DESCRIPTION

TITLE OF THE INVENTION: INTELLIGENT DIALOGUE METHOD AND APPARATUS, ELECTRONIC DEVICE AND STORAGE MEDIUM

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] The present disclosure claims the priority to the Chinese patent application with the filing No. 202510195811X filed with the Chinese Patent Office on February 21, 2025, and entitled "INTELLIGENT DIALOGUE METHOD AND APPARATUS, ELECTRONIC DEVICE AND STORAGE MEDIUM", the contents of which are incorporated herein by reference in entirety.

TECHNICAL FIELD

[0002] The present disclosure relates to the technical field of artificial intelligence, and specifically to an intelligent dialogue method and apparatus, an electronic device and a storage medium.

BACKGROUND ART

[0003] With the high-speed development of artificial intelligence technology and the remarkable performance of large language models in the dialogue field in recent years, dialogue agents (chatbots) are increasingly appearing in our daily life. For dialogue agents used by enterprises, they are mostly applied to scenes such as customer service and sales promotion. In these scenes, controllability and proactivity are critical attributes of dialogue agents. Controllability requires dialogue agents to follow SOP (Standard Operating Procedure). Proactivity requires dialogue agents to be capable of guiding dialogue toward an objective even when the user is uncooperative.

[0004] At present, machine learning techniques are often employed to use chatbots to conduct dialogues. However, such method relies on a large quantity of annotated data and predefined rules, while the rules often exhibit limitations, making it difficult for chatbots to deal with complex dialogue scenes. In addition, there is also a method for carrying out dialogues using a large language model. Although this method can cope with various dialogue scenes, it fails to avoid model hallucination, and is prone to causing the dialogue to deviate from an objective.

SUMMARY

[0005] In view of this, the present disclosure aims at providing an intelligent dialogue method and apparatus, an electronic device and a storage medium.

[0006] In order to achieve the above objective, technical solutions employed in the embodiments of the present disclosure are as follows.

[0007] In the first aspect, the present disclosure provides an intelligent dialogue method, including steps of:

[0008] in a process of conducting a dialogue with a user, obtaining each candidate dialogue node of next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree;

[0009] taking all the candidate dialogue nodes as dialogue constraint information, and determining a target dialogue node of next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information; and

[0010] generating a reply content of the next step using the large language model based on the target dialogue node, and feeding the reply content back to the user.

[0011] In an optional embodiment, each dialogue node includes dialogue action information; and

[0012] the step of determining a target dialogue node of next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information includes:

[0013] generating an action acquisition instruction based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information;

[0014] inputting the action acquisition instruction into the large language model, and generating reference dialogue action information using the large language model;

[0015] matching the dialogue action information about each of the candidate dialogue nodes with the reference dialogue action information;

[0016] if there is a candidate dialogue node matched with the reference dialogue action information, taking the candidate dialogue node as the target dialogue node; and

[0017] if there is no candidate dialogue node matched with the reference dialogue action information, generating a provisional dialogue node and dialogue action information thereabout using the large language model, and taking the provisional dialogue node as the target dialogue node, where the dialogue action information about the provisional dialogue node denotes guiding the dialogue to a dialogue node in the dialogue SOP tree.

[0018] In an optional embodiment, the step of matching the dialogue action information about each of the candidate dialogue nodes with the reference dialogue action information includes:

[0019] calculating similarity between the dialogue action information about each of the candidate dialogue nodes and the reference dialogue action information, so as to obtain corresponding similarity of each of the candidate dialogue nodes;

[0020] if there is a candidate dialogue node with the similarity reaching a predefined first threshold, taking a candidate dialogue node corresponding to a greatest similarity as the candidate dialogue node matched with the reference dialogue action information; and

[0021] if there is no candidate dialogue node with similarity reaching the first threshold, determining that there is no candidate dialogue node matched with the reference dialogue action information.

[0022] In an optional embodiment, the step of generating a reply content of the next step using the large language model based on the target dialogue node includes:

[0023] if the target dialogue node is a provisional dialogue node, generating a first script acquisition instruction based on dialogue action information about the provisional dialogue node; and

[0024] inputting the first script acquisition instruction into the large language model, generating a guidance script using the large language model, and taking the guidance script as a reply content of next step.

[0025] In an optional embodiment, each dialogue node further includes a plurality of standard scripts;

[0026] the step of generating a reply content of the next step using the large language model based on the target dialogue node includes:

[0027] when the target dialogue node is the candidate dialogue node, acquiring each standard script of the candidate dialogue node, so as to obtain each candidate script;

[0028] taking all candidate scripts as script constraint information, and generating a second script acquisition instruction based on the script constraint information;

[0029] inputting the second script acquisition instruction into the large language model, and generating a target script using the large language model; and

[0030] selecting a reply content of next step from the target script and all the candidate scripts.

[0031] In an optional embodiment, the step of selecting a reply content of next step from the target script and all the candidate scripts includes:

[0032] calculating a similarity between the target script and each of the candidate scripts;

[0033] if the similarity between the target script and at least one of the candidate scripts reaches a predefined second threshold, taking the target script as the reply content of the next step; and

[0034] if the similarities between the target script and all the candidate scripts fail to reach the second threshold, taking a candidate script with a greatest similarity as the reply content of the next step.

[0035] In the second aspect, the present disclosure provides an intelligent dialogue apparatus, including:

[0036] a dialogue planning module, configured to, in a process of conducting a dialogue with a user, obtain each candidate dialogue node of next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree;

[0037] take all the candidate dialogue nodes as dialogue constraint information, and determine a target dialogue node of next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information; and

[0038] a dialogue generation module, configured to generate a reply content of next step using the large language model based on the target dialogue node, and feed the reply content back to the user.

[0039] In an optional embodiment, each dialogue node includes dialogue action information; and the dialogue planning module is further configured to:

[0040] generate an action acquisition instruction based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information;

[0041] input the action acquisition instruction into the large language model, and generate reference dialogue action information using the large language model;

[0042] match the dialogue action information about each of the candidate dialogue nodes with the reference dialogue action information;

[0043] if there is a candidate dialogue node matched with the reference dialogue action information, take the candidate dialogue node as the target dialogue node; and

[0044] if there is no candidate dialogue node matched with the reference dialogue action information, generate a provisional dialogue node and dialogue action information thereabout using the large language model, where the dialogue action information about the provisional dialogue node denotes guiding the dialogue to a dialogue node in the dialogue SOP tree.

[0045] In the third aspect, the present disclosure provides an electronic device, including a processor and a memory, where the memory stores a computer program, and the processor,

when executing the computer program, implements the intelligent dialogue method according to any one of preceding embodiments.

[0046] In the fourth aspect, the present disclosure provides a storage medium, where the store medium stores a computer program thereon, and the computer program, when executed by a processor, implements the intelligent dialogue method according to any one of preceding embodiments.

[0047] The present disclosure provides the intelligent dialogue method and apparatus, the electronic device and the storage medium. The method includes: in a process of conducting a dialogue with a user, obtaining each candidate dialogue node of next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree; taking all the candidate dialogue nodes as dialogue constraint information, and determining a target dialogue node of next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information; and generating a reply content of the next step using the large language model based on the target dialogue node, and feeding the reply content back to the user. By taking all the candidate nodes as the dialogue constraint information, and combining the SOP tree and the historical dialogue information, the large language model is enabled to determine the dialogue node of next step and generate the reply content, so as to conduct the dialogue with the user, thus ensuring the controllability and flexibility of the dialogue, improving the fluency and accuracy of the dialogue, and elevating the user experience.

[0048] In order to make the above objective, features and advantages of the present disclosure more obvious and understandable, preferred embodiments are particularly illustrated below to give detailed description in conjunction with drawings.

BRIEF DESCRIPTION OF DRAWINGS

[0049] In order to more clearly illustrate technical solutions of embodiments of the present disclosure, drawings which need to be used in the embodiments will be briefly introduced below. It should be understood that the drawings merely show some embodiments of the present disclosure, and thus should not be construed as limitation to the scope. Those ordinarily skilled in the art still could obtain other relevant drawings according to these drawings, without using any inventive efforts.

[0050] FIG. 1 shows a schematic block diagram of an electronic device provided by embodiments of the present disclosure;

[0051] FIG. 2 is a first schematic flowchart of an intelligent dialogue method provided by embodiments of the present disclosure;

[0052] FIG. 3 is a second schematic flowchart of the intelligent dialogue method provided by embodiments of the present disclosure; and

[0053] FIG. 4 shows a functional module diagram of an intelligent dialogue apparatus provided by embodiments of the present disclosure.

[0054] Reference signs: 100-electronic device; 110-bus; 120-processor; 130-memory; 150-I/O module; 170-communication module; 300-intelligent dialogue apparatus; 310-dialogue planning module; 330-dialogue generation module.

DETAILED DESCRIPTION OF EMBODIMENTS

[0055] Technical solutions in embodiments of the present disclosure will be described clearly and completely below in conjunction with drawings in the embodiments of the present disclosure, and apparently, only some but not all embodiments of the present disclosure are described. Generally, components in the embodiments of the present disclosure described and shown in the drawings herein may be arranged and designed in different configurations.

[0056] Therefore, the detailed description below of the embodiments of the present disclosure provided in the drawings is not intended to limit the scope of protection of the present disclosure, but merely represents chosen embodiments of the present disclosure. Based on the embodiments of the present disclosure, all of other embodiments obtained by those skilled in the art without using inventive efforts shall fall within the scope of protection of the present disclosure.

[0057] It should be noted that the relational terms such as “first” and “second” are merely used for distinguishing one entity or operation from another entity or operation, while it is not required or implied that these entities or operations necessarily have any such practical relation or order. Moreover, the terms such as “comprise”, “contain” or any other variants thereof are intended to encompass non-exclusive inclusion, such that a process, a method, an article or a device including a series of elements includes not only those elements, but also includes other elements that are not explicitly listed, or further includes elements inherent in such a process, method, article or device. Without further limitations, an element defined by the phrase “comprising a...” does not exclude the presence of additional identical elements in the process, method, article or device that includes the specified element.

[0058] With reference to FIG. 1, it is a schematic block diagram of an electronic device 100 provided by embodiments of the present disclosure. The electronic device 100 includes a bus 110, a processor 120, a memory 130, an I/O module 150, and a communication module 170.

[0059] The bus 110 is a circuit that connects the above elements to each other and transmits signals between the above elements.

[0060] The processor 120 can receive commands from the other elements (e.g., the memory 130, the I/O module 150, and the communication module 170) in the above through the bus 110, can interpret received commands, and can execute calculations or data processing according to the interpreted commands. The processor 120 may be an integrated circuit chip, with a function of processing signals. The processor 120 may be a general-purpose processor, including Central Processing Unit (CPU), and Network Processor (NP), etc., and also may be a digital signal processor (DSP), an application-specific integrated circuit (ASIC), a field-programmable gate array (FPGA) or other programmable logic devices, discrete gates or transistor logic devices, and discrete hardware components.

[0061] The memory 130 may store commands or data received from or generated by the processor 120 or other elements (e.g., the I/O module 150 and the communication module 170). The memory 130 can be, but is not limited to, Random Access Memory (RAM), Read Only Memory (ROM), Programmable Read-Only Memory (PROM), Erasable Programmable Read-Only Memory (EPROM), and Electric Erasable Programmable Read-Only Memory (EEPROM).

[0062] The I/O module 150 can receive commands or data input from a user via input-output means (e.g., a sensor, a keyboard, and a touch screen), and can transmit received commands or data to the processor 120 or the memory 130 through the bus 110. It is configured to display various kinds of information (e.g., multimedia data and text data) received by, stored in, and processed by the above elements, and can display videos, images, data, and the like to the user.

[0063] The communication module 170 may be configured to perform communication of signaling or data with other devices.

[0064] It is understandable that the structure shown in FIG. 1 is merely a structural schematic view of the electronic device 100, while the electronic device 100 may further include more or fewer components than those shown in FIG. 1, or have a configuration different from that shown in FIG. 1. Various components shown in FIG. 1 may be realized by hardware, software or a combination thereof.

[0065] The above electronic device 100 is taken as an execution subject below, to execute various steps in various methods provided by embodiments of the present disclosure and achieve corresponding technical effects.

[0066] With reference to FIG. 2, it is a schematic flowchart of an intelligent dialogue method provided by embodiments of the present disclosure.

[0067] Step S202, in a process of conducting a dialogue with a user, obtaining each candidate dialogue node of next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree.

[0068] In the present embodiment, a dialogue SOP tree is preconfigured based on a dialogue scene, so as to specify various dialogue nodes of the scene. Then in the process of conducting the dialogue with the user, a dialogue node that may be executed in the next step is acquired based on all the dialogue nodes and the current dialogue node in the preconfigured dialogue SOP tree, so as to obtain each candidate dialogue node. For example, a next dialogue node of the current dialogue node may be acquired from the dialogue SOP tree, so as to obtain each candidate dialogue node.

[0069] Optionally, each dialogue node in the dialogue SOP tree can be configured based on a predefined objective of the dialogue scene and a path for the dialogue to achieve the objective.

[0070] For example, in a dialogue scene of inviting a user to test-drive vehicles at a store, following dialogue nodes may be configured: opening greeting and identification, inquiring about the user's test-driving interest, introducing store promotions and inviting the user to test-drive specific models, scheduling test-driving time and place with the user, reminding the user of relevant information about test driving, expressing gratitude to the user and closing the dialogue, and etc. It can be understood that each dialogue node in the dialogue SOP tree is an intermediate step associated with the dialogue scene that needs to be confirmed or executed.

[0071] Step S204, taking all the candidate dialogue nodes as dialogue constraint information, and determining a target dialogue node of next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information.

[0072] Step S206, generating a reply content of the next step using the large language model based on the target dialogue node, and feeding the reply content back to the user.

[0073] In the present embodiment, after each candidate dialogue node is obtained, all the candidate dialogue nodes are taken as the dialogue constraint information. Then, based on the dialogue constraint information, the SOP tree and the historical dialogue information about dialogues with the user, the target dialogue node of the next step is determined using the

preconfigured large language model, where the target dialogue node may be understood as a dialogue node to be actually executed. Subsequently, based on the target dialogue node, the reply content of the next step is generated using the large language model and fed back to the user. Each step is executed in a similar way, so as to implement an intelligent dialogue between the large language model and the user.

[0074] It can be understood that, in embodiments of the present disclosure, all the candidate dialogue nodes are taken as constraint conditions for the large language model to conduct the dialogue with the user in the next step, and the SOP tree and the historical dialogue information are combined, to enable the large language model to determine an appropriate dialogue node and generate a reply content based on the dialogue scene and context, so as to conduct the dialogue with the user, thus ensuring the controllability and flexibility of the dialogue, improving fluency and accuracy of the dialogue, and elevating user experience.

[0075] It can be seen that based on the above steps, in the process of conducting the dialogue with the user, each candidate dialogue node of next step is obtained based on all the dialogue nodes and the current dialogue node in the preconfigured dialogue SOP tree; then all the candidate dialogue nodes are taken as the dialogue constraint information, and based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information, the target dialogue node of next step is determined using the preconfigured large language model; subsequently, based on the target dialogue node, the reply content of next step is generated using the large language model, and the reply content is fed back to the user. By taking all the candidate nodes as the dialogue constraint information and combining the SOP tree and the historical dialogue information, the large language model is enabled to determine the dialogue node of next step and generate the reply content, so as to conduct the dialogue with the user, thus ensuring the controllability and flexibility of the dialogue, improving the fluency and accuracy of the dialogue, and elevating the user experience.

[0076] Optionally, for the process of determining a target dialogue node of next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information in step S204, embodiments of the present disclosure provide a possible implementation mode, with reference to FIG. 3.

[0077] Step S204-1, generating an action acquisition instruction based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information.

[0078] Step S204-2, inputting the action acquisition instruction into the large language model, and generating reference dialogue action information using the large language model.

[0079] Step S204-3, matching the dialogue action information about each candidate dialogue node with the reference dialogue action information.

[0080] Step S204-4, if there is a candidate dialogue node matched with the reference dialogue action information, taking the candidate dialogue node as the target dialogue node.

[0081] Step S204-5, if there is no candidate dialogue node matched with the reference dialogue action information, generating a provisional dialogue node and dialogue action information thereabout using the large language model, and taking the provisional dialogue node as the target dialogue node, where the dialogue action information about the provisional dialogue node denotes guiding the dialogue to a dialogue node in the dialogue SOP tree.

[0082] In a possible implementation mode, the dialogue SOP tree, the dialogue constraint information, the historical dialogue information, a current SOP node and the objective to be finally achieved are taken as context information about a large language model instruction, to enable the large language model to generate an appropriate dialogue node of next step. Meanwhile, in a decoding and generating process of the large language model, a subsequent node of the current node is taken as a decoding-generation result, and the large language model is used to perform single-pass batch decoding. If a certain node has higher decoding result similarity than a predefined threshold, this node is taken as a dialogue node of next step. If there is no node with decoding result similarity higher than the predefined threshold, a dialogue node generated by the large language model is taken as the target dialogue node, and this node in this case is a “provisional dialogue node”.

[0083] The provisional dialogue node can be understood as an intermediate node, at which a provisional dialogue and guidance are performed, to guide the dialogue to a certain existing node in the dialogue SOP tree that facilitates achieving the final objective.

[0084] In the present embodiment, each dialogue node of the dialogue SOP tree includes the dialogue action information, where the dialogue action information may be understood as an action to be executed by the dialogue node.

[0085] For ease of understanding, a dialogue scene of inviting a user to test-drive vehicles at a store is taken as an example for illustration below, and dialogue nodes and dialogue action information in the dialogue SOP tree of the dialogue scene are listed in the following table.

Dialogue node	Dialogue action information
Dialogue node 1	Opening greeting and identification
Dialogue node 2	Inquiring about the user’s test-driving interest

Dialogue node 3	Introducing store promotions and inviting the user to test-drive model A or model B
Dialogue node 4-1	Scheduling time with the user for test-driving model A
Dialogue node 4-2	Scheduling time with the user for test-driving model B
Dialogue node 5	Reminding the user of relevant information about test driving
Dialogue node 6	Expressing gratitude to the user and closing the dialogue

[0086] It should be understood that, the above dialogue scene and contents in the above table are only examples, and the dialogue scene and the dialogue SOP tree may be configured according to actual situations, which are not limited in the embodiments of the present disclosure.

[0087] For example, assume that the current dialogue node is the dialogue node 3, its next dialogue nodes, i.e., dialogue node 4-1 and dialogue node 4-2, are taken as candidate dialogue nodes, and the dialogue node 4-1 and the dialogue node 4-2 are taken as the dialogue constraint information. Then the dialogue SOP tree, dialogue constraint information, historical dialogue information and a predefined action generation rule are concatenated, so as to obtain the action acquisition instruction.

[0088] Subsequently, the action acquisition instruction is input into the large language model, so as to enable the large language model to generate a dialogue action that it deems should be executed next based on the dialogue SOP tree, the dialogue constraint information, i.e., the dialogue node 4-1 and the dialogue node 4-2, and the history dialogue information, i.e., dialogue contents of the dialogue nodes 1 to 3, so as to output the reference dialogue action information.

[0089] Then, respective dialogue action information about the dialogue node 4-1 and the dialogue node 4-2 is matched with the reference dialogue action information, respectively, so as to determine whether there is a candidate dialogue node matched with the reference dialogue action information. If existing, it is indicated that the current dialogue follows the dialogue SOP tree, and then the matched candidate dialogue node is taken as the target dialogue node. If not existing, it is indicated that the current dialogue deviates from the dialogue SOP tree, then a provisional dialogue node and dialogue action information thereof are generated using the large language model, and the provisional dialogue node is taken as a target dialogue node, where the dialogue action information denotes guiding the dialogue to the dialogue node in the dialogue SOP tree.

[0090] In a possible implementation mode, the dialogue SOP tree, the dialogue constraint information, the historical dialogue information, a current SOP node and the objective to be finally achieved can be taken as context information about the large language model instruction, so as to enable the large language model to generate an appropriate dialogue node of next step.

[0091] Meanwhile, in a decoding and generating process of the large language model, two nodes, i.e., "dialogue node 4-1: scheduling time with the user for test-driving model A" and "dialogue node 4-2: scheduling time with the user for test-driving model B", are taken as a decoding-generation result, and the large language model is used to perform single-pass batch decoding. If a certain node has higher decoding result similarity than a predefined threshold, this node is taken as a dialogue node of next step. If there is no node with decoding result similarity higher than the predefined threshold, a dialogue node generated by the large language model is taken as the target dialogue node, and this node in this case is a "provisional dialogue node".

[0092] For example, if the user inquires about test-driving models other than models A and B, the decoding result similarities of the two nodes "dialogue node 4-1: scheduling time with the user for test-driving model A" and "dialogue node 4-2: scheduling time with the user for test-driving model B" will both fall below the threshold, and in this case, the large language model is enabled to generate a provisional dialogue node word by word. This node is autonomously generated by the large language model, and should be "Introducing test-driving information about other models", and then the provisional dialogue node is taken as a dialogue node in current stage, and a dialogue action is generated, i.e., using the large language model to generate a specific reply and an inquiry script based on test-driving information about other models.

[0093] If the user expresses interest still in model A, the decoding result similarity of the node "dialogue node 4-1: scheduling time with the user for test-driving model A" in the decoding result of next selected node of the large language model is higher than the predefined threshold, in which case, this node can be taken as a dialogue node of next step, indicating that the dialogue gets back to the SOP tree.

[0094] In addition, if the user has any request or question during the dialogue, which causes the dialogue to deviate from the dialogue SOP tree, a provisional dialogue node is also generated for provisionally addressing the request or question of the user, and meanwhile, it is required to guide the dialogue to a certain node in the dialogue SOP tree, so as to continue subsequent dialogue.

[0095] It can be understood as that the embodiments of the present disclosure ensure the controllability and flexibility of the dialogue by matching the reference dialogue action information generated by the large language model respectively with the dialogue action information about the candidate dialogue nodes, so as to determine whether the dialogue deviates from the dialogue SOP tree, and if the dialogue deviates from the dialogue SOP tree, guiding the dialogue to the dialogue SOP tree by generating a provisional dialogue node and dialogue action information thereof.

[0096] Optionally, for step S204-3, embodiments of the present disclosure provide a possible implementation mode.

[0097] Step S204-3-1, calculating similarity between the dialogue action information about each candidate dialogue node and the reference dialogue action information, so as to obtain corresponding similarity of each candidate dialogue node.

[0098] Step S204-3-2, if there is a candidate dialogue node with the similarity reaching a predefined first threshold, taking a candidate dialogue node corresponding to the greatest similarity as the candidate dialogue node matched with the reference dialogue action information.

[0099] Step S204-3-3, if there is no candidate dialogue node with similarity reaching the first threshold, determining that there is no candidate dialogue node matched with the reference dialogue action information.

[0100] For ease of understanding, the above example is further used for illustration. For example, in an implementation mode, assume that content replied by the user to the dialogue node 3 is model A, and the reference dialogue action information output by the large language model is: confirming relevant information for the user to test-drive model A. Then the similarity between the reference dialogue action information and respective dialogue action information about the dialogue node 4-1 and the dialogue node4-2 is calculated, respectively, so as to obtain respective corresponding similarities of the dialogue node 4-1 and the dialogue node 4-2.

[0101] Next, respective corresponding similarities of the dialogue node 4-1 and the dialogue node 4-2 are respectively compared with the predefined first threshold. If corresponding similarity of only one dialogue node reaches the first threshold, this candidate dialogue node is taken as the candidate dialogue node matched with the reference dialogue action information. For example, if model A is Tesla and model B is Xiaomi, and upon similarity calculation, only corresponding similarity of the dialogue node 4-1 reaches the first threshold, the dialogue node 4-1 is taken as the matched candidate dialogue node.

[0102] If the corresponding similarities of a plurality of dialogue nodes all reach the first threshold, the candidate dialogue node corresponding to the greatest similarity is taken as the candidate dialogue node matched with the reference dialogue action information. For example, model A is Tesla Model X, model B is Tesla Model Y, and upon similarity calculation, respective corresponding similarities of the dialogue node 4-1 and the dialogue node 4-2 reach the first threshold, and the dialogue node 4-1 has greater similarity than the dialogue node 4-2, the dialogue node 4-1 is taken as the matched candidate dialogue node.

[0103] In another implementation mode, assume that content replied by the user to the dialogue node 3 is model C, and the reference dialogue action information output by the large language model is: confirming relevant information for the user to test-drive model C. Then the similarity between the reference dialogue action information and respective dialogue action information about the dialogue node 4-1 and the dialogue node4-2 is calculated, respectively, so as to obtain respective corresponding similarities of the dialogue node 4-1 and the dialogue node 4-2.

[0104] Next, respective corresponding similarities of the dialogue node 4-1 and the dialogue node 4-2 are compared with the predefined first threshold, respectively. If corresponding similarities of all the candidate dialogue nodes fail to reach the first threshold, it is determined that there is no candidate dialogue node matched with the reference dialogue action information. For example, if model A is Tesla, model B is Xiaomi, and model C is NIO, and upon similarity calculation, both respective corresponding similarities of the dialogue node 4-1 and the dialogue node 4-2 fail to reach the first threshold, it is determined that there is no matched candidate dialogue node.

[0105] It can be understood that, the large language model outputs a text by outputting tokens, i.e., characters, one by one, and then similarity between the tokens already output and the dialogue action information about each candidate dialogue node can be calculated in real time based on the tokens output by the large language model, so as to obtain the similarity between the reference dialogue action information and the dialogue action information about each candidate dialogue node.

[0106] Optionally, for step S206, embodiments of the present disclosure provide a possible implementation mode.

[0107] Step S206-a1, if the target dialogue node is a provisional dialogue node, generating a first script acquisition instruction based on the dialogue action information about the provisional dialogue node.

[0108] Step S206-a2, inputting the first script acquisition instruction into the large language model, generating a guidance script using the large language model, and taking the guidance script as a reply content of next step.

[0109] For ease of understanding, the above example is further used for illustration. For example, if the content replied by the user to the dialogue node 3 is model C, there is no candidate dialogue node matched with the reference dialogue action information output by the large language model, and the provisional dialogue node generated by the large language model is taken as the target dialogue node.

[0110] If the target dialogue node is the provisional dialogue node, the dialogue action information about the provisional dialogue node is concatenated with a predefined first script generation rule, so as to obtain a first script acquisition instruction. Moreover, the first script acquisition instruction is input into the large language model, so as to enable the large language model to generate a guidance script. Then the guidance script is taken as a reply content of next step and fed back to the user, so as to get the dialogue back to the dialogue SOP tree.

[0111] Optionally, the dialogue SOP tree, the dialogue constraint information, the historical dialogue information, a current SOP node and the target to be finally achieved are taken as context information about the large language model instruction, and the provisional dialogue node is taken as an objective of current reply, so as to enable the large language model to generate appropriate reply content. Moreover, a preceding SOP node and adjacent node are given as target nodes of the guidance script, to enable the large language model to generate an appropriate reply and guide the user back to a certain node therein. The two parts of contents are autonomously generated by the large language model based on the context information.

[0112] For example, the large language model can generate a guidance script stating that promotions are exclusively available for model A and model B and recommending test-driving the two models, so as to guide the dialogue to the dialogue node 3. The large language model may further generate a guidance script stating that model A or model B is more advantageous than model C, so as to guide the dialogue to the dialogue node 4-1 or the dialogue node 4-2.

[0113] Optionally, for step S206, embodiments of the present disclosure further provide a possible implementation mode.

[0114] Step S206-b1, if the target dialogue node is the candidate dialogue node, acquiring each standard script of the candidate dialogue node, so as to obtain each candidate script.

[0115] Step S206-b2, taking all the candidate scripts as script constraint information, and generating a second script acquisition instruction based on the script constraint information.

[0116] Step S206-b3, inputting the second script acquisition instruction into the large language model, and generating a target script using the large language model.

[0117] Step S206-b4, selecting a reply content of next step from the target script and all candidate scripts.

[0118] In the present embodiment, each dialogue node of the dialogue SOP tree further includes a plurality of standard scripts.

[0119] For ease of understanding, the above example is further used for illustration. For example, if the content replied by the user to the dialogue node 3 is model A, the dialogue node 4-1 is the candidate dialogue node matched with the reference dialogue action information, and the dialogue node 4-1 is taken as the target dialogue node.

[0120] When the target dialogue node is the candidate dialogue node, i.e., the dialogue node 4-1, each standard script of the dialogue node 4-1 is acquired, so as to obtain each candidate script. Then, all the candidate scripts are taken as script constraint information, and the script constraint information is concatenated with a predefined second script generation rule, so as to obtain the second script acquisition instruction. Then the second script acquisition instruction is input into the large language model, so as to enable the large language model to generate the target script with reference to the candidate scripts. Next, reply content of next step is selected from the target script and all the candidate scripts.

[0121] In some embodiments, the second script acquisition instruction can be input into the large language model, and in a process of the large language model decoding and generating the script, all candidate scripts are taken as a decoding-generation result, and the large language model is used to perform single-pass batch decoding. If a certain candidate script has higher decoding result similarity than a predefined threshold, the candidate script is considered as meeting a set standard, that is, if the target script is matched with the dialogue action information about the dialogue node 4-1, the candidate script is taken as the reply content of next step and fed back to the user. If there is no candidate script with decoding result similarity higher than the predefined threshold, the candidate script with the greatest similarity is taken as reply content of next step and fed back to the user.

[0122] It may be understood as that the embodiments of the present disclosure take all standard scripts of the matched candidate dialogue nodes as the script constraint information to guide the large language model to generate a similar script, thus precluding the model hallucination, and improving the accuracy and fluency of the dialogue.

[0123] Optionally, for step S206-b4, embodiments of the present disclosure provide a possible implementation mode.

[0124] Step S206-b4-1, calculating the similarity between the target script and each candidate script.

[0125] Step S206-b4-2, if the similarity between the target script and at least one candidate script reaches a predefined second threshold, taking the target script as reply content of next step.

[0126] Step S206-b4-3, if the similarities between the target script and all the candidate scripts fail to reach the second threshold, taking the candidate script with the greatest similarity as reply content of next step.

[0127] For ease of understanding, the above example is further used for illustration. For example, assume that the dialogue node 4-1 includes three standard scripts, and then three candidate scrips are obtained. Then the similarities between the target script and the three candidate scripts are calculated, respectively.

[0128] If the similarity between the target script and at least one candidate script reaches the predefined second threshold, it is indicated that the target script meets the set standard, that is, the target script is matched with the dialogue action information about the dialogue node 4-1, and the candidate script is taken as the reply content of next step and fed back to the user.

[0129] If the similarities between the target script and all the candidate scripts fail to reach the second threshold, it is indicated that the target script does not meet the set standard, that is, the target script is not matched with the dialogue action information about the dialogue node 4-1, and the candidate script with the greatest similarity is taken as reply content of next step and fed back to the user.

[0130] It may be understood that, by calculating the similarities between the scripts generated by the large language model and the standard script, and selecting the reply content, the embodiments of the present disclosure further preclude the model hallucination, and improve the accuracy and fluency of the dialogue.

[0131] In order to execute corresponding steps in the above embodiments and various possible embodiments, an implementation mode of an intelligent dialogue apparatus is provided below. With reference to FIG. 4, it is a functional module diagram of the intelligent dialogue apparatus provided by embodiments of the present disclosure. It should be noted that an implementation principle of an intelligent dialogue apparatus 300 provided by the present embodiment and technical effects produced are the same as those of the above embodiments. For brevity of description, for parts that are not mentioned in the present embodiment, reference can be made to corresponding contents in the above embodiments. The intelligent dialogue apparatus 300 includes:

[0132] a dialogue planning module 310, configured to, in a process of conducting a dialogue with a user, obtain each candidate dialogue node of next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree; take all the candidate dialogue nodes as the dialogue constraint information, and determine a target dialogue node of next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information; and

[0133] a dialogue generation module 330, configured to generate a reply content of next step using the large language model based on the target dialogue node, and feed the reply content back to the user.

[0134] Optionally, the dialogue planning module 310 is further configured to: generate an action acquisition instruction based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information; input the action acquisition instruction into the large language model, and generate reference dialogue action information using the large language model; match the dialogue action information about each candidate dialogue node with the reference dialogue action information; if there is a candidate dialogue node matched with the reference dialogue action information, take the candidate dialogue node as the target dialogue node; if there is no candidate dialogue node matched with the reference dialogue action information, generate a provisional dialogue node and dialogue action information thereabout using the large language model, and take the provisional dialogue node as the target dialogue node, where the dialogue action information about the provisional dialogue node denotes guiding the dialogue to a dialogue node in the dialogue SOP tree.

[0135] Optionally, the dialogue planning module 310 is further configured to: calculate similarity between the dialogue action information about each candidate dialogue node and the reference dialogue action information, so as to obtain corresponding similarity of each candidate dialogue node; if there is a candidate dialogue node with the similarity reaching a predefined first threshold, take a candidate dialogue node corresponding to the greatest similarity as the candidate dialogue node matched with the reference dialogue action information; and if there is no candidate dialogue node with similarity reaching the first threshold, determine that there is no candidate dialogue node matched with the reference dialogue action information.

[0136] Optionally, the dialogue generation module 330 is further configured to: if the target dialogue node is a provisional dialogue node, generate a first script acquisition instruction based on the dialogue action information about the provisional dialogue node; and input the

first script acquisition instruction into the large language model, generate a guidance script using the large language model, and take the guidance script as a reply content of next step.

[0137] Optionally, the dialogue generation module 330 is further configured to: if the target dialogue node is the candidate dialogue node, acquire each standard script of the candidate dialogue node, so as to obtain each candidate script; and take all the candidate scripts as script constraint information, and generating a second script acquisition instruction based on the script constraint information; input the second script acquisition instruction into the large language model, and generate a target script using the large language model; and select reply content of next step from the target script and all candidate scripts.

[0138] Optionally, the dialogue generation module 330 is further configured to: calculate the similarity between the target script and each candidate script; if the similarity between the target script and at least one candidate script reaches a predefined second threshold, take the target script as reply content of next step; and if the similarities between the target script and all the candidate scripts fail to reach the second threshold, take the candidate script with the greatest similarity as reply content of next step.

[0139] Embodiments of the present disclose further provide an electronic device, including a processor and a memory, where the memory stores a computer program, and the processor, when executing the computer program, implements the intelligent dialogue method disclosed in the embodiments of the present disclosure.

[0140] Embodiments of the present disclosure further provide a storage medium, storing a computer program thereon, where the computer program, when executed by a processor, implements the intelligent dialogue method disclosed in the embodiments of the present disclosure.

[0141] In the several embodiments provided by the present disclosure, it should be understood that the apparatus and the method disclosed further may be implemented in other manners. The apparatus embodiments described in the above are merely illustrative. For example, flowcharts and block diagrams in the drawings show system structures, functions and operations that can be implemented in the apparatus, method and computer program products according to multiple embodiments of the present disclosure. In this regard, each block in the flowcharts or the block diagrams may represent a part of a module, a program segment or a code, which contains one or more executable instructions configured to realize a specified logical function. It also should be noted that in some implementation modes as substitutions, the functions marked in blocks also can take effect in an order different than that marked in the drawings. For example, two consecutive blocks substantially can be concurrently executed, or

they sometimes can be executed in a reverse order, which depends upon the functions involved. It is also noted that each block of the block diagrams and/or flowcharts, and combinations of blocks in the block diagrams and/or flowcharts, can be implemented in dedicated hardware-based systems that perform the specified functions or actions, or can be implemented in a combination of dedicated hardware and computer instructions.

[0142] Besides, various functional modules in various embodiments of the present disclosure can be integrated together to form one independent part, and it is also possible that the various modules exist independently, or that two or more modules are integrated to form one independent part.

[0143] If the function is realized in a form of software functional module and is sold or used as an individual product, it may be stored in one computer-readable storage medium. Based on such understanding, the essence of the technical solutions of the present disclosure or the parts that make contribution to the prior art thereof may be embodied in form of a software product, and the computer software product is stored in a storage medium, including several instructions for making one computer device (which can be a personal computer, a server or a network device etc.) execute all or part of the steps of the methods of various embodiments of the present disclosure. The aforementioned storage medium includes various media in which program codes can be stored, such as USB flash disk, mobile hard disk, Read-Only Memory (ROM), Random Access Memory (RAM), diskette or compact disk.

[0144] The above-mentioned are only preferred embodiments of the present disclosure and are not intended to limit the present disclosure. For those skilled in the art, the present disclosure may have various modifications and changes. Any modifications, equivalent substitutions, improvements and the like made within the spirit and principle of the present disclosure should be covered within the scope of protection of the present disclosure.

INDUSTRIAL APPLICABILITY

[0145] Through the above solutions, the large language model is used to generate the reply content of next step, thus ensuring the controllability and flexibility of the dialogue, improving fluency and accuracy of the dialogue, and elevating the user experience.

CLAIMS

Claim 1: An intelligent dialogue method, comprising steps of:

in a process of conducting a dialogue with a user, obtaining each candidate dialogue node of a next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree;

taking all the candidate dialogue nodes as dialogue constraint information, and determining a target dialogue node of the next step using a preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information; and

generating a reply content of the next step using the large language model based on the target dialogue node, and feeding the reply content back to the user.

Claim 2: The intelligent dialogue method according to claim 1, wherein the step of obtaining each candidate dialogue node of next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree comprises:

acquiring a next dialogue node of the current dialogue node from all the dialogue nodes, so as to obtain each candidate dialogue node of the next step.

Claim 3: The intelligent dialogue method according to claim 1 or 2, wherein each dialogue node comprises dialogue action information, wherein the dialogue action information is configured to indicate an action to be executed by the dialogue node; and

the step of determining a target dialogue node of the next step using a preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information comprises:

generating an action acquisition instruction based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information;

inputting the action acquisition instruction into the large language model, and generating reference dialogue action information using the large language model;

matching the dialogue action information about each of the candidate dialogue nodes with the reference dialogue action information;

if there is a candidate dialogue node matched with the reference dialogue action information, taking the candidate dialogue node as the target dialogue node; and

if there is no candidate dialogue node matched with the reference dialogue action information, generating a provisional dialogue node and dialogue action information thereabout using the large language model, and taking the provisional dialogue node as the target dialogue node, where the dialogue action information about the provisional dialogue node denotes guiding the dialogue to a dialogue node in the dialogue SOP tree.

Claim 4: The intelligent dialogue method according to claim 3, wherein the step of generating an action acquisition instruction based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information comprises:

concatenating the dialogue SOP tree, the dialogue constraint information, the historical dialogue information and a predefined action generation rule, so as to obtain an action acquisition instruction.

Claim 5: The intelligent dialogue method according to claim 3 or 4, wherein the step of matching the dialogue action information about each of the candidate dialogue nodes with the reference dialogue action information comprises:

calculating similarity between the dialogue action information about each of the candidate dialogue nodes and the reference dialogue action information, so as to obtain corresponding similarity of each of the candidate dialogue nodes;

if there is a candidate dialogue node with a similarity reaching a predefined first threshold, taking a candidate dialogue node corresponding to a greatest similarity as the candidate dialogue node matched with the reference dialogue action information; and

if there is no candidate dialogue node with a similarity reaching the first threshold, determining that there is no candidate dialogue node matched with the reference dialogue action information.

Claim 6: The intelligent dialogue method according to claim 5, wherein the step of, if there is a candidate dialogue node with a similarity reaching a predefined first threshold, taking a candidate dialogue node corresponding to a greatest similarity as the candidate dialogue node matched with the reference dialogue action information comprises:

if similarity of one candidate dialogue node reaches the first threshold, taking the candidate dialogue node as the candidate dialogue node matched with the reference dialogue action information; and

if similarities of a plurality of candidate dialogue nodes reach the first threshold, taking the candidate dialogue node corresponding to the greatest similarity as the candidate dialogue node matched with the reference dialogue action information.

Claim 7: The intelligent dialogue method according to any one of claims 1–6, wherein the step of generating a reply content of the next step using the large language model based on the target dialogue node comprises:

if the target dialogue node is a provisional dialogue node, generating a first script acquisition instruction based on dialogue action information about the provisional dialogue node; and

inputting the first script acquisition instruction into the large language model, generating a guidance script using the large language model, and taking the guidance script as a reply content of the next step.

Claim 8: The intelligent dialogue method according to claim 7, wherein the step of inputting the first script acquisition instruction into the large language model, generating a guidance script using the large language model, and taking the guidance script as a reply content of the next step comprises:

if the target dialogue node is the candidate dialogue node, the large language model generating a guidance script based on a standard script of the target dialogue node, and taking the guidance script as the reply content of the next step; and

if the target dialogue node is the provisional dialogue node, the large language model generating a guidance script based on an adjacent node of the

target dialogue node, and taking the guidance script as the reply content of the next step.

Claim 9: The intelligent dialogue method according to claim 8, wherein the step of generating a first script acquisition instruction based on dialogue action information about the provisional dialogue node comprises:

concatenating the dialogue action information about the provisional dialogue node with a predefined first script generation rule, so as to obtain a first script acquisition instruction.

Claim 10: The intelligent dialogue method according to any one of claims 1–9, wherein each dialogue node further includes a plurality of standard scripts;

the step of generating a reply content of the next step using the large language model based on the target dialogue node comprises:

when the target dialogue node is the candidate dialogue node, acquiring each standard script of the candidate dialogue node, so as to obtain each candidate script;

taking all the candidate scripts as script constraint information, and generating a second script acquisition instruction based on the script constraint information;

inputting the second script acquisition instruction into the large language model, and generating a target script using the large language model; and

selecting a reply content of the next step from the target script and all the candidate scripts.

Claim 11: The intelligent dialogue method according to claim 10, wherein the step of selecting a reply content of next step from the target script and all the candidate scripts comprises:

calculating a similarity between the target script and each of the candidate scripts;

if a similarity between the target script and at least one of the candidate scripts reaches a predefined second threshold, taking the target script as the reply content of the next step; and

if similarities between the target script and all the candidate scripts fail to reach the second threshold, taking a candidate script with a greatest similarity as the reply content of the next step.

Claim 12: The intelligent dialogue method according to any one of claims 1–11, wherein a creation process of the preconfigured dialogue SOP tree is as follows:

configuring each dialogue node in the dialogue SOP tree based on a predefined objective of a dialogue scene and a path for a dialogue to achieve an objective; and

generating the preconfigured dialogue SOP tree based on each dialogue node.

Claim 13: The intelligent dialogue method according to any one of claims 1–12, wherein a creation process of the preconfigured dialogue SOP tree is as follows:

batch-generating simulated dialogic data within a dialogue scene using the large language model;

determining dialogue branch nodes; and

assembling and generating the preconfigured dialogue SOP tree based on the simulated dialogic data and the dialogue branch nodes.

Claim 14: An intelligent dialogue apparatus, comprising:

a dialogue planning module, configured to, in a process of conducting a dialogue with a user, obtain each candidate dialogue node of a next step based on all dialogue nodes and a current dialogue node in a preconfigured dialogue SOP tree;

take all the candidate dialogue nodes as dialogue constraint information, and determine a target dialogue node of the next step using the preconfigured large language model based on the dialogue SOP tree, the dialogue constraint information and historical dialogue information; and

a dialogue generation module, configured to generate a reply content of the next step using the large language model based on the target dialogue node, and feed the reply content back to the user.

Claim 15: The intelligent dialogue apparatus according to claim 14, wherein each dialogue node comprises dialogue action information, wherein the dialogue action information is configured to indicate an action to be executed by the dialogue node; and the dialogue planning module is further configured to:

generate an action acquisition instruction based on the dialogue SOP tree, the dialogue constraint information and the historical dialogue information;

input the action acquisition instruction into the large language model, and generate reference dialogue action information using the large language model;

match the dialogue action information about each of the candidate dialogue nodes with the reference dialogue action information;

if there is a candidate dialogue node matched with the reference dialogue action information, take the candidate dialogue node as the target dialogue node; and

if there is no candidate dialogue node matched with the reference dialogue action information, generate a provisional dialogue node and dialogue action information thereabout using the large language model, wherein the dialogue action information about the provisional dialogue node denotes guiding the dialogue to a dialogue node in the dialogue SOP tree.

Claim 16: An electronic device, comprising a processor and a memory, wherein the memory stores a computer program, and the processor, when executing the computer program, implements the intelligent dialogue method according to any one of claims 1–13.

Claim 17: A storage medium, wherein the store medium stores a computer program thereon, and the computer program, when executed by a processor, implements the intelligent dialogue method according to any one of claims 1–13.