.

Nhìn chung sáng chế số US7624007B2 và US20160188292A1 đáp ứng được một cách riêng rẽ một vài chức năng cụ thể nào đấy, chúng không xử lý được các biến thể của các câu hỏi/yêu cầu đầu vào; đồng thời cả hai sáng chế không có các phương án tối ưu hóa tốc độ hồi đáp/trả lời từ phía người dùng.

Do đó, điều cần thiết là cần có một hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng đáp ứng được biến thể của nhiều câu hỏi/yêu cầu đầu vào.

Điều cần thiết nữa là cần có một hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng tối ưu hóa tốc độ hồi đáp/trả lời các câu hỏi/yêu cầu từ người dùng bằng cách phân chia nhiều cấp độ xử lý khác nhau với độ khó tăng dần.

Sáng chế này đạt các mục tiêu trên.

Bản chất kỹ thuật của sáng chế

Mục đích thứ nhất của sáng chế là cung cấp một hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng bao gồm: một mạng dựa trên đám mây (cloud); ít nhất một thiết bị điện tử thông minh; một hệ thống thiết bị điện toán; trong đó, thiết bị điện tử thông minh và hệ thống thiết bị điện toán kết nối vào mạng dựa trên đám mây thông qua kênh truyền dẫn internet;

trong đó, thiết bị điện tử thông minh dùng để nhận câu hỏi/yêu cầu đầu vào là các lệnh/các yêu cầu từ người dùng;

trong đó, hệ thống thiết bị điện toán dùng để hồi đáp/trả lời các câu hỏi/yêu cầu đầu vào từ người dùng thông qua thiết bị điện tử thông minh;

trong đó, thiết bị điện tử thông minh bao gồm điện thoại thông minh, máy tính bảng, máy tính để bàn, thiết bị đeo thông minh (smart wearable device), loa thông minh/trợ lý ảo, các thiết bị IoT, robot/xe tự hành;

trong đó, người dùng tương tác với thiết bị điện tử thông minh để ra lệnh/yêu cầu trên ứng dụng Web và/hoặc App thông qua bàn phím cơ/bàn phím ảo và/hoặc giao diện âm thanh (audio interface);

trong đó, câu hỏi đáp/trả lời của hệ thống điện toán 120 bao gồm dạng văn bản, âm thanh và/hoặc ra lệnh cho thiết bị điện tử thông minh thực thi một hành động như mở/tắt nhạc, bật đèn, mở/tắt một chức năng nào đấy trong các thiết bị IoT; khác biệt ở chỗ câu hỏi đáp/trả lời của hệ thống thiết bị điện toán còn bao gồm lập trình nhúng vào trong các thiết bị điện tử thông minh bằng giao diện lập trình ứng dụng (“API” - Application Programming Interface).

Theo đó, hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng còn bao các phương án thực hiện như sau:

Phương án thứ nhất: trong đó hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng còn bao gồm bên thứ ba, là những hàm API được cung ứng từ những hệ thống điện toán của những nhà cung cấp thông tin trực tuyến theo thời gian thực như thông tin tỷ giá, thời tiết, chăm sóc sức khỏe, danh bạ địa chỉ các cơ quan/tổ chức, v.v.

Phương án thứ hai: trong đó hệ thống thiết bị điện toán bao gồm: giao diện mạng đầu vào/đầu ra (I/O), cụm các đơn vị xử lý trung tâm và đơn vị xử lý đồ họa, trình quản lý và thực thi các truy vấn SQL, trình phát triển các ứng dụng phân tán và chuyển đổi cấu trúc dữ liệu, ứng dụng hỏi và đáp Web/App/API, trung tâm lưu trữ dữ liệu ngôn ngữ, và lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML);

trong đó, trung tâm dữ liệu ngôn ngữ bao gồm kho dữ liệu và trình xử lý và lưu trữ dữ liệu; trong đó, kho dữ liệu bao gồm kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản, kho dữ liệu bộ câu hỏi và trả lời, kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh và tập ngữ liệu và từ vựng dùng để huấn luyện máy học;

trong đó, ứng dụng hỏi và đáp Web/App/API khi được thực thi bởi cụm các đơn vị xử lý trung tâm và các đơn vị xử lý đồ họa, nó sẽ khởi động ứng dụng hỏi và đáp Web/App/API;

trong đó, ứng dụng hỏi và đáp Web/App/API bao gồm mô-đun truy xuất bộ nhớ truy cập nhanh, mô-đun so khớp câu hỏi/yêu cầu đơn giản, mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào và mô-đun phân tích câu;

trong đó, khi mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào hoặc mô-đun phân tích câu được kích hoạt, chúng ra lệnh cho lớp mô-đun trí tuệ nhân tạo và máy học tính độ tương đồng và/hoặc phân tích câu hỏi/yêu cầu đầu vào bằng ít nhất một mô hình ngôn ngữ (language model) và thực hiện phân lớp câu hỏi/yêu cầu đầu vào để hỏi đáp/trả lời cho người dùng dựa vào việc truy xuất/truy vấn trung tâm dữ liệu ngôn ngữ.

Phương án thứ ba: trong đó giao diện mạng đầu vào/đầu ra (I/O) được sử dụng là công nghệ Cloudflare, trình quản lý và thực thi các truy vấn SQL sử dụng công nghệ Hadoop và Hive, trình phát triển các ứng dụng phân tán và chuyển đổi cấu trúc dữ liệu sử dụng công nghệ Spring cloud và Rest API, trình xử lý và lưu trữ dữ liệu sử dụng công nghệ HDFS và Redis.

Phương án thứ tư: trong đó lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML) còn bao gồm mô hình ngôn ngữ, thuật toán SymSS, thuật toán máy vector hỗ trợ (SVM - Support Vector Machine), thuật toán trường ngẫu nhiên có điều kiện (CRF - Conditional Random Field).

Mục đích thứ hai của sáng chế là cung cấp một phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng bao gồm các bước:

i) cung ứng một hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

ii) nhận một câu hỏi/yêu cầu đầu vào từ người dùng thông qua thiết bị điện tử thông minh; trong đó, câu hỏi/yêu cầu đầu vào từ người dùng là lệnh/các yêu cầu gửi đến thiết bị điện tử thông minh trên ứng dụng Web và/hoặc App thông qua bàn phím cơ/bàn phím ảo và/hoặc giao diện âm thanh (audio interface);

iii) mô-đun truy xuất bộ nhớ truy cập nhanh thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh (cache);

trong đó, bước iii) còn được thực thi từ trình xử lý và lưu trữ dữ liệu;

trong đó, trình xử lý và lưu trữ dữ liệu được thực hiện bởi công nghệ Redis;

iv) xét xem việc truy vấn đã tìm thấy được câu hỏi đáp/trả lời: nếu truy vấn tìm thấy được

câu hỏi đáp/trả lời thì thực hiện bước xi); ngược lại thì thực hiện bước v);

v) xét xem phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản đã được thực hiện hay chưa: nếu phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản đã thực thi thực hiện bước vii); ngược lại thì thực hiện bước vi);

vi) thực hiện phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản và quay lại bước iv);

trong đó, bước vi) do mô-đun so khớp câu hỏi/yêu cầu đơn giản thực hiện;

vii) xét xem phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào đã được thực hiện hay chưa: nếu phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào đã thực thi thì thực hiện bước ix); ngược lại thì thực hiện bước viii);

viii) thực hiện phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào và quay lại bước iv);

trong đó, bước viii) do mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào thực hiện;

ix) xét xem phương pháp phân tích câu hỏi/yêu cầu đầu vào đã được thực hiện hay chưa: nếu phương pháp phân tích câu hỏi/yêu cầu đầu vào đã thực thi thì thực hiện bước xi); ngược lại thì thực hiện bước x);

x) thực hiện phương pháp phân tích câu hỏi/yêu cầu đầu vào và quay lại bước iv);

trong đó, bước x) do mô-đun phân tích câu thực hiện;

xi) trả kết quả truy vấn/hỏi đáp yêu cầu đến người dùng và/hoặc thiết bị điện tử thông minh;

trong đó, kết quả truy vấn/hỏi đáp bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh;

trong đó, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu trả lời (true) hoặc không tìm thấy câu trả lời (false);

trong đó, tùy vào ngữ cảnh câu hỏi/yêu cầu đầu vào từ người dùng mà hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng sẽ hỏi đáp bằng văn bản, hình ảnh, âm thanh hoặc không đưa ra bất kỳ một hành vi nào hướng đến người dùng và/hoặc thiết bị điện tử thông minh;

xii) xét xem kết quả trả về ở bước xi) là một câu hỏi đáp/trả lời đã được tìm thấy/xác định thì thực hiện bước xiii); ngược lại thì kết thúc phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

xiii) thực hiện sắp xếp lại thứ tự dữ liệu ưu tiên trong bộ nhớ truy cập nhanh nhằm tối ưu hóa dung lượng bộ nhớ truy cập nhanh và nhằm tối ưu hóa thời gian truy vấn các câu hỏi đáp/trả lời các câu yêu cầu đầu vào của người dùng;

trong đó, bước xiii) sử dụng thuật toán LRU (Least Recently Used); đó là xóa các dữ liệu ít được dùng gần đây nhất trong bộ nhớ truy cập nhanh và thuật toán LRU được lập trình xử lý ở trình xử lý và lưu trữ dữ liệu.

Ngoài ra, phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng còn bao gồm các phương án sau đây:

Phương án thứ nhất: trong đó phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản bao gồm các bước:

i) nhận một câu hỏi/yêu cầu đầu vào từ phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

ii) truy vấn câu hỏi/yêu cầu đầu vào từ kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản;

iii) xét xem dữ liệu truy vấn ở bước ii) có được tìm thấy: nếu truy vấn tìm thấy dữ liệu thì thực hiện bước iv); ngược lại thì thực hiện bước vi);

iv) truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời;

trong đó, mẫu câu hỏi/yêu cầu đơn giản là câu có một số ít các từ nhưng câu vẫn đủ ngữ/nghĩa;

trong đó, mẫu câu hỏi/yêu cầu đơn giản này người dùng có thể tự cập nhật và/hoặc là những mẫu câu hỏi/yêu cầu đã được hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng thiết lập sẵn;

v) thực hiện cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh;

vi) trả kết quả truy vấn và quay lại phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

trong đó, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh;

trong đó, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu hỏi đáp/trả lời (true) hoặc không tìm thấy câu hỏi đáp/trả lời (false).

Phương án thứ hai: trong đó phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào bao gồm các bước:

i) nhận một câu hỏi/yêu cầu đầu vào từ phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

ii) thực hiện tính và so sánh độ tương đồng giữa câu hỏi/yêu cầu đầu vào với các mẫu câu hỏi/yêu cầu đơn giản trong kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản với ngưỡng tương đồng N được cho trước;

trong đó, N là số thực và $N \in [0,1]$;

trong đó, bước ii) gồm các bước con như sau:

a) thực hiện tính độ tương đồng X về cấu trúc giữa câu hỏi/yêu cầu đầu vào với mẫu câu đơn giản; cụ thể là hai câu sẽ lần lượt được phân tích cú pháp, sau đó so sánh hai cây cú pháp với nhau bằng thuật toán SyMSS; trong đó X là số thực và $X \in [0,1]$;

b) thực hiện tính độ tương đồng Y về nội dung giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản; cụ thể là hai câu sẽ lần lượt được vector hóa (vectorization - biểu diễn các câu thành một cấu trúc vector) bằng mô hình ngôn ngữ, sau đó tính độ tương đồng cosin (cosine similarity) giữa hai vector này với nhau; trong đó Y là số thực và $Y \in [0,1]$;

c) thực hiện tính độ tương đồng Z giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản bằng cách tính trung bình các kết quả ở bước a) và bước b); cụ thể $Z = \frac{(X+Y)}{2}$ và $Z \in [0,1]$; trong đó, nếu Z có giá trị càng lớn, nghĩa là Z tiến gần 1, thì sự tương đồng/giống nhau về ngữ nghĩa giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản càng lớn;

d) thực hiện so sánh độ tương đồng Z với ngưỡng tương đồng N cho trước;

iii) xét xem có tìm thấy bất kỳ độ tương đồng Z nào vượt ngưỡng hay không, nghĩa là $Z \geq N$: nếu có thì thực hiện bước iv); ngược lại thì thực hiện bước viii);

iv) lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh;

v) xét xem bước iv) có truy vấn được câu hỏi đáp/trả lời hay không: nếu truy vấn được câu hỏi đáp/trả lời thì thực hiện bước vii); ngược lại thì thực hiện bước vi);

vi) lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời;

vii) thực hiện cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh; và kho dữ liệu bộ câu hỏi/yêu cầu và trả lời;

viii) trả kết quả truy vấn và quay lại phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

trong đó, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đây trên thiết bị điện tử thông minh;

trong đó, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu hỏi đáp/trả lời (true) hoặc không tìm thấy câu hỏi đáp/trả lời (false).

Phương án thứ ba: trong đó phương pháp phân tích câu đầu vào bao gồm các bước:

i) nhận một câu hỏi/yêu cầu đầu vào từ phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

ii) thực hiện phân loại câu hỏi/yêu cầu đầu vào và bao gồm các bước con như sau:

a) thực hiện trích xuất đặc trưng của câu hỏi/yêu cầu đầu vào;

trong đó, bước được thực hiện bằng mô hình ngôn ngữ, cụ thể là mô-đun phân tích câu ra lệnh cho lớp mô-đun trí tuệ nhân tạo và máy học thực hiện trích xuất đặc trưng của câu hỏi/yêu cầu đầu vào;

b) từ các đặc trưng được trích xuất ở bước a), mô-đun phân tích câu tiếp tục phân lớp/lựa chọn các đặc trưng được trích xuất ở bước a) để tìm kiếm câu hỏi đáp/trả lời chính xác nhất;

trong đó, bước b) được thực hiện bởi thuật toán máy vector hỗ trợ (SVM - Support Vector Machine), chúng được lưu trữ và thực thi ở lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML).

iii) xét xem bước ii) có phân loại được câu hỏi/yêu cầu đầu vào: nếu có thì thực bước iv); ngược lại thì thực hiện bước ix);

iv) thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời dựa trên các đặc trưng đã phân loại được ở bước ii);

v) xét xem truy vấn câu hỏi đáp/trả lời ở bước iv) đã có đủ thông tin hay không: nếu có thì thực hiện bước ix); ngược lại thì thực hiện bước vi);

vi) thực hiện truy vấn lấy thêm các thông tin từ kho dữ liệu bộ câu hỏi và trả lời;

vii) thực hiện rút trích thông tin;

trong đó, việc rút trích thông tin sẽ được hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng sử dụng kỹ thuật của máy học để trích xuất thông tin;

trong đó, việc rút trích thông tin còn bao gồm thuật toán trường ngẫu nhiên có điều kiện (CRF - Conditional Random Field);

viii) xét xem việc rút trích thông tin ở bước vii) có thành công hay không: nếu rút trích thành công, nghĩa là thông tin rút trích được đủ dùng để truy vấn câu hỏi đáp/trả lời thì thực hiện bước xiii); ngược lại thì thực hiện bước ix);

ix) thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh;

x) xét xem bước ix) có hay không truy vấn được câu hỏi đáp/trả lời: nếu truy vấn được câu hỏi đáp/trả lời thì đi đến thực hiện bước xiii); ngược lại thì thực hiện bước xi);

xi) truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời;

xii) cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh; và kho dữ liệu bộ câu hỏi và trả lời;

xiii) trả kết quả truy vấn và quay lại phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

trong đó, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh;

trong đó, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu hỏi đáp/trả lời (true) hoặc không tìm thấy câu hỏi đáp/trả lời (false).

Mục đích thứ hai của sáng chế là cung cấp một hệ thống và phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng có khả năng hỏi đáp/trả lời nhanh câu hỏi/yêu cầu của người dùng. Ưu điểm này của sáng chế có được bằng cách sử dụng dữ liệu mẫu câu đơn giản và bộ nhớ truy cập nhanh (cache).

Mục đích thứ ba của sáng chế là cung cấp một hệ thống và phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng đáp ứng được nhiều thể loại cấu trúc câu hỏi/đầu vào từ đơn giản đến phức tạp.

Những ưu điểm này và những ưu điểm khác của sáng chế sẽ trở nên rõ ràng đối với những người có hiểu biết thông thường trong lĩnh vực kỹ thuật sau khi đọc phần mô tả chi tiết sau đây về các phương án được ưu tiên, được minh họa trên các hình vẽ khác nhau.

Mô tả vắn tắt các hình vẽ

Các bản vẽ đi kèm, được kết hợp và tạo thành một phần của đặc điểm kỹ thuật này, minh họa các phương án của sáng chế và cùng với mô tả, phục vụ để giải thích các nguyên

tắc của sáng chế.

Hình 1 là sơ đồ hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo phương án mẫu mực của sáng chế;

Hình 2 là lưu đồ minh họa phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo phương án mẫu mực của sáng chế;

Hình 3 là lưu đồ minh họa phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo phương án mẫu mực của sáng chế;

Hình 4 là lưu đồ minh họa phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo phương án mẫu mực của sáng chế;

Hình 5 là lưu đồ minh họa phương pháp phân tích câu hỏi/yêu cầu đầu vào trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo phương án mẫu mực của sáng chế.

Các hình vẽ mô tả chỉ nhằm mục đích minh họa các phương án khác nhau của công nghệ. Một người có kỹ năng thông thường trong kỹ thuật sẽ dễ dàng công nhận từ các mô tả chi tiết sau đây rằng các phương án thay thế của các cấu trúc và phương pháp được minh họa trong tài liệu này có thể được sử dụng mà không rời khỏi các nguyên lý của công nghệ được mô tả trong tài liệu này.

Mô tả chi tiết sáng chế

Phần sau đây sẽ xem xét chi tiết các phương án được ưu tiên của sáng chế, các ví dụ về chúng được minh họa trên các hình vẽ đi kèm. Mặc dù sáng chế sẽ được mô tả cùng với các phương án được ưu tiên, nhưng sẽ được hiểu rằng sáng chế không có ý định giới hạn đối với các phương án này. Ngược lại, sáng chế nhằm mục đích đảm bảo các lựa chọn thay thế, cải tiến và định lượng, có thể được bao gồm trong tinh thần và phạm vi của sáng chế được xác định bởi các yêu cầu bảo hộ được thêm vào. Hơn nữa, trong phần mô tả chi tiết sau đây của sáng chế, nhiều chi tiết cụ thể được đặt ra để cung cấp sự hiểu biết thấu đáo về sáng chế. Tuy nhiên, sẽ là hiển nhiên đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này rằng sáng chế có thể được thực hành mà không có các chi tiết cụ thể này. Trong các trường hợp khác, các phương pháp đã biết, quy trình, thành phần và phạm vi ứng dụng đã không được mô tả chi tiết để không làm mờ đi các khía cạnh không cần thiết của sáng chế.

Cần phải hiểu thêm rằng thuật ngữ được sử dụng ở đây được sử dụng cho mục đích chỉ mô tả các phương án cụ thể và không nhằm mục đích giới hạn phạm vi của sáng chế. Cần lưu ý rằng các thuật ngữ được sử dụng ở đây và trong các yêu cầu bảo hộ, các hình thức số ít, “một” và “một số ít” bao gồm các tham chiếu số nhiều trừ khi bối cảnh chỉ ra rõ ràng. Do đó, ví dụ, một tài liệu tham khảo về một phần tử, một phần tử tham chiếu đến một hoặc nhiều phần tử và bao gồm các phần tử tương đương được biết đến với những người có kỹ năng thông thường trong lĩnh vực kỹ thuật tương đương. Tương tự, đối với một ví dụ khác, một tham chiếu đến một bước đi hay một cách có nghĩa là một cách tham chiếu đến một hoặc nhiều bước hoặc phương tiện và có thể bao gồm các bước phụ và phương tiện phụ thuộc. Tất cả các liên từ được sử dụng phải được hiểu theo nghĩa bao hàm nhất có thể. Do đó, từ ngữ “ưu tiên” hay “nên” được hiểu là có định nghĩa về một logic hoặc một logic chứ không phải là một từ hay logic độc quyền, trừ khi bối cảnh chỉ ra rõ ràng và cần thiết. Các

cấu trúc được mô tả trong tài liệu này cũng được hiểu là để chỉ các chức năng tương đương của các cấu trúc đó. Ngôn ngữ có thể được hiểu để diễn đạt gần đúng nên được hiểu như vậy trừ khi bối cảnh khác chỉ rõ.

Trừ khi được định nghĩa khác, tất cả các thuật ngữ khoa học và kỹ thuật được sử dụng ở đây đều có cùng ý nghĩa như thường được hiểu bởi một trong những kỹ năng thông thường trong kỹ thuật mà sáng chế này đề cập. Các cấu trúc được mô tả trong tài liệu này cũng được hiểu là để chỉ các chức năng tương đương của các cấu trúc đó. Các tài liệu tham khảo về phương án của một trong những phương án, hiện thực, ví dụ, có thể chỉ ra rằng các phương án của sáng chế được mô tả có thể bao gồm một tính năng, cấu trúc hoặc đặc điểm cụ thể, nhưng không phải mọi phương án nhất thiết phải bao gồm các tính năng, cấu trúc hoặc đặc tính cụ thể. Hơn nữa, việc sử dụng lặp đi lặp lại cụm từ trong một phương án, trực tiếp hoặc trong một phương án mẫu mực, không nhất thiết phải đề cập đến cùng một phương án, mặc dù chúng có thể.

Tiêu đề được cung cấp ở đây là để thuận tiện trong việc mô tả sáng chế và không được coi là hạn chế tiết lộ dưới bất kỳ hình thức nào.

Các phương án mẫu mực và khía cạnh của sáng chế được mô tả với tham chiếu từ Hình 1 đến Hình 5. Công bố hiện tại tiết lộ các tính năng sau của sáng chế: (1) hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng, (2) phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng, (3) phương pháp so khớp mẫu câu hỏi/yêu cầu đơn giản trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng, (4) phương pháp tính độ tương đồng giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng, và (5) phương pháp phân tích câu hỏi/yêu cầu đầu vào trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng.

Tham chiếu đến Hình 1, là sơ đồ minh họa hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng (“hệ thống OLLI 100”) theo phương án của sáng chế. Hệ thống OLLI 100 bao gồm một mạng dựa trên đám mây 102 (cloud), ít nhất một thiết bị điện tử thông minh 110, một hệ thống thiết bị điện toán 120; trong đó, thiết bị điện tử thông minh 110 và hệ thống thiết bị điện toán 120 kết nối vào mạng dựa trên đám mây 102 thông qua kênh truyền dẫn internet 103. Thiết bị điện tử thông minh 110 dùng để nhận câu hỏi/yêu cầu đầu vào là các lệnh/các yêu cầu từ người dùng 101; hệ thống thiết bị điện toán 120 dùng để hồi đáp/trả lời các câu hỏi/yêu cầu đầu vào từ người dùng 101 thông qua thiết bị điện tử thông minh 110. Theo các phương án của sáng chế, thiết bị điện tử thông minh 110 bao gồm, nhưng không giới hạn, là điện thoại thông minh, máy tính bảng, máy tính để bàn, thiết bị đeo thông minh (smart wearable device), loa thông minh/trợ lý ảo, các thiết bị IoT, robot/xe tự hành, v.v. Người dùng 101 tương tác với thiết bị điện tử thông minh 110 để ra lệnh/yêu cầu trên ứng dụng Web và/hoặc App thông qua bàn phím cơ/bàn phím ảo và/hoặc giao diện âm thanh (audio interface), v.v. Theo các phương án của sáng chế, phần hồi đáp/trả lời của hệ thống điện toán 120 bao gồm dạng văn bản, âm thanh và/hoặc ra lệnh cho thiết bị điện tử thông minh 110 thực thi một hành động như mở/tắt nhạc, bật đèn, mở/tắt một chức năng nào đó trong các thiết bị IoT, v.v. Theo một phương án được ưu tiên của sáng chế, phần hồi đáp/trả lời của hệ thống thiết bị điện toán 120 được lập trình nhúng vào trong các thiết bị điện tử thông minh 110 bằng giao diện lập trình ứng dụng (“API” - Application Programming Interface).

Theo các phương án của sáng chế, hệ thống OLLI 100 còn bao gồm bên thứ ba 130, là những hàm API được cung ứng từ những hệ thống điện toán của những nhà cung cấp thông tin trực tuyến theo thời gian thực như thông tin tỷ giá, thời tiết, chăm sóc sức khỏe, danh bạ địa chỉ các cơ quan/tổ chức, v.v.

Tiếp tục với Hình 1, hệ thống thiết bị điện toán 120 bao gồm giao diện mạng đầu vào/đầu ra (I/O) 121, cụm các đơn vị xử lý trung tâm và đơn vị xử lý đồ họa 122, trình quản lý và thực thi các truy vấn SQL 123, trình phát triển các ứng dụng phân tán và chuyển đổi cấu trúc dữ liệu 124, trung tâm lưu trữ dữ liệu ngôn ngữ 125 và lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML) 126. Theo phương án của sáng chế, trung tâm lưu trữ dữ liệu ngôn ngữ 125 bao gồm kho dữ liệu 125a và trình xử lý và lưu trữ dữ liệu 125b; trong đó, kho dữ liệu 125a bao gồm kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản 125a1, kho dữ liệu bộ câu hỏi và trả lời 125a2, kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh 125a3 và tập ngữ liệu và từ vựng dùng để huấn luyện máy học 125a4. Theo các phương án của sáng chế, lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML) 126 còn bao gồm mô hình ngôn ngữ, thuật toán SyMSS, thuật toán máy vector hỗ trợ (SVM - Support Vector Machine), thuật toán trường ngẫu nhiên có điều kiện (CRF - Conditional Random Field); đây là những kỹ thuật được ứng dụng rộng rãi trong lĩnh vực trí thông minh nhân tạo nên chúng sẽ không được diễn giải chi tiết ở đây, trong sáng chế này, để không làm mờ đi các khía cạnh khác của sáng chế.

Vẫn với Hình 1, thiết bị điện toán 120 còn bao gồm ứng dụng hỏi và đáp Web/App/API 127 (“ứng dụng OLLI Mika 127”) khi được thực thi bởi cụm các đơn vị xử lý trung tâm và các đơn vị xử lý đồ họa 122, nó sẽ khởi động ứng dụng OLLI Mika 127. Các mô-đun ứng dụng OLLI Mika 127 bao gồm mô-đun truy xuất bộ nhớ truy cập nhanh 127a, mô-đun so khớp câu hỏi/yêu cầu đơn giản 127b, mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào 127c và mô-đun phân tích câu 127d. Tùy vào ngữ cảnh theo yêu cầu của người dùng 101 mà các mô-đun ở trong ứng dụng OLLI Mika 127 sẽ được kích hoạt. Chi tiết về phương thức hoạt động của ứng dụng OLLI Mika 127 sẽ được thảo luận chi tiết trong các hình, từ Hình 2 - Hình 5 và sẽ thảo luận sau. Theo các phương án của sáng chế, ứng dụng OLLI Mika 127, lớp mô-đun trí tuệ nhân tạo và máy học 126 có thể được viết bằng bất kỳ sự kết hợp nào của một hoặc nhiều ngôn ngữ lập trình, bao gồm ngôn ngữ lập trình hướng đối tượng như Oracle, DB2, Microsoft SQL server, MySQL, noQuery, MongoDB, BigTable, Redis, Earn DB, Cassandra, Hbase, Neo4J, Python, Java, Smalltalk, C ++ hoặc các ngôn ngữ lập trình thủ tục thông thường và tương tự, chẳng hạn ngôn ngữ lập trình “C”.

Theo các phương án của sáng chế, các lệnh ở trong ứng dụng OLLI Mika 127, lớp mô-đun trí tuệ nhân tạo và máy học 126 có thể được lưu trữ và được tải vào máy tính theo một dạng ứng dụng có thể đọc được trên máy tính, hoặc thiết bị xử lý dữ liệu lập trình khác hoặc các thiết bị khác hoạt động theo một cách cụ thể sao cho các câu lệnh được lưu trữ trong ứng dụng mà máy tính vừa có thể đọc được vừa có thể thực thi các hàm/thủ tục được chỉ định trong lưu đồ và/hoặc khối hoặc biểu đồ khối.

Theo phương án được ưu tiên của sáng chế, giao diện mạng đầu vào/đầu ra (I/O) 121 được sử dụng là công nghệ Cloudflare, trình quản lý và thực thi các truy vấn SQL 123 sử dụng công nghệ Hadoop và Hive, trình phát triển các ứng dụng phân tán và chuyển đổi cấu trúc dữ liệu 124 sử dụng công nghệ Spring cloud và Rest API, trình xử lý và lưu trữ dữ liệu 125b sử dụng công nghệ HDFS và Redis. Các kỹ thuật/công nghệ cloudflare, Spring cloud, Rest API, HDFS, Redis, Hadoop, Hive và API là quá rõ ràng, chúng được ứng dụng phổ

biến trong lĩnh vực công nghệ thông tin nên sẽ không được mô tả chi tiết ở đây, trong sáng chế này, nhất là sáng chế ở dạng tiết lộ để không làm mờ đi các khía cạnh khác của sáng chế.

Vấn với Hình 1, trong một số trường hợp, khi mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào 127c hoặc mô-đun phân tích câu 127d được kích hoạt, chúng ra lệnh cho lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML) 126 tính độ tương đồng và/hoặc phân tích câu hỏi/yêu cầu đầu vào bằng ít nhất một mô hình ngôn ngữ (language model) và thực hiện phân lớp câu hỏi/yêu cầu đầu vào để hỏi đáp/trả lời cho người dùng 101 dựa vào việc truy xuất/truy vấn trung tâm dữ liệu ngôn ngữ 125.

Bây giờ thảo luận đến Hình 2, là lưu đồ minh họa phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng 101 được thực hiện trên hệ thống OLLI 100 theo phương án mẫu mực của sáng chế. Phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng (“phương pháp 200”) bắt đầu bằng bước 201, đó là cung ứng một hệ thống OLLI 100 như đã được mô tả ở Hình 1;

Ở bước 202, nhận một câu hỏi/yêu cầu đầu vào từ người dùng 110 thông qua thiết bị điện tử thông minh 110; theo phương án của sáng chế, câu hỏi/yêu cầu đầu vào từ người dùng 101 là lệnh/các yêu cầu gửi đến thiết bị điện tử thông minh 110 trên ứng dụng Web và/hoặc App thông qua bàn phím cơ/bàn phím ảo và/hoặc giao diện âm thanh (audio interface), v.v.

Ở bước 203, mô-đun truy xuất bộ nhớ truy cập nhanh 127a thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh (cache) 125a3. Theo phương án của sáng chế, việc truy vấn câu hỏi đáp/trả lời từ bộ nhớ truy cập nhanh nhằm giảm thiểu thời gian xử lý của hệ thống OLLI 100, tối ưu hóa việc hỏi đáp/trả lời kết quả đến người dùng 101. Ngoài ra, việc tối ưu hóa thực hiện bước 202 còn có sự thực thi từ trình xử lý và lưu trữ dữ liệu 125b. Theo một phương án được ưu tiên, trình xử lý và lưu trữ dữ liệu 125b được thực hiện bởi công nghệ Redis;

Ở bước 204, xét xem việc truy vấn đã tìm thấy được câu hỏi đáp/trả lời: nếu truy vấn tìm thấy được câu hỏi đáp/trả lời thì thực hiện bước 211; ngược lại thì thực hiện bước 205. Ở các bước tiếp theo, từ bước 205 đến bước 210, phương pháp 200 thực hiện duyệt qua các phương pháp khác nhau để truy vấn câu hỏi đáp/trả lời phù hợp với yêu cầu của câu hỏi/yêu cầu đầu vào từ người dùng 101; các phương pháp này được xây dựng từ mức độ đơn giản đến phức tạp nhằm tối ưu hóa thời gian truy vấn câu hỏi đáp/trả lời để hỏi đáp cho người dùng 101. Theo phương án của sáng chế, phương pháp 100, từ bước 205 đến bước 210, sẽ ngừng truy vấn câu hỏi đáp/trả lời khi một trong hai điều kiện sau đây được thỏa mãn:

- a) một trong các phương pháp là phương pháp 300, phương pháp 400, phương pháp 500 truy vấn được câu hỏi đáp/trả lời; hoặc
- b) cả ba phương pháp là phương pháp 300, phương pháp 400, phương pháp 500 đều không truy vấn được câu hỏi đáp/trả lời;

Ở bước 205, xét xem phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản (“phương pháp 300”) đã được thực hiện hay chưa: nếu phương pháp 300 đã thực thi thực hiện bước 207; ngược lại thì thực hiện bước 206;

Ở bước 206, thực hiện phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản và quay lại bước 204. Theo phương án của sáng chế, bước 206 do

mô-đun so khớp câu hỏi/yêu cầu đơn giản 127b thực hiện;

Ở bước 207, xét xem phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào (“phương pháp 400”) đã được thực hiện hay chưa: nếu phương pháp 400 đã thực thi thì thực hiện bước 209; ngược lại thì thực hiện bước 208;

Ở bước 208, thực hiện phương pháp 400 và quay lại bước 204. Theo phương án của sáng chế, bước 208 do mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào 127c thực hiện;

Ở bước 209, xét xem phương pháp phân tích câu hỏi/yêu cầu đầu vào (“phương pháp 500”) đã được thực hiện hay chưa: nếu phương pháp 500 đã thực thi thì thực hiện bước 211; ngược lại thì thực hiện bước 210;

Ở bước 210, thực hiện phương pháp 500 và quay lại bước 204. Theo phương án của sáng chế, bước 210 do mô-đun phân tích câu 127d thực hiện;

Ở bước 211, trả kết quả truy vấn/hỏi đáp yêu cầu đến người dùng 101 và/hoặc thiết bị điện tử thông minh 110. Theo phương án của sáng chế, kết quả truy vấn/hỏi đáp bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh 110. Theo phương án của sáng chế, kết quả truy vấn/hỏi đáp bao gồm hai trạng thái: Truy vấn được câu trả lời (True) hoặc không tìm thấy câu trả lời (False); tùy vào ngữ cảnh câu hỏi/yêu cầu đầu vào từ người dùng 101 mà hệ thống 100 sẽ hỏi đáp bằng văn bản, hình ảnh, âm thanh hoặc không đưa ra bất kỳ một hành vi nào hướng đến người dùng 101 và/hoặc thiết bị điện tử thông minh 110;

Ở bước 212, xét xem kết quả trả về ở bước 211 là một câu hỏi đáp/trả lời đã được tìm thấy/xác định thì thực hiện bước 213; ngược lại thì kết thúc phương pháp 200;

Ở bước 213, Thực hiện sắp xếp lại thứ tự dữ liệu ưu tiên trong bộ nhớ truy cập nhanh nhằm tối ưu hóa dung lượng bộ nhớ truy cập nhanh và nhằm tối ưu hóa thời gian truy vấn các câu hỏi đáp/trả lời các câu yêu cầu đầu vào của người dùng 101. Theo một phương án được ưu tiên, bước 213 sử dụng thuật toán LRU (Least Recently Used); đó là xóa các dữ liệu ít được dùng gần đây nhất trong bộ nhớ truy cập nhanh và thuật toán LRU được lập trình xử lý ở trình xử lý và lưu trữ dữ liệu 125b.

Bây giờ thảo luận đến Hình 3, là lưu đồ minh họa phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng 101 theo phương án mẫu mực của sáng chế. Phương pháp so khớp câu hỏi/yêu cầu đầu vào đơn giản với mẫu câu hỏi/yêu cầu đơn giản (“phương pháp 300”) bao gồm các bước sau đây:

Bước 301, nhận một câu hỏi/yêu cầu đầu vào ở bước 206 của phương pháp 200;

Ở bước 302, truy vấn câu hỏi/yêu cầu đầu vào từ kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản 125a1;

Ở bước 303, xét xem dữ liệu truy vấn ở bước 302 có được tìm thấy: nếu truy vấn tìm thấy dữ liệu thì thực hiện bước 304; ngược lại thì thực hiện bước 306;

Ở bước 304, truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời 125a2;

Theo các khía cạnh của sáng chế, mẫu câu hỏi/yêu cầu đơn giản là câu có một số ít các từ nhưng câu vẫn đủ ngữ/nghĩa như: mở/tắt nhạc, chạy, ngừng, bây giờ là mấy giờ, thời

tiết Sài Gòn hôm nay thế nào, v.v..

Theo phương án của sáng chế, mẫu câu hỏi/yêu cầu đơn giản này người dùng 101 có thể tự cập nhật và/hoặc là những mẫu câu hỏi/yêu cầu đã được hệ thống OLLI 100 thiết lập sẵn; và giả thuyết những mẫu câu hỏi/yêu cầu đơn giản này luôn có câu trả lời từ kho dữ liệu bộ câu hỏi và trả lời 125a2;

Ở bước 305, thực hiện cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh 125a3;

Ở bước 306, trả kết quả truy vấn và quay lại phương pháp 200. Theo phương án của sáng chế, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh 110. Theo phương án của sáng chế, kết quả truy vấn/hỏi đáp bao gồm hai trạng thái: Truy vấn được câu hỏi đáp/trả lời (True) hoặc không tìm thấy câu hỏi đáp/trả lời (False).

Tiếp theo là thảo luận đến Hình 4, là lưu đồ minh họa phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo phương án mẫu mực của sáng chế. Phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào (“phương pháp 400”) bao gồm các bước sau đây:

Bước 410, nhận một câu hỏi/yêu cầu đầu vào ở bước 208 của phương pháp 200;

Bước 420, thực hiện tính và so sánh độ tương đồng giữa câu hỏi/yêu cầu đầu vào với các mẫu câu hỏi/yêu cầu đơn giản trong kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản 125a1 với ngưỡng tương đồng N được cho trước; trong đó N là số thực và $N \in [0,1]$. Theo các phương án của sáng chế, bước tính và so sánh độ tương đồng giữa câu hỏi/yêu cầu đầu vào với các mẫu câu hỏi/yêu cầu đơn giản gồm các bước con như sau:

Ở bước 421, thực hiện tính độ tương đồng X về cấu trúc giữa câu hỏi/yêu cầu đầu vào với mẫu câu đơn giản; cụ thể là hai câu sẽ lần lượt được phân tích cú pháp, sau đó so sánh hai câu cú pháp với nhau bằng thuật toán SyMSS; trong đó X là số thực và $X \in [0,1]$;

Ở bước 422, thực hiện tính độ tương đồng Y về nội dung giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản; cụ thể là hai câu sẽ lần lượt được vector hóa (vectorization - biểu diễn các câu thành một cấu trúc vector) bằng mô hình ngôn ngữ, sau đó tính độ tương đồng cosin (cosine similarity) giữa hai vector này với nhau; trong đó Y là số thực và $Y \in [0,1]$;

Ở bước 423, thực hiện tính độ tương đồng Z giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản bằng cách tính trung bình các kết quả ở bước 421 và bước 422; cụ thể $Z = \frac{(X+Y)}{2}$ và $Z \in [0,1]$; trong đó, nếu Z có giá trị càng lớn, nghĩa là Z càng tiến gần 1, thì sự tương đồng/giống nhau về ngữ nghĩa giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản càng lớn;

Cần lưu ý rằng thuật ngữ vector hóa văn bản/cấu trúc câu, độ tương đồng cosin là những kỹ thuật được ứng dụng rộng rãi trong lĩnh vực trí thông minh nhân tạo nên chúng sẽ không được diễn giải chi tiết ở đây để không làm mờ đi các khía cạnh khác của sáng chế.

Ở bước 424, thực hiện so sánh độ tương đồng Z với ngưỡng tương đồng N cho trước;

Bước 430, xét xem có tìm thấy bất kỳ độ tương đồng Z nào vượt ngưỡng hay không,

nghĩa là $Z \geq N$: nếu có thì thực hiện bước 440; ngược lại thì thực hiện bước 480;

Bước 440, lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh 125a3;

Bước 450, xét xem bước 440 có truy vấn được câu hỏi đáp/trả lời hay không: nếu truy vấn được câu hỏi đáp/trả lời thì thực hiện bước 470; ngược lại thì thực hiện bước 460;

Bước 460, lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời 125a2;

Bước 470, thực hiện cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh 125a3; và kho dữ liệu bộ câu hỏi/yêu cầu và trả lời 125a2;

Bước 480, trả kết quả truy vấn và quay lại phương pháp 200. Kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh 110. Theo phương án của sáng chế, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: Truy vấn được câu hỏi đáp/trả lời (True) hoặc không tìm thấy câu hỏi đáp/trả lời (False).

Cuối cùng là thảo luận đến Hình 5, là lưu đồ minh họa phương pháp phân tích câu đầu vào trong phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp các yêu cầu từ người dùng theo phương án mẫu mực của sáng chế. Phương pháp phân tích câu (“phương pháp 500”) bao gồm các bước sau đây:

Bước 501, nhận một câu hỏi/yêu cầu đầu vào ở bước 210 của phương pháp 200;

Bước 502, thực hiện phân loại câu hỏi/yêu cầu đầu vào và bao gồm các bước con như sau:

Ở bước 502a, thực hiện trích xuất đặc trưng của câu hỏi/yêu cầu đầu vào. Theo phương án của sáng chế, bước 502a được thực hiện bằng mô hình ngôn ngữ, cụ thể là mô-đun phân tích câu 127d ra lệnh cho lớp mô-đun trí tuệ nhân tạo và máy học 126 thực hiện trích xuất đặc trưng của câu hỏi/yêu cầu đầu vào;

Ở bước 502b, từ các đặc trưng được trích xuất ở bước 502a, mô-đun phân tích câu 127d tiếp tục phân lớp/lựa chọn các đặc trưng được trích xuất ở bước 502a để tìm kiếm câu hỏi đáp/trả lời chính xác nhất. Theo phương án của sáng chế, bước 502b được thực hiện bởi thuật toán máy vector hỗ trợ (SVM - Support Vector Machine), chúng được lưu trữ và thực thi ở lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML) 126. Ví dụ sau đây sẽ minh họa một trường hợp cụ thể của bước 502, nó nhằm giải thích điểm kỹ thuật để người trung bình trong cùng lĩnh vực có thể hiểu được hiện thân của sáng chế; ví dụ không nhằm làm hạn chế phạm vi các yêu cầu bảo hộ:

- a) câu hỏi/yêu cầu đầu vào là "mở cho tôi nghe bài Diễm Xưa của nhạc sĩ Trịnh Công Sơn";
- b) bước 502a, bằng mô hình ngôn ngữ, sẽ biểu diễn đặc trưng của các loại câu ở dạng ma trận $[0.1, 0.52, 0.24, \dots]$;
- c) bước 502b, từ ma trận đặc trưng $[0.1, 0.52, 0.24, \dots]$ sẽ thực hiện phân lớp/lựa chọn các đặc trưng mà câu có ý nghĩa, trong trường hợp này, câu hỏi/yêu cầu có một bộ phân lớp là “mở nhạc”

Bước 503, xét xem bước 502 có phân loại được câu hỏi/yêu cầu đầu vào: nếu có thì thực bước 504; ngược lại thì thực hiện bước 509;

Bước 504, thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời dựa trên các đặc trưng đã phân loại được ở bước 502;

Bước 505, xét xem truy vấn câu hỏi đáp/trả lời ở bước 504 đã có đủ thông tin hay không: nếu có thì thực hiện bước 509; ngược lại thì thực hiện bước 506;

Bước 506, thực hiện truy vấn lấy thêm các thông tin từ kho dữ liệu bộ câu hỏi và trả lời. Như một ví dụ minh họa, thông tin cần bổ sung để hỏi đáp/trả lời cho câu hỏi/yêu cầu đầu vào là bài nhạc nào? tác giả nào?;

Bước 507, thực hiện rút trích thông tin. Theo phương án của sáng chế, việc rút trích thông tin sẽ được hệ thống OLLI 100 sử dụng kỹ thuật của máy học để trích xuất thông tin; cụ thể là dùng thuật toán trường ngẫu nhiên có điều kiện (CRF - Conditional Random Field). Như một ví dụ minh họa, thông tin rút trích bao gồm: bài hát Diễm Xưa, nhạc sĩ Trịnh Công Sơn;

Bước 508, xét xem việc rút trích thông tin ở bước 507 có thành công hay không: nếu rút trích thành công, nghĩa là thông tin rút trích được sẽ đủ dùng để truy vấn câu hỏi đáp/trả lời thì thực hiện bước 513; ngược lại thì thực hiện bước 509;

Bước 509, thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh 125a3;

Bước 510, xét xem bước 509 có hay không truy vấn được câu hỏi đáp/trả lời: nếu truy vấn được câu hỏi đáp/trả lời thì đi đến thực hiện bước 513; ngược lại thì thực hiện bước 511;

Bước 511, truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời 125a2;

Bước 512, cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh 125a3; và kho dữ liệu bộ câu hỏi và trả lời 125a2;

Bước 513, trả kết quả truy vấn và quay lại phương pháp 200. Theo phương án của sáng chế, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh 110. Theo phương án của sáng chế, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: Truy vấn được câu hỏi đáp/trả lời (True) hoặc không tìm thấy câu hỏi đáp/trả lời (False).

Các lưu đồ giải thuật và các bước được mô tả trên đây chỉ là một ví dụ. Có thể có nhiều biến thể cho lưu đồ này hoặc các bước (hoặc thao tác) được mô tả trong đó mà không xuất phát từ tinh thần của sáng chế. Ví dụ, các bước có thể được thực hiện theo thứ tự khác nhau hoặc các bước có thể được thêm, xóa hoặc sửa đổi. Tất cả các biến thể này được coi là một phần yêu cầu của sáng chế.

Những ích lợi (hiệu quả) có thể đạt được

Sáng chế đề cập đến một hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng hoàn toàn tự động. Xét về độ khó của ngôn ngữ, hệ thống và phương pháp của sáng chế sẽ xử lý ở ba cấp độ khác nhau, và ứng với mỗi cấp độ này, hệ thống và phương pháp của sáng chế sẽ dùng những cách thức khác nhau. Nếu một câu hỏi/yêu cầu

đầu vào đã được hệ thống và phương pháp của sáng chế hồi đáp/trả lời được ở một cấp xử lý nào đấy, câu đó sẽ không cần thiết phải được xử lý ở các cấp khác có độ phức tạp lớn hơn. Điều này giúp hệ thống và phương pháp của sáng chế hồi đáp/trả lời được chính xác nội dung của câu hỏi/yêu cầu đầu vào từ người dùng ở đúng độ phức tạp của ngôn ngữ và có khả năng hồi đáp/trả lời một cách phù hợp. Ngoài ra, hệ thống và phương pháp của sáng chế còn sử dụng cơ chế của bộ nhớ truy cập nhanh (cache) nhằm tăng tốc độ xử lý của toàn hệ thống, giúp giảm đáng kể độ trễ của quá trình xử lý, nhất là đối với các câu hỏi/yêu cầu đầu vào phức tạp từ phía người dùng.

Các thảo luận/tuyên bố bên trên đã mô tả đầy đủ ít nhất một phương án của sáng chế, các phương án tương đương hoặc thay thế khác để thực hiện một hệ thống và phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng theo sáng chế sẽ rõ ràng với những người có kỹ năng trung bình trong lĩnh vực tương đương. Các khía cạnh khác nhau của sáng chế đã được mô tả ở trên bằng cách minh họa, và các phương án cụ thể đã được tiết lộ không nhằm mục đích giới hạn sáng chế đối với các hình thức cụ thể. Việc triển khai cụ thể một hệ thống và phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng có thể khác nhau tùy thuộc vào bối cảnh hoặc ứng dụng cụ thể. Do đó, sáng chế bao gồm tất cả các sửa đổi, tương đương và các lựa chọn thay thế nằm trong tinh thần và phạm vi của các yêu cầu bảo hộ sau đây. Cần phải hiểu thêm rằng không phải tất cả các phương án được tiết lộ trong mô tả đã nêu ở trên sẽ nhất thiết phải thỏa mãn hoặc đạt được từng đối tượng.

Lưu đồ giải thuật và biểu đồ khối được mô tả minh họa cấu trúc, chức năng và hoạt động của việc triển khai các hệ thống, các phương pháp và ứng dụng Web/App theo các phương án khác nhau của sáng chế. Về vấn đề này, mỗi khối trong lưu đồ giải thuật hoặc biểu đồ khối có thể đại diện cho một mô-đun, phân đoạn hoặc một phần mã, bao gồm một hoặc nhiều câu lệnh thực thi để thực hiện (các) hàm/thủ tục lô-gíc đã chỉ định. Cũng cần lưu ý rằng, trong một số triển khai thay thế, các hàm/thủ tục ghi trong khối có thể xảy ra không theo trật tự được ghi trong các hình. Ví dụ, trên thực tế, hai khối được hiển thị có thể được thực thi đồng thời hoặc các khối đôi khi có thể được thực hiện theo thứ tự ngược lại tùy thuộc vào chức năng liên quan. Cũng cần lưu ý rằng mỗi khối của lưu đồ và/hoặc biểu đồ khối, và sự kết hợp các khối có thể được thực hiện bởi các hệ thống dựa trên phần cứng được thực hiện với mục đích đặc biệt có các hàm hoặc thủ tục được chỉ định, hoặc kết hợp các câu lệnh đặc biệt về phần cứng máy tính.

Thuật ngữ được sử dụng ở đây chỉ nhằm mục đích mô tả các phương án cụ thể và không nhằm mục đích giới hạn sáng chế. Các thuật ngữ được sử dụng trong tài liệu này như các hình thức số ít, một khu vực, một khu vực và một số ít, được dự định bao gồm cả các hình thức số nhiều, trừ khi bối cảnh chỉ ra rõ ràng khác. Nó sẽ được hiểu thêm rằng các điều khoản bao gồm và/hoặc bao gồm, khi được sử dụng trong thông số kỹ thuật này, chỉ định sự hiện diện của các tính năng, số nguyên, bước, hoạt động, yếu tố và/hoặc thành phần đã nêu, nhưng không loại trừ sự hiện diện hoặc bổ sung của một hoặc nhiều tính năng khác, số nguyên, bước, thao tác, thành phần phần tử và/hoặc nhóm của chúng.

Yêu cầu các yếu tố và các bước trong tài liệu này có thể đã được đánh số và/hoặc ký tự chỉ như một sự trợ giúp cho khả năng đọc và hiểu. Bất kỳ việc đánh số và ký tự nào như vậy không nhằm mục đích và không nên được thực hiện để chỉ ra thứ tự của các yếu tố và/hoặc các bước trong yêu cầu bảo hộ.

Tóm tắt sáng chế được cung cấp sẽ cho phép người đọc xác định bản chất và ý chính

của công bố kỹ thuật. Nó được trình bày với sự hiểu biết rằng nó sẽ không được sử dụng để giới hạn hoặc giải thích phạm vi hoặc ý nghĩa của các yêu cầu bảo hộ. Các yêu cầu bảo hộ sau đây đã được đưa vào mô tả chi tiết, với mỗi yêu cầu bảo hộ đứng riêng như một phương án độc lập.

Chú thích hình vẽ

- 100 Hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp các yêu cầu từ người dùng
- 101 Người dùng
- 102 Mạng dựa trên đám mây
- 103 Kênh truyền dẫn internet
- 110 Thiết bị điện tử thông minh
- 120 Hệ thống thiết bị điện toán
- 121 Giao diện mạng đầu vào/đầu ra (I/O)
- 122 Cụm các đơn vị xử lý trung tâm và đơn vị xử lý đồ họa
- 123 Trình quản lý và thực thi các truy suất SQL
- 124 Trình phát triển các ứng dụng phân tán và chuyển đổi cấu trúc dữ liệu
- 125 Trung tâm lưu trữ dữ liệu ngôn ngữ
- 125a Kho dữ liệu
- 125a1 Kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản
- 125a2 Kho dữ liệu bộ câu hỏi và trả lời
- 125a3 Kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh
- 125a4 Tập ngữ liệu và từ vựng dùng để huấn luyện máy học
- 125b Trình xử lý và lưu trữ dữ liệu
- 126 Lớp mô-đun trí tuệ nhân tạo và máy học
- 127 Ứng dụng hỏi và đáp Web/App/API (ứng dụng OLLI Mika)
- 127a Mô-đun truy xuất bộ nhớ truy cập nhanh
- 127b Mô-đun so khớp câu hỏi/yêu cầu đơn giản
- 127c Mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào
- 127d Mô-đun phân tích câu
- 130 Bên thứ ba

YÊU CẦU BẢO HỘ

1. Hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng bao gồm:

một mạng dựa trên đám mây (cloud);

ít nhất một thiết bị điện tử thông minh;

một hệ thống thiết bị điện toán, bao gồm ứng dụng hỏi và đáp Web/App/API, và trung tâm lưu trữ dữ liệu ngôn ngữ;

trong đó, trung tâm dữ liệu ngôn ngữ bao gồm kho dữ liệu và trình xử lý và lưu trữ dữ liệu;

trong đó, kho dữ liệu bao gồm kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản, kho dữ liệu bộ câu hỏi và trả lời, kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh, và tập ngữ liệu và từ vựng dùng để huấn luyện máy học;

trong đó, ứng dụng hỏi và đáp Web/App/API khi được thực thi bởi cụm các đơn vị xử lý trung tâm và các đơn vị xử lý đồ họa, nó sẽ khởi động ứng dụng hỏi và đáp Web/App/API; trong đó, ứng dụng hỏi và đáp Web/App/API bao gồm mô-đun truy xuất bộ nhớ truy cập nhanh, mô-đun so khớp câu hỏi/yêu cầu đơn giản, mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào, và mô-đun phân tích câu;

trong đó, mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào thực hiện phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào bao gồm các bước sau:

i) nhận một câu hỏi/yêu cầu đầu vào;

ii) thực hiện tính và so sánh độ tương đồng giữa câu hỏi/yêu cầu đầu vào với các mẫu câu hỏi/yêu cầu đơn giản trong kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản với ngưỡng tương đồng N được cho trước; trong đó, N là số thực và $N \in [0,1]$;

trong đó, bước ii) gồm các bước con như sau:

a) thực hiện tính độ tương đồng X về cấu trúc giữa câu hỏi/yêu cầu đầu vào với mẫu câu đơn giản; cụ thể là hai câu sẽ lần lượt được phân tích cú pháp, sau đó so sánh hai cây cú pháp với nhau bằng thuật toán SyMSS; trong đó X là số thực và $X \in [0,1]$;

b) thực hiện tính độ tương đồng Y về nội dung giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản; cụ thể là hai câu sẽ lần lượt được vector hóa (vectorization - biểu diễn các câu thành một cấu trúc vector) bằng mô hình ngôn ngữ, sau đó tính độ tương đồng cosin (cosine similarity) giữa hai vector này với nhau; trong đó Y là số thực và $Y \in [0,1]$;

c) thực hiện tính độ tương đồng Z giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản bằng cách tính trung bình các kết quả ở bước a) và bước b); cụ thể $Z = \frac{(X+Y)}{2}$ và $Z \in [0,1]$; trong đó, nếu Z có giá trị càng lớn, nghĩa là Z tiến gần càng gần 1, thì sự tương đồng/giống nhau về ngữ nghĩa giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản càng lớn;

- d) thực hiện so sánh độ tương đồng Z với ngưỡng tương đồng N cho trước;
- iii) xét xem có tìm thấy bất kỳ độ tương đồng Z nào vượt ngưỡng hay không, nghĩa là $Z \geq N$: nếu có thì thực hiện bước iv); ngược lại thì thực hiện bước viii);
- iv) lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh;
- v) xét xem bước iv) có truy vấn được câu hỏi đáp/trả lời hay không: nếu truy vấn được câu hỏi đáp/trả lời thì thực hiện bước vii); ngược lại thì thực hiện bước vi);
- vi) lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời;
- vii) thực hiện cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh; và kho dữ liệu bộ câu hỏi/yêu cầu và trả lời;
- viii) trả kết quả truy vấn;

trong đó, thiết bị điện tử thông minh và hệ thống thiết bị điện toán kết nối vào mạng dựa trên đám mây thông qua kênh truyền dẫn internet;

trong đó, thiết bị điện tử thông minh dùng để nhận câu hỏi/yêu cầu đầu vào là các lệnh/các yêu cầu từ người dùng;

trong đó, hệ thống thiết bị điện toán dùng để hỏi đáp/trả lời các câu hỏi/yêu cầu đầu vào từ người dùng thông qua thiết bị điện tử thông minh;

trong đó, thiết bị điện tử thông minh bao gồm điện thoại thông minh, máy tính bảng, máy tính để bàn, thiết bị đeo thông minh (smart wearable device), loa thông minh/trợ lý ảo, các thiết bị IoT, robot/xe tự hành;

trong đó, người dùng tương tác với thiết bị điện tử thông minh để ra lệnh/yêu cầu trên ứng dụng Web và/hoặc App thông qua bàn phím cơ/bàn phím ảo và/hoặc giao diện âm thanh (audio interface);

trong đó, câu hỏi đáp/trả lời của hệ thống điện toán bao gồm dạng văn bản, âm thanh và/hoặc ra lệnh cho thiết bị điện tử thông minh thực thi một hành động như mở/tắt nhạc, bật đèn, mở/tắt một chức năng nào đấy trong các thiết bị IoT; khác biệt ở chỗ câu hỏi đáp/trả lời của hệ thống thiết bị điện toán còn bao gồm lập trình nhúng vào trong các thiết bị điện tử thông minh bằng giao diện lập trình ứng dụng (“API” - Application Programming Interface).

2. Hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo điểm 1, còn bao gồm bên thứ ba, là những hàm API được cung ứng từ những hệ thống điện toán của những nhà cung cấp thông tin trực tuyến theo thời gian thực như thông tin tỷ giá, thời tiết, chăm sóc sức khỏe, danh bạ địa chỉ các cơ quan/tổ chức.

3. Hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo điểm 1, trong đó hệ thống thiết bị điện toán còn bao gồm:

giao diện mạng đầu vào/đầu ra (I/O);

cụm các đơn vị xử lý trung tâm và đơn vị xử lý đồ họa;

trình quản lý và thực thi các truy vấn SQL;

trình phát triển các ứng dụng phân tán và chuyển đổi cấu trúc dữ liệu;

và lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML);

trong đó, khi mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào hoặc mô-đun phân tích câu được kích hoạt, chúng ra lệnh cho lớp mô-đun trí tuệ nhân tạo và máy học tính độ tương đồng và/hoặc phân tích câu hỏi/yêu cầu đầu vào bằng ít nhất một mô hình ngôn ngữ (language model) và thực hiện phân lớp câu hỏi/yêu cầu đầu vào để hồi đáp/trả lời cho người dùng dựa vào việc truy xuất/truy vấn trung tâm dữ liệu ngôn ngữ.

4. Hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng theo điểm 3, trong đó giao diện mạng đầu vào/đầu ra (I/O) được sử dụng là công nghệ Cloudflare, trình quản lý và thực thi các truy vấn SQL sử dụng công nghệ Hadoop và Hive, trình phát triển các ứng dụng phân tán và chuyển đổi cấu trúc dữ liệu sử dụng công nghệ Srping cloud và Rest API, trình xử lý và lưu trữ dữ liệu sử dụng công nghệ HDFS và Redis.

5. Hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng theo điểm 3, trong đó lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML) còn bao gồm mô hình ngôn ngữ, thuật toán SyMSS, thuật toán máy vector hỗ trợ (SVM - Support Vector Machine), thuật toán trường ngẫu nhiên có điều kiện (CRF - Conditional Random Field).

6. Phương pháp xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng bao gồm các bước:

i) cung ứng một hệ thống xử lý ngôn ngữ tự nhiên dùng để hồi đáp/trả lời các yêu cầu từ người dùng;

ii) nhận một câu hỏi/yêu cầu đầu vào từ người dùng thông qua thiết bị điện tử thông minh; trong đó, câu hỏi/yêu cầu đầu vào từ người dùng là lệnh/các yêu cầu gửi đến thiết bị điện tử thông minh trên ứng dụng Web và/hoặc App thông qua bàn phím cơ/bàn phím ảo và/hoặc giao diện âm thanh (audio interface);

iii) mô-đun truy xuất bộ nhớ truy cập nhanh thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh (cache);

trong đó, bước iii) còn được thực thi từ trình xử lý và lưu trữ dữ liệu;

trong đó, trình xử lý và lưu trữ dữ liệu được thực hiện bởi công nghệ Redis;

iv) xét xem việc truy vấn đã tìm thấy được câu hỏi đáp/trả lời: nếu truy vấn tìm thấy được câu hỏi đáp/trả lời thì thực hiện bước xi); ngược lại thì thực hiện bước v);

v) xét xem phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản đã được thực hiện hay chưa: nếu phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản đã thực thi thực hiện bước vii); ngược lại thì thực hiện bước vi);

vi) thực hiện phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản và quay lại bước iv);

trong đó, bước vi) do mô-đun so khớp câu hỏi/yêu cầu đơn giản thực hiện;

vii) xét xem phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu

hỏi/yêu cầu đầu vào đã được thực hiện hay chưa: nếu phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào đã thực thi thì thực hiện bước ix); ngược lại thì thực hiện bước viii);

viii) thực hiện phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào và quay lại bước iv); trong đó phương pháp tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào bao gồm các bước:

a) nhận một câu hỏi/yêu cầu đầu vào;

b) thực hiện tính và so sánh độ tương đồng giữa câu hỏi/yêu cầu đầu vào với các mẫu câu hỏi/yêu cầu đơn giản trong kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản với ngưỡng tương đồng N được cho trước; trong đó, N là số thực và $N \in [0,1]$;

trong đó, bước b) gồm các bước con như sau:

b1) thực hiện tính độ tương đồng X về cấu trúc giữa câu hỏi/yêu cầu đầu vào với mẫu câu đơn giản; cụ thể là hai câu sẽ lần lượt được phân tích cú pháp, sau đó so sánh hai cây cú pháp với nhau bằng thuật toán SyMSS; trong đó X là số thực và $X \in [0,1]$;

b2) thực hiện tính độ tương đồng Y về nội dung giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản; cụ thể là hai câu sẽ lần lượt được vector hóa (vectorization - biểu diễn các câu thành một cấu trúc vector) bằng mô hình ngôn ngữ, sau đó tính độ tương đồng cosin (cosine similarity) giữa hai vector này với nhau; trong đó Y là số thực và $Y \in [0,1]$;

b3) thực hiện tính độ tương đồng Z giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản bằng cách tính trung bình các kết quả ở bước a) và bước b); cụ thể $Z = \frac{(X+Y)}{2}$ và $Z \in [0,1]$; trong đó, nếu Z có giá trị càng lớn, nghĩa là Z tiến gần càng gần 1, thì sự tương đồng/giống nhau về ngữ nghĩa giữa câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản càng lớn;

b4) thực hiện so sánh độ tương đồng Z với ngưỡng tương đồng N cho trước;

c) xét xem có tìm thấy bất kỳ độ tương đồng Z nào vượt ngưỡng hay không, nghĩa là $Z \geq N$: nếu có thì thực hiện bước d); ngược lại thì thực hiện bước h);

d) lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh;

e) xét xem bước d) có truy vấn được câu hỏi đáp/trả lời hay không: nếu truy vấn được câu hỏi đáp/trả lời thì thực hiện bước g); ngược lại thì thực hiện bước f);

f) lấy mẫu câu hỏi/yêu cầu đơn giản có tính tương đồng với câu hỏi/yêu cầu đầu vào để truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời;

g) thực hiện cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh; và kho dữ liệu bộ câu hỏi/yêu cầu và trả lời;

h) trả kết quả truy vấn; trong đó, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh; trong đó, kết quả

truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu hỏi đáp/trả lời (true) hoặc không tìm thấy câu hỏi đáp/trả lời (false);

trong đó, bước viii) do mô-đun tính độ tương đồng giữa mẫu câu hỏi/yêu cầu đơn giản với câu hỏi/yêu cầu đầu vào thực hiện;

ix) xét xem phương pháp phân tích câu hỏi/yêu cầu đầu vào đã được thực hiện hay chưa: nếu phương pháp phân tích câu hỏi/yêu cầu đầu vào đã thực thi thì thực hiện bước xi); ngược lại thì thực hiện bước x);

x) thực hiện phương pháp phân tích câu hỏi/yêu cầu đầu vào và quay lại bước iv);

trong đó, bước x) do mô-đun phân tích câu thực hiện;

xi) trả kết quả truy vấn/hỏi đáp yêu cầu đến người dùng và/hoặc thiết bị điện tử thông minh; trong đó, kết quả truy vấn/hỏi đáp bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh;

trong đó, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu trả lời (true) hoặc không tìm thấy câu trả lời (false);

trong đó, tùy vào ngữ cảnh câu hỏi/yêu cầu đầu vào từ người dùng mà hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng sẽ hỏi đáp bằng văn bản, hình ảnh, âm thanh hoặc không đưa ra bất kỳ một hành vi nào hướng đến người dùng và/hoặc thiết bị điện tử thông minh;

xii) xét xem kết quả trả về ở bước xi) là một câu hỏi đáp/trả lời đã được tìm thấy/xác định thì thực hiện bước xiii); ngược lại thì kết thúc phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

xiii) thực hiện sắp xếp lại thứ tự dữ liệu ưu tiên trong bộ nhớ truy cập nhanh nhằm tối ưu hóa dung lượng bộ nhớ truy cập nhanh và nhằm tối ưu hóa thời gian truy vấn các câu hỏi đáp/trả lời các câu yêu cầu đầu vào của người dùng;

trong đó, bước xiii) sử dụng thuật toán LRU (Least Recently Used); đó là xóa các dữ liệu ít được dùng gần đây nhất trong bộ nhớ truy cập nhanh và thuật toán LRU được lập trình xử lý ở trình xử lý và lưu trữ dữ liệu.

7. Phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo điểm 6, trong đó phương pháp so khớp câu hỏi/yêu cầu đầu vào với mẫu câu hỏi/yêu cầu đơn giản bao gồm các bước:

i) nhận một câu hỏi/yêu cầu đầu vào từ phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

ii) truy vấn câu hỏi/yêu cầu đầu vào từ kho dữ liệu mẫu câu hỏi/yêu cầu đơn giản;

iii) xét xem dữ liệu truy vấn ở bước ii) có được tìm thấy: nếu truy vấn tìm thấy dữ liệu thì thực hiện bước iv); ngược lại thì thực hiện bước vi);

iv) truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời;

trong đó, mẫu câu hỏi/yêu cầu đơn giản là câu có một số ít các từ nhưng câu vẫn đủ ngữ/nghĩa;

trong đó, mẫu câu hỏi/yêu cầu đơn giản này người dùng có thể tự cập nhật và/hoặc là những

mẫu câu hỏi/yêu cầu đã được hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng thiết lập sẵn;

v) thực hiện cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh;

vi) trả kết quả truy vấn và quay lại phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

trong đó, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh;

trong đó, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu hỏi đáp/trả lời (true) hoặc không tìm thấy câu hỏi đáp/trả lời (false).

8. Phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng theo điểm 6, trong đó phương pháp phân tích câu đầu vào bao gồm các bước:

i) nhận một câu hỏi/yêu cầu đầu vào từ phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

ii) thực hiện phân loại câu hỏi/yêu cầu đầu vào và bao gồm các bước con như sau:

a) thực hiện trích xuất đặc trưng của câu hỏi/yêu cầu đầu vào;

trong đó, bước được thực hiện bằng mô hình ngôn ngữ, cụ thể là mô-đun phân tích câu ra lệnh cho lớp mô-đun trí tuệ nhân tạo và máy học thực hiện trích xuất đặc trưng của câu hỏi/yêu cầu đầu vào;

b) từ các đặc trưng được trích xuất ở bước a), mô-đun phân tích câu tiếp tục phân lớp/lựa chọn các đặc trưng được trích xuất ở bước a) để tìm kiếm câu hỏi đáp/trả lời chính xác nhất;

trong đó, bước b) được thực hiện bởi thuật toán máy vector hỗ trợ (SVM - Support Vector Machine), chúng được lưu trữ và thực thi ở lớp mô-đun trí tuệ nhân tạo và máy học (AI/ML);

iii) xét xem bước ii) có phân loại được câu hỏi/yêu cầu đầu vào: nếu có thì thực bước iv); ngược lại thì thực hiện bước ix);

iv) thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời dựa trên các đặc trưng đã phân loại được ở bước ii);

v) xét xem truy vấn câu hỏi đáp/trả lời ở bước iv) đã có đủ thông tin hay không: nếu có thì thực hiện bước ix); ngược lại thì thực hiện bước vi);

vi) thực hiện truy vấn lấy thêm các thông tin từ kho dữ liệu bộ câu hỏi và trả lời;

vii) thực hiện rút trích thông tin;

trong đó, việc rút trích thông tin sẽ được hệ thống xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng sử dụng kỹ thuật của máy học để trích xuất thông tin;

trong đó, việc rút trích thông tin còn bao gồm thuật toán trường ngẫu nhiên có điều kiện (CRF - Conditional Random Field);

viii) xét xem việc rút trích thông tin ở bước vii) có thành công hay không: nếu rút trích thành công, nghĩa là thông tin rút trích được đủ dùng để truy vấn câu hỏi đáp/trả lời thì

thực hiện bước xiii); ngược lại thì thực hiện bước ix);

ix) thực hiện truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh;

x) xét xem bước ix) có hay không truy vấn được câu hỏi đáp/trả lời: nếu truy vấn được câu hỏi đáp/trả lời thì đi đến thực hiện bước xiii); ngược lại thì thực hiện bước xi);

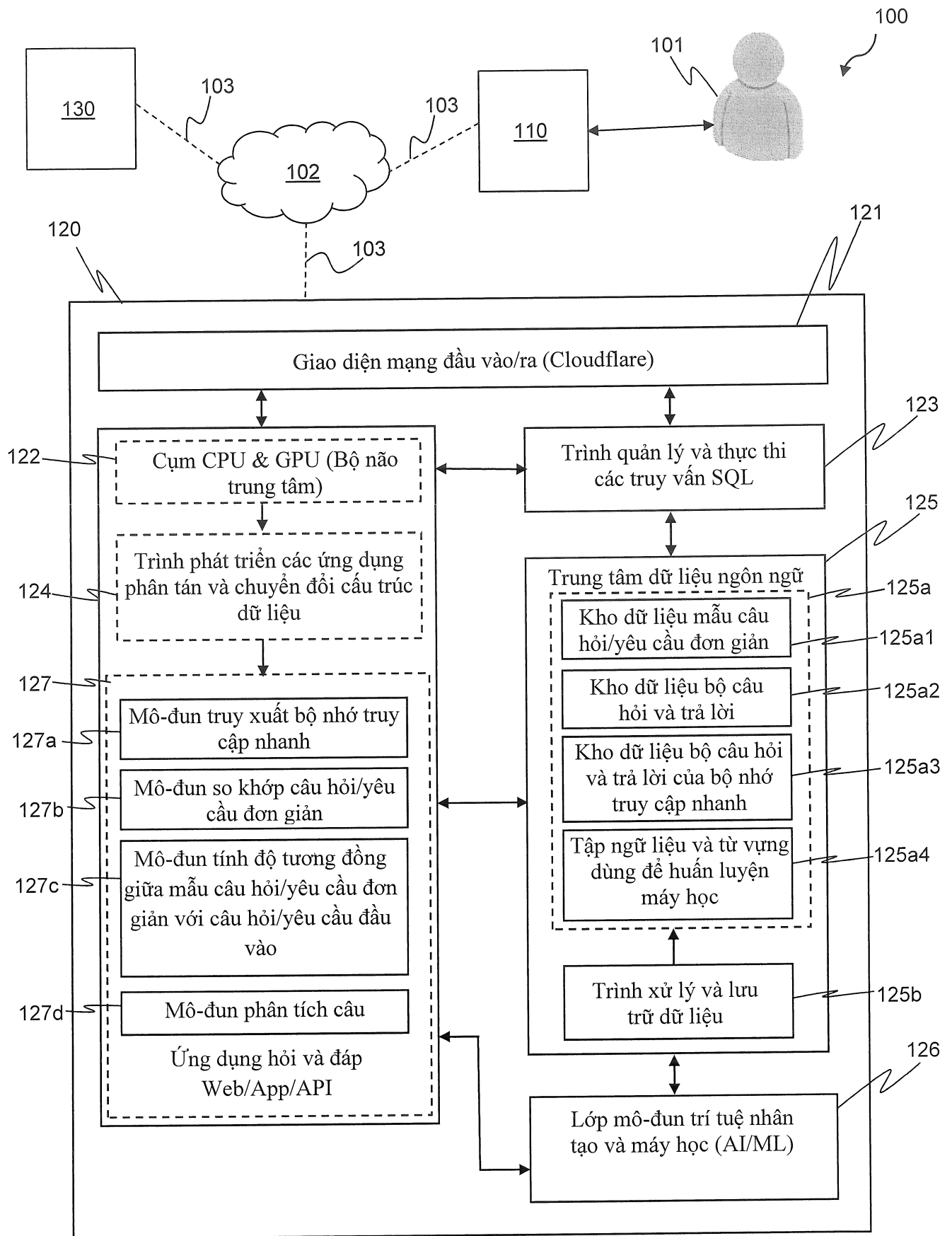
xi) truy vấn câu hỏi đáp/trả lời từ kho dữ liệu bộ câu hỏi và trả lời;

xii) cập nhật câu hỏi/yêu cầu đầu vào và câu hỏi đáp/trả lời vào trong kho dữ liệu bộ câu hỏi và trả lời của bộ nhớ truy cập nhanh; và kho dữ liệu bộ câu hỏi và trả lời;

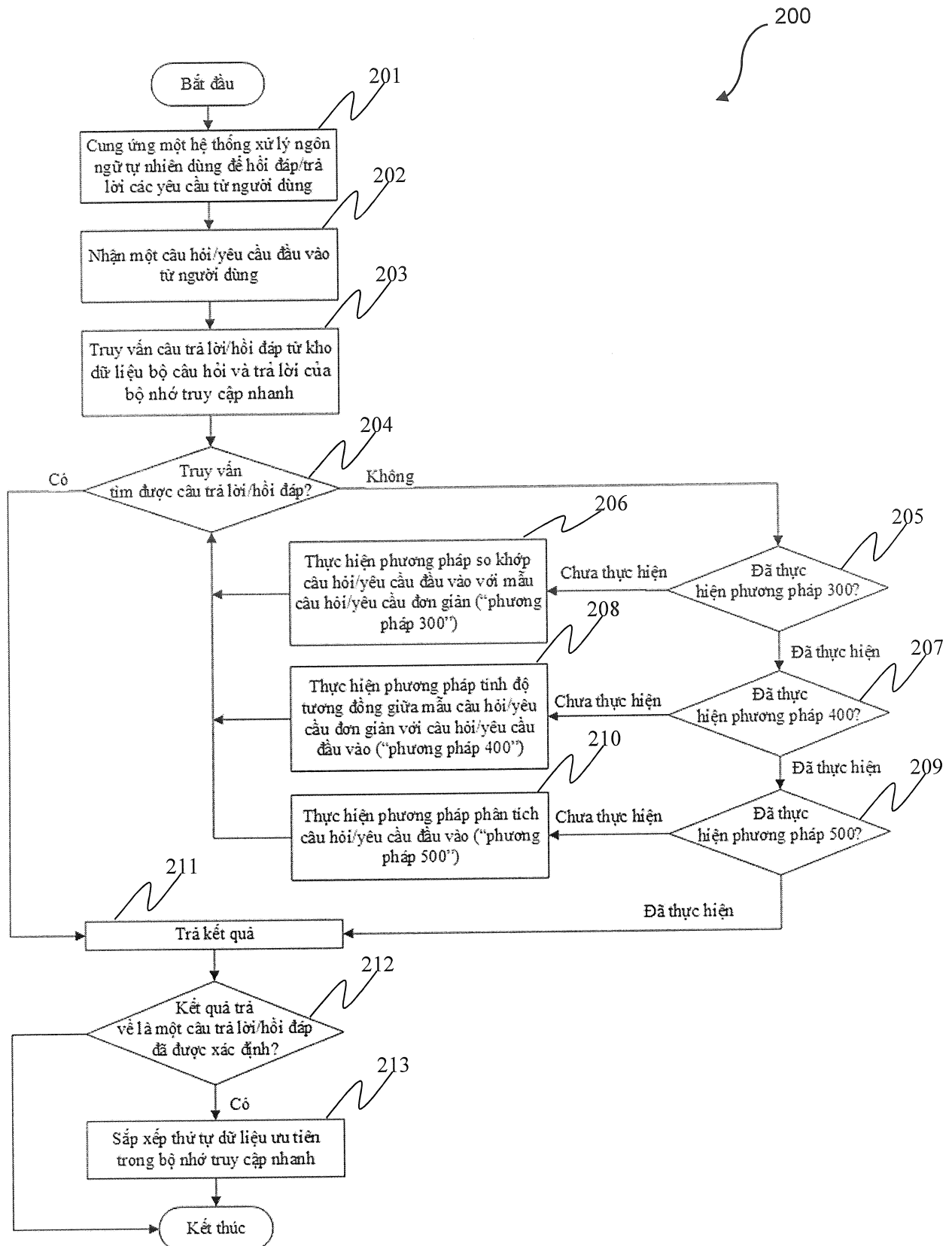
xiii) trả kết quả truy vấn và quay lại phương pháp xử lý ngôn ngữ tự nhiên dùng để hỏi đáp/trả lời các yêu cầu từ người dùng;

trong đó, kết quả truy vấn bao gồm dạng văn bản và/hoặc âm thanh và/hoặc một câu lệnh thực thi một hành vi cụ thể như mở/tắt nhạc, bật/tắt một chức năng cụ thể nào đấy trên thiết bị điện tử thông minh;

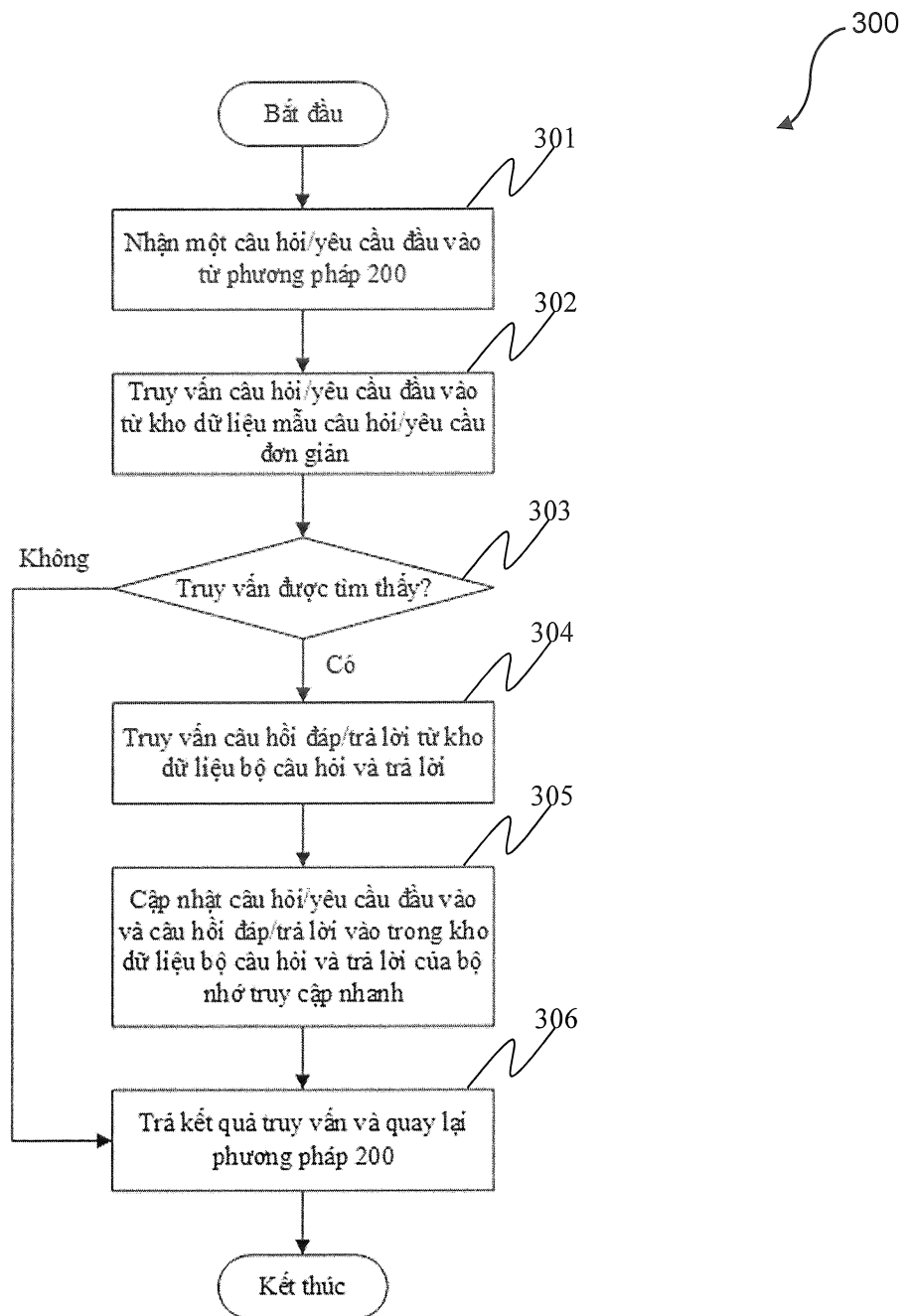
trong đó, kết quả truy vấn/hỏi đáp còn bao gồm hai trạng thái: truy vấn được câu hỏi đáp/trả lời (true) hoặc không tìm thấy câu hỏi đáp/trả lời (false).



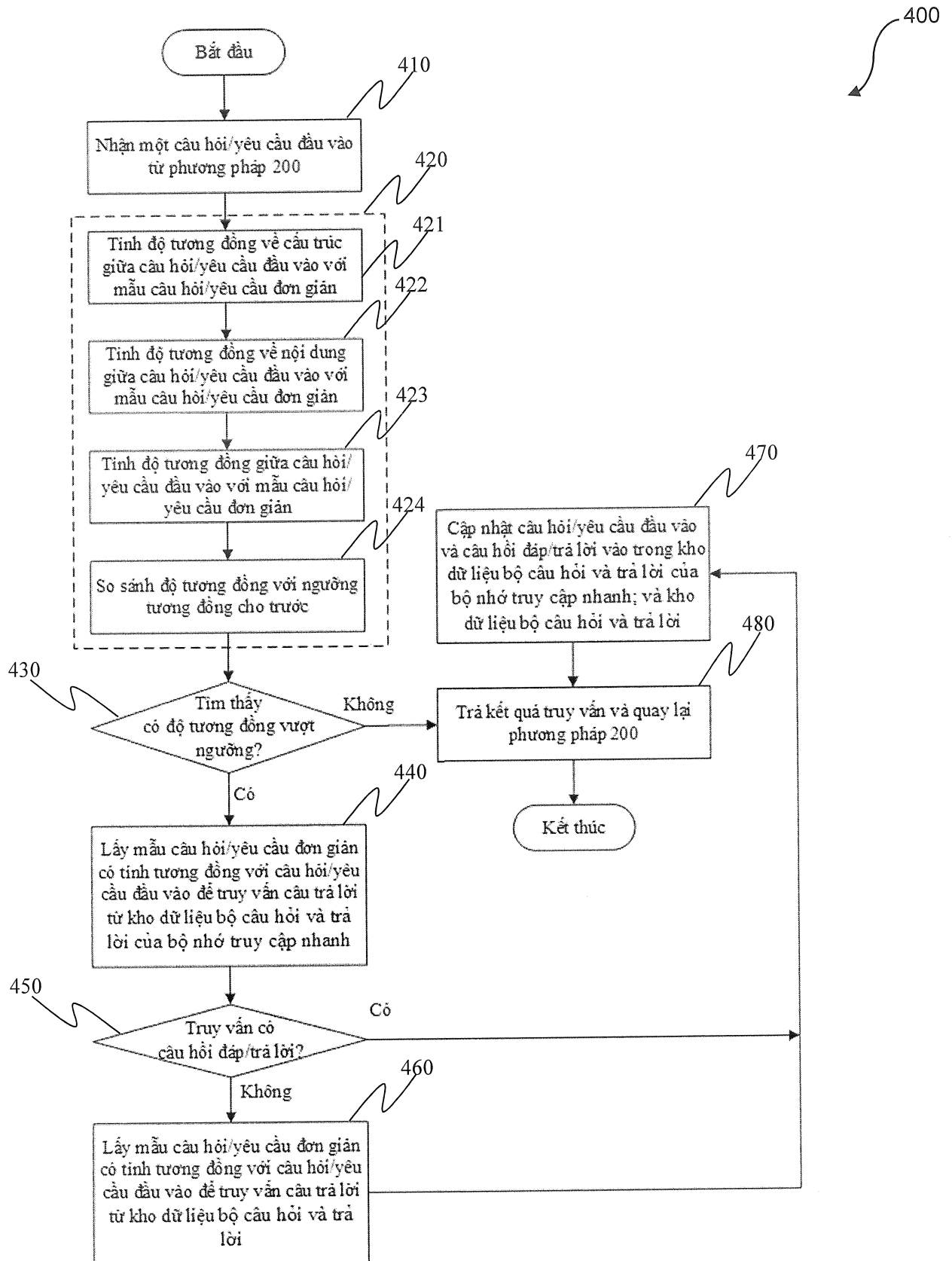
HÌNH 1



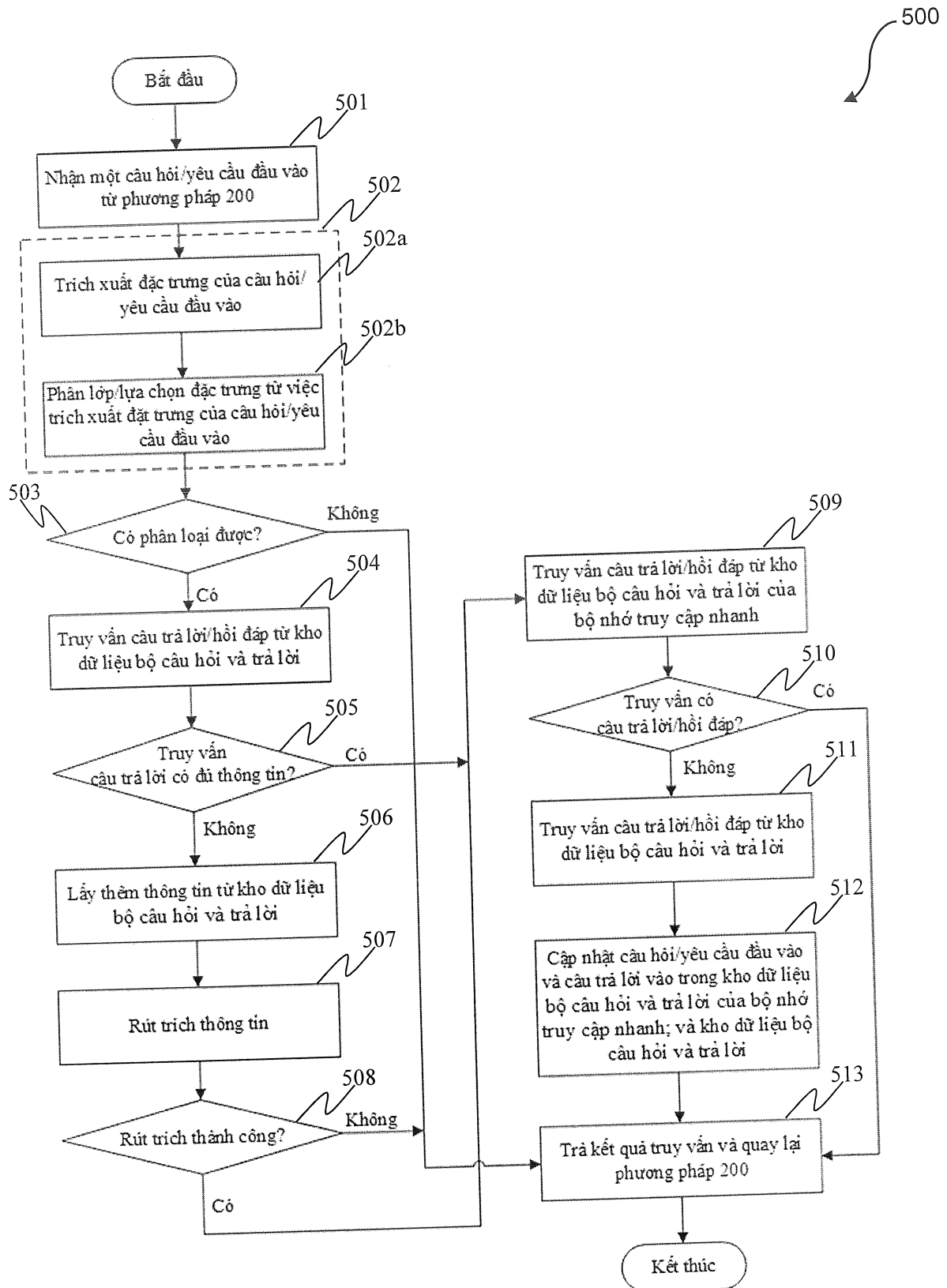
HÌNH 2



HÌNH 3



HÌNH 4



HÌNH 5