

HUMAN-COMPUTER INTERACTION METHOD AND SYSTEM BASED ON KNOWLEDGE GRAPH

RELATED APPLICATIONS

[0001] The present application is a national phase application of the international application PCT/CN2016/094905 filed August 12, 2016, which claims the benefit of the Chinese application CN201610532831.2 filed July 7, 2016, each of which is incorporated herein by reference in its entirety.

FIELD OF THE INVENTION

[0002] The present invention relates to the technical field of artificial intelligence, and particularly relates to a human-computer interaction method and system based on a knowledge graph.

BACKGROUND OF THE INVENTION

[0003] Multi-round conversations cannot be smoothly conducted by the existing human-computer interaction technology based on a knowledge graph, for example, a user firstly inputs: what is the temperature today in Shenzhen, a robot finishes an answer, and the user continuously inputs: how about tomorrow. Since information is incomplete in the question, the robot cannot identify the question, thereby causing a situation that the multi-round conversations cannot continue.

OBJECTS AND SUMMARY OF THE INVENTION

[0004] The present invention aims at providing a human-computer interaction method and system based on a knowledge graph, so that the problem that conversations cannot continue by the existing human-computer interaction technology based on the knowledge graph when information of a question raised by a user is incomplete is solved.

[0005] In order to solve the technical problem, the present invention adopts the technical scheme as follows:

[0006] the human-computer interaction method based on the knowledge graph comprises:

[0007] acquiring current multi-modal input records of a user, wherein the current multi-modal input records comprise input data and input time;

[0008] carrying out intention classification on the current multi-modal input records, storing the current multi-modal input records into a corresponding intention repository and carrying out, by the intention repository, feature classification on multi-modal input records by N feature tags;

[0009] carrying out feature classification on the current multi-modal input records in the intention repository, determining feature tags and a number M of the current multi-modal input records and putting the current multi-modal input records and feature values thereof under the corresponding feature tags;

[0010] comparing M and N; determining current feature classification information according to the current multi-modal input records when M is equal to N and making a response to the user based on the knowledge graph according to the current feature classification information;

[0011] determining to-be-complemented features of the current multi-modal input records and judging complementing feasibilities of the to-be-complemented features when M is less than N; when a judgment result is that the features can be complemented, complementing feature values of the to-be-complemented features by adopting an automatic complementing manner to obtain updated current multi-modal input records, determining current feature classification information according to the updated current multi-modal input records and making a response to the user based on the knowledge graph according to the current feature classification information; and when the judgment result is that the features cannot be complemented, pushing a question about the to-be-complemented features to the user, acquiring feature values of the to-be-complemented features to obtain updated current multi-modal input records, determining current feature classification information according to the updated current multi-modal input records and making a response to the user based on the knowledge graph according to the current feature classification information.

[0012] Based on this, further, the input data comprise at least one of characters, voices, cartoons, expressions, actions, gestures, pictures and videos.

[0013] Based on any of above embodiments, further, the automatic complementing manner comprises a historical multi-modal input record complementing manner, a knowledge repository complementing manner, a user attribute repository complementing manner and a third-party acquired data complementing manner.

[0014] Based on this, further, the historical multi-modal input record complementing manner is specifically described as follows:

[0015] inquiring historical multi-modal input records with a last input time in classification results under tags of the to-be-complemented features in the intention repository and acquiring feature values of to-be-complemented features in the historical multi-modal input records as the feature values of the to-be-complemented features in the current multi-modal input records.

[0016] Based on any of the above embodiments, further, the method further comprises:

[0017] generating response records and storing the response records into a response repository.

[0018] The human-computer interaction system based on the knowledge graph comprises:

[0019] an input module configured to acquire current multi-modal input records of a user, wherein the current multi-modal input records comprise input data and input time;

[0020] an intention identification module configured to carry out intention classification on the current multi-modal input records and store the current multi-modal input records into a corresponding intention repository, wherein the intention repository is configured to carry out feature classification on multi-modal input records by N feature tags;

[0021] a feature classification module configured to carry out feature classification on the current multi-modal input records in the intention repository, determine feature tags and a number M of the current multi-modal input records and put the current multi-modal input records and feature values thereof under the corresponding feature tags; and

[0022] an intention response module configured to compare M and N, wherein: when M is equal to N, current feature classification information is determined according to the current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; when M is less than N, to-be-complemented features of the current multi-modal input records are determined, and complementing feasibilities of the to-be-complemented features are judged; when a judgment result is that the features can be complemented, feature values of the to-be-complemented features are complemented by adopting an automatic complementing

manner to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; and when the judgment result is that the features cannot be complemented, a question about the to-be-complemented features is pushed to the user, feature values of the to-be-complemented features are acquired to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information.

[0023] Based on this, further, the input data comprise at least one of characters, voices, cartoons, expressions, actions, gestures, pictures and videos.

[0024] Based on any of above embodiments, further, the automatic complementing manner comprises a historical multi-modal input record complementing manner, a knowledge repository complementing manner, a user attribute repository complementing manner and a third-party acquired data complementing manner.

[0025] Based on this, further, the historical multi-modal input record complementing manner is specifically described as follows:

[0026] inquiring historical multi-modal input records with a last input time in classification results under tags of the to-be-complemented features in the intention repository and acquiring feature values of to-be-complemented features in the historical multi-modal input records as the feature values of the to-be-complemented features in the current multi-modal input records.

[0027] Based on any of the above embodiments, further, the system further comprises:

[0028] a response storage module configured to generate response records and store the response records into a response repository.

[0029] The present invention has the beneficial effects that:

[0030] the present invention provides the human-computer interaction method and system based on the knowledge graph; when features used for inquiring information are incomplete in the multi-modal input records of the user, the to-be-complemented features of the current multi-modal input records are determined, and the complementing feasibilities of

the to-be-complemented features are judged; when the judgment result is that the features can be complemented, the feature values of the to-be-complemented features are complemented by adopting an automatic complementing manner, so that features required when the user inquires information are automatically complemented, and the response is made to the user based on the knowledge graph, therefore, multi-round conversations can be realized; and when the judgment result is that the features cannot be complemented, the question about the to-be-complemented features is pushed to the user, and the feature values of the to-be-complemented features are acquired, so that features required when the user inquires information are complemented, and the response is made to the user based on the knowledge graph, therefore, the multi-round conversations can be realized.

BRIEF DESCRIPTION OF FIGURES

[0031] The present invention is further described below in combination with figures and embodiments.

[0032] Fig. 1 is a flow chart of a human-computer interaction method based on a knowledge graph provided by an embodiment of the present invention; and

[0033] Fig. 2 is a structural schematic diagram of a human-computer interaction system based on a knowledge graph provided by an embodiment of the present invention.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0034] In order that the purpose, the technical scheme and the advantages of the present invention can be more clearly understood, the present invention is further described below in detail in combination with drawings and embodiments. It should be understood that the specific embodiments described herein are only used for explaining the present invention, not limiting the present invention.

[0035] Specific embodiment I

[0036] As shown in Fig. 1, an embodiment of present invention provides a human-computer interaction method based on a knowledge graph, comprising:

[0037] step S101, acquiring current multi-modal input records of a user, wherein the current multi-modal input records comprise input data and input time;

[0038] step S102, carrying out intention classification on the current multi-modal input records, storing the current multi-modal input records into a corresponding intention repository and carrying out, by the intention repository, feature classification on multi-modal input records by N feature tags;

[0039] step S103, carrying out feature classification on the current multi-modal input records in the intention repository, determining feature tags and a number M of the current multi-modal input records and putting the current multi-modal input records and feature values thereof under the corresponding feature tags; and

[0040] step S104, comparing M and N, wherein when M is equal to N, current feature classification information is determined according to the current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; when M is less than N, to-be-complemented features of the current multi-modal input records are determined, and complementing feasibilities of the to-be-complemented features are judged; when a judgment result is that the features can be complemented, feature values of the to-be-complemented features are complemented by adopting an automatic complementing manner to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; and when the judgment result is that the features cannot be complemented, a question about the to-be-complemented features is pushed to the user, feature values of the to-be-complemented features are acquired to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information.

[0041] A system can judge whether a question can be complemented before the features are complemented; not all situations can be completed by complementing, and the system can automatically complement the question only when the question can be complemented; if the system cannot automatically complement the question, the system asks

a customer in turn and pushes a question to the user, for example, 1, for a question that ‘what is the weather in Shenzhen?’, the question can be automatically complemented as ‘what is the weather today in Shenzhen?’ according to a user habit in a user attribute repository; and 2, for a question that ‘do you have a railway ticket for this weekend?’, the system inquires the user ‘where is the destination?’ at the moment.

[0042] According to the embodiment of the present invention, when features used for inquiring information are incomplete in the multi-modal input records of the user, the to-be-complemented features of the current multi-modal input records are determined, and the complementing feasibilities of the to-be-complemented features are judged; when the judgment result is that the features can be complemented, the feature values of the to-be-complemented features are complemented by adopting an automatic complementing manner, so that features required when the user inquires information are automatically complemented, and the response is made to the user based on the knowledge graph, therefore, multi-round conversations can be realized; and when the judgment result is that the features cannot be complemented, the question about the to-be-complemented features is pushed to the user, and the feature values of the to-be-complemented features are acquired, so that features required when the user inquires information are complemented, and the response is made to the user based on the knowledge graph, therefore, the multi-round conversations can be realized.

[0043] In the embodiment of the present invention, a form of the input data in the step S101 is not limited; and preferably, the input data can be at least one of characters, voices, cartoons, expressions, actions, gestures, pictures and videos.

[0044] In the embodiment of the present invention, the automatic complementing manner in the step S104 is not limited; and based on any of above embodiments, preferably, the automatic complementing manner can comprise a historical multi-modal input record complementing manner, a knowledge repository complementing manner, a user attribute repository complementing manner and a third-party acquired data complementing manner. The complementing manner can be the complementing based on the above contents and can also be complementing based on other methods, for example, for a question that ‘what is the weather today?’, the system can acquire a positioned place by GPS (Global Position System) positioning to complement the question. Or, when the user asks a question that ‘is there

anything delicious nearby?’, the system can complement the question as ‘is there delicious hot pot nearby?’ by a personal attribute repository of the user, such as the preference of the user (if the user prefers to eat hot pot).

[0045] Based on this, preferably, the historical multi-modal input record complementing manner can be specifically described as follows: inquiring historical multi-modal input records with a last input time in classification results under tags of the to-be-complemented features in the intention repository and acquiring feature values of to-be-complemented features in the historical multi-modal input records as the feature values of the to-be-complemented features in the current multi-modal input records. According to the configuration of the system, a conversation can comprise one or more of multi-modal input records, and current intentions can be classified by the historical multi-modal input records. For example, a first multi-modal input record of the user is: what is the temperature today in Shenzhen, a robot finishes an answer, and the user continuously inputs: how about tomorrow. The question is incomplete, therefore, at the moment, the robot classifies the multi-modal input record as weather inquiry according to the first multi-modal input record, and a weather inquiry intention repository is configured to carry out classification on the multi-modal input record by two features tags of time and place; in the weather inquiry intention repository, feature classification is carried out on a second input data that ‘how about tomorrow?’, a feature tag of the second input data is determined as time, and the quantity of the feature tag is one; and since one is less than two, and the place feature tag of does not exist, the place needs to be complemented for the second input data, the multi-modal input record that ‘what is the temperature today in Shenzhen?’ with the last input time in a classification result under the place tag is inquired in the weather inquiry intention repository, and a place feature value ‘Shenzhen’ is acquired as a to-be-complemented place feature value of the multi-modal input record, therefore, the system can inquire weather conditions of tomorrow in Shenzhen and send the weather conditions to the user according to the complemented place feature value ‘Shenzhen’ and the time feature value ‘tomorrow’ in the original input data.

[0046] Based on any of the above embodiments, preferably, in the embodiment of the present invention, after the step S104, the method further comprises: generating response

records and storing the response records into a response repository. Doing like this can facilitate the user to inquire historical response records.

[0047] In the embodiment of the present invention, the contents of the response records are not limited, and preferably, the response records can comprise state classification information. Doing like this can facilitate the user to inquire state classification information corresponding to the responses from the historical response records.

[0048] In the embodiment of the present invention, a storage position of the intention repository is not limited, and preferably, the intention repository can be stored into a cloud server. The intention repository can be stored into the cloud server, so that intention repositories of users among different equipment can be shared.

[0049] Specific embodiment II

[0050] As shown in Fig. 2, an embodiment of the present invention provides a human-computer interaction system based on a knowledge graph, comprising:

[0051] an input module configured to acquire current multi-modal input records of a user, wherein the current multi-modal input records comprise input data and input time;

[0052] an intention identification module configured to carry out intention classification on the current multi-modal input records and store the current multi-modal input records into a corresponding intention repository, wherein the intention repository is configured to carry out feature classification on multi-modal input records by N feature tags;

[0053] a feature classification module configured to carry out feature classification on the current multi-modal input records in the intention repository, determine feature tags and a number M of the current multi-modal input records and put the current multi-modal input records and feature values thereof under the corresponding feature tags; and

[0054] an intention response module, configured to compare M and N, wherein: when M is equal to N, current feature classification information is determined according to the current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; when M is less than N, to-be-complemented features of the current multi-modal input records are determined, and complementing feasibilities of the to-be-complemented features are judged; when a judgment result is that the features can be complemented, feature values of the to-be-complemented features are complemented by adopting an automatic complementing

manner to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; and when the judgment result is that the features cannot be complemented, a question about the to-be-complemented features is pushed to the user, feature values of the to-be-complemented features are acquired to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information.

[0055] According to the embodiment of the present invention, when features used for inquiring information are incomplete in the multi-modal input records of the user, the to-be-complemented features of the current multi-modal input records are determined, and the complementing feasibilities of the to-be-complemented features are judged; when the judgment result is that the features can be complemented, the feature values of the to-be-complemented features are complemented by adopting an automatic complementing manner, so that features required when the user inquires information are automatically complemented, and the response is made to the user based on the knowledge graph, therefore, multi-round conversations can be realized; and when the judgment result is that the features cannot be complemented, the question about the to-be-complemented features is pushed to the user, and the feature values of the to-be-complemented features are acquired, so that features required when the user inquires information are complemented, and the response is made to the user based on the knowledge graph, therefore, the multi-round conversations can be realized.

[0056] In the embodiment of the present invention, a form of the input data is not limited; and preferably, the input data can be at least one of characters, voices, cartoons, expressions, actions, gestures, pictures and videos.

[0057] In the embodiment of the present invention, the automatic complementing manner is not limited; and based on any of above embodiments, preferably, the automatic complementing manner can comprise a historical multi-modal input record complementing manner, a knowledge repository complementing manner, a user attribute repository complementing manner and a third-party acquired data complementing manner. The

complementing manner can be the complementing based on the above contents and can also be complementing based on other methods, for example, for a question that ‘what is the weather today?’, the system can acquire a positioned place by GPS positioning to complement the question. Or, when the user asks a question that ‘is there anything delicious nearby?’, the system can complement the question a ‘is there delicious hot pot nearby?’ by a personal attribute repository of the user, such as the preference of the user (if the user prefers to eat hot pot).

[0058] Based on this, preferably, the historical multi-modal input record complementing manner can be specifically described as follows: inquiring historical multi-modal input records with the last input time in classification results under tags of the to-be-complemented features in the intention repository and acquiring feature values of to-be-complemented features in the historical multi-modal input records as the feature values of the to-be-complemented features in the current multi-modal input records. According to the configuration of the system, a conversation can comprise one or more of multi-modal input records, and current intentions can be classified by the historical multi-modal input records. For example, a first multi-modal input record of the user is: what is the temperature today in Shenzhen, a robot finishes an answer, and the user continuously inputs: how about tomorrow. The question is incomplete, therefore, at the moment, the robot classifies the multi-modal input record as weather inquiry according to the first multi-modal input record, and a weather inquiry intention repository is configured to carry out classification on the multi-modal input record by two features tags of time and place; in the weather inquiry intention repository, feature classification is carried out on a second input data that ‘how about tomorrow?’, a feature tag of the second input data is determined as time, and the quantity of the feature tag is one; and since one is less than two, and the place feature tag of does not exist, the place needs to be complemented for the second input data, the multi-modal input record that ‘what is the temperature today in Shenzhen?’ with the last input time in a classification result under the place tag is inquired in the weather inquiry intention repository, and a place feature value ‘Shenzhen’ is acquired as a to-be-complemented place feature value of the multi-modal input record, therefore, the system can inquire weather conditions of tomorrow in Shenzhen and send the weather

conditions to the user according to the complemented place feature value ‘Shenzhen’ and the time feature value ‘tomorrow’ in the original input data.

[0059] Based on any of the above embodiments, preferably, in the embodiment of the present invention, the system further comprises: a response storage module configured to generate response records and store the response records into a response repository. Doing like this can facilitate the user to inquire historical response records.

[0060] In the embodiment of the present invention, the contents of the response records are not limited, and preferably, the response records can comprise state classification information. Doing like this can facilitate the user to inquire state classification information corresponding to the responses from the historical response records.

[0061] In the embodiment of the present invention, a storage position of the intention repository is not limited, and preferably, the intention repository can be stored into a cloud server. The intention repository can be stored into the cloud server, so that intention repositories of users among different equipment can be shared.

[0062] Even if the present invention is described to a certain degree, obviously, appropriate changes of all conditions can be performed under the condition of not departing from the spirit and the scope of the present invention. It can be understood that the present invention is not limited to the described implementation schemes, belongs to the scope of the claims and comprises equivalent replacements of all described factors.

WHAT IS CLAIMED IS:

1. A human-computer interaction method based on a knowledge graph, comprising:
 - acquiring current multi-modal input records of a user, wherein the current multi-modal input records comprise input data and input time;
 - carrying out intention classification on the current multi-modal input records, storing the current multi-modal input records into a corresponding intention repository and carrying out, by the intention repository, feature classification on multi-modal input records by N feature tags;
 - carrying out feature classification on the current multi-modal input records in the intention repository, determining feature tags and a number M of the current multi-modal input records and putting the current multi-modal input records and feature values thereof under the corresponding feature tags;
 - comparing M and N; determining current feature classification information according to the current multi-modal input records when M is equal to N and making a response to the user based on the knowledge graph according to the current feature classification information;
 - determining to-be-complemented features of the current multi-modal input records and judging complementing feasibilities of the to-be-complemented features when M is less than N; when a judgment result is that the features can be complemented, complementing feature values of the to-be-complemented features by adopting an automatic complementing manner to obtain updated current multi-modal input records, determining current feature classification information according to the updated current multi-modal input records and making a response to the user based on the knowledge graph according to the current feature classification information; and when the judgment result is that the features cannot be complemented, pushing a question about the to-be-complemented features to the user, acquiring feature values of the to-be-complemented features to obtain updated current multi-modal input records, determining current feature classification information according to the updated current multi-modal input records and making a response to the user based on the knowledge graph according to the current feature classification information.

2. The human-computer interaction method based on the knowledge graph according to claim 1, wherein the input data comprise at least one of characters, voices, cartoons, expressions, actions, gestures, pictures and videos.
3. The human-computer interaction method based on the knowledge graph according to claims 1 or 2, wherein the automatic complementing manner comprises a historical multi-modal input record complementing manner, a knowledge repository complementing manner, a user attribute repository complementing manner and a third-party acquired data complementing manner.
4. The human-computer interaction method based on the knowledge graph according to claim 3, wherein the historical multi-modal input record complementing manner is specifically described as follows:
 - inquiring historical multi-modal input records with a last input time in classification results under tags of the to-be-complemented features in the intention repository; and
 - acquiring feature values of to-be-complemented features in the historical multi-modal input records as the feature values of the to-be-complemented features in the current multi-modal input records.
5. The human-computer interaction method based on the knowledge graph according to claims 1 or 2, further comprising generating response records and storing the response records into a response repository.
6. A human-computer interaction system based on a knowledge graph, comprising:
 - an input module configured to acquire current multi-modal input records of a user, wherein the current multi-modal input records comprise input data and input time;
 - an intention identification module configured to carry out intention classification on the current multi-modal input records and store the current multi-modal input records into a corresponding intention repository, wherein the intention repository is configured to carry out feature classification on multi-modal input records by N feature tags;
 - a feature classification module configured to carry out feature classification on the current multi-modal input records in the intention repository, determine feature

tags and a number M of the current multi-modal input records and put the current multi-modal input records and feature values thereof under the corresponding feature tags; and

an intention response module configured to compare M and N , wherein: when M is equal to N , current feature classification information is determined according to the current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; when M is less than N , to-be-complemented features of the current multi-modal input records are determined, and complementing feasibilities of the to-be-complemented features are judged; when a judgment result is that the features can be complemented, feature values of the to-be-complemented features are complemented by adopting an automatic complementing manner to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information; and when the judgment result is that the features cannot be complemented, a question about the to-be-complemented features is pushed to the user, feature values of the to-be-complemented features are acquired to obtain updated current multi-modal input records, current feature classification information is determined according to the updated current multi-modal input records, and a response is made to the user based on the knowledge graph according to the current feature classification information.

7. The human-computer interaction system based on the knowledge graph according to claim 6, wherein the input data comprise at least one of characters, voices, cartoons, expressions, actions, gestures, pictures and videos.
8. The human-computer interaction system based on the knowledge graph according to claims 6 or 7, wherein the automatic complementing manner comprises a historical multi-modal input record complementing manner, a knowledge repository complementing manner, a user attribute repository complementing manner and a third-party acquired data complementing manner.

9. The human-computer interaction system based on the knowledge graph according to claim 8, wherein the historical multi-modal input record complementing manner is specifically described as follows:
 - inquiring historical multi-modal input records with a last input time in classification results under tags of the to-be-complemented features in the intention repository; and
 - acquiring feature values of to-be-complemented features in the historical multi-modal input records as the feature values of the to-be-complemented features in the current multi-modal input records.
10. The human-computer interaction system based on the knowledge graph according to claims 6 or 7, further comprising a response storage module configured to generate response records and store the response records into a response repository.