DESCRIPTION

TITLE OF THE INVENTION: POST-TRAINING METHOD AND APPARATUS FOR LARGE QUESTION-ANSWERING MODEL, 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. 2025102875932 filed on March 12, 2025 with the Chinese Patent Office, and entitled "POST-TRAINING METHOD AND APPARATUS FOR LARGE QUESTION-ANSWERING MODEL, 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 a post-training method and apparatus for a large question-answering model, an electronic device and a storage medium.

BACKGROUND ART

[0003] In the prior art, models are often enabled to respond (answer) more accurately through post-training. For example, typical approaches are RLHF (Reinforcement Learning from Human Feedback) method and DPO (Direct Preference Optimization) method. However, the RLHF method and the DPO method both in essence enable models to answer questions according to human preferences, that is, making response contents of models close to human preferences. Neither of the two methods substantially enhances the reasoning capability of models.

[0004] Therefore, how to improve the reasoning capability of models so as to enable the models to respond more accurately is a problem to be solved urgently.

SUMMARY

[0005] Embodiments of the present disclosure at least provide a post-training method for a large question-answering model and related devices. The related devices may include a post-training apparatus for the large question-answering model, an electronic device, a computer-

readable storage medium and a computer program product, thus enabling the large question-answering model to develop capability of performing step-by-step reasoning on questions, thereby fundamentally elevating a logical reasoning capability of the model.

[0006] In order to achieve at least one of the above objectives, technical solutions employed by the present disclosure are as follows.

[0007] In the first aspect, the present disclosure provides a post-training method for a large question-answering model, including steps of:

[0008] acquiring multiple pieces of inference data, where the inference data is obtained by logically splitting an original question and an original answer thereof;

[0009] constructing an inference dataset based on all the inference data; and

[0010] performing post-training processing on the large question-answering model using the inference dataset, where the inference dataset is configured to guide the large question-answering model to learn to perform step-by-step reasoning on an input question.

[0011] In an optional embodiment, the inference dataset includes a first subset, a second subset, a third subset and a fourth subset; and the step of constructing an inference dataset based on all the inference data includes:

[0012] obtaining at least one first sample for each piece of the inference data using a progressive reasoning type construction rule based on the inference data, where the first subset includes all first samples corresponding to each piece of the inference data;

[0013] obtaining at least one second sample for each piece of the inference data using a progressive decomposition type construction rule based on the inference data, where the second subset includes all second samples corresponding to each piece of the inference data;

[0014] obtaining at least one third sample for each piece of the inference data using a mid-segment extraction type construction rule based on the inference data, where the third subset includes all third samples corresponding to each piece of the inference data; and

[0015] obtaining at least one fourth sample for each piece of the inference data using a step-by-step reasoning type construction rule based on the inference data, where the fourth subset includes all fourth samples corresponding to each piece of the inference data.

[0016] In an optional embodiment, the first subset is configured to guide the large question-answering model to learn to answer the input question step by step, the second subset is configured to guide the large question-answering model to learn to logically decompose the input question, the third subset is configured to guide the large question-answering model to learn to complement a missing intermediate process based on a step-by-step reasoning mode,

and the fourth subset is configured to guide the large question-answering model to learn to answer step by step following logical decomposition on the input question.

[0017] In an optional embodiment, the step of performing post-training processing on the large question-answering model using the inference dataset includes:

[0018] performing the post-training processing on the large question-answering model using the first subset, the second subset, the third subset and the fourth subset, respectively.

[0019] In an optional embodiment, the step of performing post-training processing on the large question-answering model using the inference dataset includes:

[0020] performing the post-training processing on the large question-answering model using the first subset, the second subset, the third subset and the fourth subset in a sequential order, where the sequential order includes: a sequentially arranged order or a randomly arranged order.

[0021] In an optional embodiment, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question;

[0022] the step of obtaining at least one first sample using a progressive reasoning type construction rule based on the inference data includes:

[0023] concatenating a first prompt text, and the original question and a first sub-question in the inference data to obtain a first one of first sample, and setting a label of the first one of first sample as a first sub-answer in the inference data, where the first prompt text is configured to instruct the large question-answering model to infer an answer to a last sub-question in an input content based on the input content; and

[0024] concatenating the first prompt text, and the original question, an n-th sub-question, and each sub-question preceding the n-th sub-question and sub-answers thereof in the inference data to obtain an n-th first sample, and setting a label of the n-th first sample as an n-th sub-answer in the inference data, where n = 2, 3, …, N.

[0025] In an optional embodiment, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question;

[0026] the step of obtaining at least one second sample using a progressive decomposition type construction rule based on the inference data includes:

[0027] concatenating a second prompt text, and the original question and a first sub-question in the inference data to obtain a first second sample, and setting a label of the first second sample as a second sub-answer in the inference data, where the second prompt text is configured to instruct the large question-answering model to obtain a sub-question immediately

succeeding a last sub-question in an input content by decomposing the original question in the input content based on the input content; and

[0028] concatenating the second prompt text, and the original question, an m-th sub-question and each preceding sub-question in the inference data to obtain an m-th second sample, and setting a label of the m-th second sample as an (m+1)-th sub-question in the inference data, where m = 2, 3, …, N-1.

[0029] In an optional embodiment, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question;

[0030] the step of obtaining at least one third sample using a mid-segment extraction type construction rule based on the inference data includes:

[0031] concatenating a third prompt text, and each sub-question and sub-answer in the inference data, excluding a first sub-question and a first sub-answer, to obtain a first third sample, and setting a label of the first third sample as the first sub-question and the first sub-answer in the inference data, where the third prompt text is configured to instruct the large question-answering model to infer missing intermediate sub-questions and intermediate sub-answers thereof in an input content based on the input content; and

[0032] concatenating the third prompt text, and each sub-question and sub-answer in the inference data, excluding an m-th sub-question and an m-th sub-answer, to obtain an m-th third sample, and setting a label of the m-th third sample as the m-th sub-question and the m-th sub-answer in the inference data, where m = 2, 3, …, N-1.

[0033] In an optional embodiment, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question;

[0034] the step of obtaining at least one fourth sample using a step-by-step reasoning type construction rule based on the inference data includes:

[0035] concatenating a fourth prompt text, and the original question in the inference data to obtain a first fourth sample, and setting a label of the first fourth sample as each sub-question and sub-answer thereof in the inference data, where the fourth prompt text is configured to instruct the large question-answering model to infer subsequent sub-questions and sub-answers step by step based on an input content; and

[0036] concatenating the fourth prompt text, and the original question, each sub-question preceding an n-th sub-question and sub-answers thereof in the inference data to obtain an n-th

fourth sample, and setting a label of the n-th fourth sample as each sub-question subsequent to the (n-1)-th sub-question and sub-answers thereof in the inference data, where n = 2, 3, …, N.

[0037] In an optional embodiment, a process of the post-training processing includes: a supervised fine-tuning mode or an instruction tuning mode.

[0038] In the second aspect, the present disclosure further provides a post-training apparatus for a large question-answering model, including:

[0039] a data acquisition module, configured to acquire multiple pieces of inference data, where the inference data is obtained by logically splitting an original question and an original answer thereof;

[0040] a construction module, configured to construct an inference dataset based on all the inference data; and

[0041] a post-training module, configured to perform post-training processing on the large question-answering model using the inference dataset, where the inference dataset is configured to guide the large question-answering model to learn to perform step-by-step reasoning on an input question.

[0042] In the third aspect, the present disclosure further provides an electronic device, including: a memory and a processor, where the memory stores a software program, and when the electronic device runs, the processor executes the software program so as to implement the post-training method for the large question-answering model in the first aspect.

[0043] In the fourth aspect, the present disclosure further provides a computer-readable storage medium, where the computer-readable storage medium stores a computer program, and the computer program, when executed by the processor, implements the post-training method for the large question-answering model in the first aspect.

[0044] Compared with the prior art, the present disclosure provides the post-training method and apparatus for the large question-answering model, the electronic device and the storage medium. Firstly, multiple pieces of inference data are acquired; then the inference dataset is constructed based on all the inference data; and finally the post-training processing is performed on the large question-answering model using the inference dataset. As each piece of inference data is obtained by logically splitting the original question and the original answer thereof, the inference dataset can be configured to guide the large question-answering model to learn to perform step-by-step reasoning on the input question. The present disclosure enables the large question-answering model to develop the capability of performing step-by-step reasoning on questions, thereby fundamentally improving the logical reasoning capability of the model.

[0045] In order to make the above objectives, features and advantages of the present disclosure more obvious and understandable, preferred embodiments are particularly illustrated below to give detailed description in conjunction with the drawings.

BRIEF DESCRIPTION OF DRAWINGS

[0046] 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. The drawings herein are incorporated into the description and constitute a part of the description. The drawings show embodiments complying with the present disclosure, and are used to illustrate the technical solutions of the present disclosure together with the description. It should be understood that the drawings merely show some embodiments of the present disclosure, and thus should not be considered 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.

[0047] FIG. 1 is a first schematic flowchart of a post-training method for a large question-answering model provided by embodiments of the present disclosure.

[0048] FIG. 2 is a second schematic flowchart of a post-training method for a large question-answering model provided by embodiments of the present disclosure.

[0049] FIG. 3 is a structural schematic view of a post-training apparatus for a large question-answering model provided by embodiments of the present disclosure.

[0050] FIG. 4 is a structural schematic view of an electronic device provided by embodiments of the present disclosure.

DETAILED DESCRIPTION OF EMBODIMENTS

[0051] In order to make objectives, technical solutions and advantages of embodiments of the present disclosure clearer, the technical solutions in the embodiments of the present disclosure will be described clearly and completely below in conjunction with the drawings in the embodiments of the present disclosure. Apparently, the described embodiments are only some but not all of the embodiments of the present disclosure. Generally, components in the embodiments of the present disclosure described and shown in the drawings herein may be arranged and designed in different configurations. 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.

[0052] It should be noted that like reference signs and letters represent like items in the following drawings, and thus, once a certain item is defined in one drawing, it is unnecessary to further define and explain the same in subsequent drawings.

[0053] Besides, the terms such as “first” and “second”, if appear, are merely intended for differentiated description, but should not be construed as an indication or an implication of importance in the relativity.

[0054] It should be noted that features in the embodiments of the present disclosure may be combined with each other without conflict.

[0055] In the prior art, in order to make contents of model responses more accurate, there are two approaches.

[0056] The first approach is to enhance reasoning capability of models through post-training, so as to make model responses more accurate. Typical methods are the RLHF method and the DPO method. The RLHF method is implemented according to the following principle: firstly, models need to be fed with data following instruction fine-tuning, and upon acquisition of an answer to the data, preference data is annotated artificially, for example, annotated data (x, ylose, ywin); then a reward model modelaward is trained using the annotated preference data. In order to make distribution of a trained model not differ too much from that of an original model, a reference model modelrefer is further needed. Therefore, three models are required in a training process. The DPO method fine-tunes the model directly on a preference dataset, without the need for a reward model or a reference model, and is a simplified RLHF method.

[0057] The second approach is to increase reasoning time, so as to enable the model to respond more accurately. The most common practice is to link the model to an external knowledge base or search engine, so that the model takes acquired external knowledge as a reference to answer questions. To link to the external knowledge base, a BGE (Bidirectional Generative Embedding) module is demanded to segment and quantize contents of the external knowledge base in advance, so as to enable the model to efficiently retrieve and utilize the knowledge.

[0058] With regard to the first approach in the above, it in essence simply lets the model answer questions according to human preferences, but fails to substantially enhance the reasoning capability of the model. Moreover, a standard RLHF method demands three models at the same time during training, and has higher requirements on hardware.

[0059] The second approach in the above just provides more references for the model, and also fails to substantially improve the reasoning capability of the model. Moreover, a reasoning process requires integration of the bge module, which introduces additional reasoning time.

[0060] Therefore, how to enable the model to respond more accurately by improving the reasoning capability of the model is a problem to be solved urgently.

[0061] Based on the finding of the above technical problems, the inventors propose the following technical solutions through inventive efforts so as to solve or address the above problems. It should be noted that the above defects in the prior art solutions are all results obtained by the inventors through practice and deliberate study. Therefore, the discovery process of the above problems and following solutions for the above problems proposed in embodiments of the present disclosure should be contributions made by the inventors to the present disclosure during creation of the invention, and should not be construed as technical contents well known to those skilled in the art.

[0062] Embodiments of the present disclosure provide a post-training method for a large question-answering model, capable of acquiring inference data obtained by logically splitting an original question and an original answer thereof, and constructing an inference dataset based on multiple pieces of such inference data, so as to perform post-training on the large question-answering model using the inference dataset, thereby enabling the large question-answering model to gradually develop capability of performing step-by-step reasoning on questions in a post-training process, thus fundamentally elevating the logical reasoning capability of the model.

[0063] With reference to FIG. 1, FIG. 1 is a schematic flowchart of a post-training method for a large question-answering model provided by embodiments of the present disclosure. An execution subject of the method may be, but is not limited to, a computing device such as a smartphone, a personal notebook, a personal computer, and a server. The post-training method for the large question-answering model includes steps as follows.

[0064] S101, acquiring multiple pieces of inference data.

[0065] In the present embodiment, each piece of inference data can be obtained by logically splitting an original question and an original answer thereof.

[0066] S102, constructing an inference dataset based on all the inference data.

[0067] Herein, the inference dataset is a structured collection of data configured to train and evaluate logical reasoning and multi-step problem-solving capabilities of the large question-answering model.

[0068] S103, performing post-training processing on the large question-answering model using the inference dataset.

[0069] Herein, the large question-answering model is an artificial intelligence model based on large-scale pre-training, configured to comprehend and respond to natural language from users, such as GPT.

[0070] In the present embodiment, the inference dataset is configured to guide the large question-answering model to learn to perform step-by-step reasoning on an input question. Therefore, in a post-training stage, the large question-answering model can gradually develop the capability of performing step-by-step logical reasoning on the input question.

[0071] The post-training method for the large question-answering model according to embodiments of the present disclosure firstly acquires multiple pieces of inference data, then constructs the inference dataset based on all the inference data, and finally performs the post-training processing on the large question-answering model using the inference dataset. As each piece of inference data is obtained by logically splitting an original question and an original answer thereof, the inference dataset can be configured to guide the large question-answering model to learn to perform step-by-step reasoning on the input question. The present disclosure enables the large question-answering model to develop the capability of performing step-by-step reasoning on questions, thus fundamentally improving the logical reasoning capability of the model.

[0072] Optionally, several original questions and original answers thereof can be collected, and then each original question is logically split into a plurality of sub-questions and sub-answers thereof through manual splitting or using an additional logic split model (e.g., progressive inference type). Herein, if the logic split model is employed, the sub-questions and sub-answers thereof need to be manually checked, so as to prevent logical errors.

[0073] Therefore, each piece of inference data may include an original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question, where an N-th sub-answer in the inference data is the original answer of the original question.

[0074] In an optional example, assume that the original question is “What is the result of 1+3×9+9÷2?”, the original answer is 32.5, four sub-questions and sub-answers thereof can be obtained by splitting:

[0075] First sub-question Q1: What is the result of 3×9? First sub-answer A1: 27;

[0076] Second sub-question Q2: What is the result of 9÷2? Second sub-answer A2: 4.5;

[0077] Third sub-question Q3: What is the result of 1+27? Third sub-answer A3: 28; and

[0078] Fourth sub-question Q4: What is the result of 28+4.5? Fourth sub-answer A4: 32.5.

[0079] This example is only a simple instance of splitting a logical reasoning question, and is not limited herein.

[0080] In an optional implementation mode, if a process of the large question-answering model developing the capability of performing logical reasoning is split, the large question-answering model may be expected to progressively develop capabilities of progressive question decomposition, progressive sub-question reasoning, and complete step-by-step logical reasoning. Moreover, in some scenes, a logical question that needs to be answered by the large question-answering model may be that a complete question and some sub-questions of the complete question are given, and missing intermediate questions and answers thereof need to be deduced. Thus, the inference dataset constructed in the present disclosure may include four subsets: a first subset, a second subset, a third subset and a fourth subset, respectively constructed by four different methods.

[0081] With reference to FIG. 2, a process of "constructing an inference dataset based on all the inference data" in the above step S102 may include sub-steps S1021–S1024 as follows.

[0082] S1021, obtaining at least one first sample for each piece of the inference data using a progressive reasoning type construction rule based on the inference data.

[0083] In the present embodiment, the first subset includes all first samples corresponding to each piece of the inference data. That is, each piece of inference data is corresponding to at least one first sample in the first subset.

[0084] Herein, the first subset is obtained by a first construction method. The first construction method adopts the progressive reasoning type construction rule. Specifically, for each piece of reasoning sample, a process of "obtaining at least one first sample using a progressive reasoning type construction rule based on the inference data" in step S1021 may include sub-steps S10211 and S10212 as follows.

[0085] S10211, concatenating a first prompt text, and the original question and a first sub-question in the inference data to obtain a first one of first sample, and setting a label of the first one of first sample as a first sub-answer in the inference data.

[0086] S10212, concatenating the first prompt text, and the original question, an n-th sub-question, and each sub-question preceding the n-th sub-question and sub-answers thereof in the inference data to obtain an n-th first sample, and setting a label of the n-th first sample as an n-th sub-answer in the inference data, where n = 2, 3, …, N.

[0087] In the present embodiment, the preset first prompt text may be configured to instruct the large question-answering model to infer an answer to a last sub-question in an input content

based on the input content. In the above, an actual text content of the first prompt text is not limited herein.

[0088] S1022, obtaining at least one second sample for each piece of the inference data using a progressive decomposition type construction rule based on the inference data.

[0089] In the present embodiment, the second subset includes all second samples corresponding to each piece of the inference data. That is, each piece of inference data is corresponding to at least one second sample in the second subset.

[0090] Herein, the second subset is obtained by a second construction method. The second construction method adopts the progressive decomposition type construction rule. Specifically, for each piece of reasoning sample, a process of "obtaining at least one second sample using a progressive decomposition type construction rule based on the inference data" in step S1022 may include sub-steps S10221–S10222 as follows.

[0091] S10221, concatenating a second prompt text, and the original question and the first sub-question in the inference data to obtain a first second sample, and setting a label of the first second sample as a second sub-answer in the inference data.

[0092] S10222, concatenating the second prompt text, and the original question, an m-th sub-question and each preceding sub-question in the inference data to obtain an m-th second sample, and setting a label of the m-th second sample as an (m+1)-th sub-question in the inference data, where m = 2, 3, …, N-1.

[0093] In the present embodiment, the preset second prompt text may be configured to instruct the large question-answering model to obtain a sub-question immediately succeeding the last sub-question in the input content by decomposing the original question in the input content based on the input content. In the above, an actual text content of the second prompt text is not limited herein.

[0094] S1023, obtaining at least one third sample for each piece of the inference data using a mid-segment extraction type construction rule based on the inference data.

[0095] In the present embodiment, the third subset includes all third samples corresponding to each piece of the inference data. That is, each piece of inference data is corresponding to at least one third sample in the third subset.

[0096] Herein, the third subset is obtained by a third construction method. The third construction method adopts the mid-segment extraction type construction rule. Specifically, for each piece of reasoning sample, a process of "obtaining at least one third sample using a mid-segment extraction type construction rule based on the inference data" in step S1023 may include sub-steps S10231–S10232 as follows.

[0097] S10231, concatenating a third prompt text, and each sub-question and sub-answer in the inference data, excluding the first sub-question and the first sub-answer, to obtain a first third sample, and setting a label of the first third sample as the first sub-question and the first sub-answer in the inference data.

[0098] S10232, concatenating the third prompt text, and each sub-question and sub-answer in the inference data, excluding an m-th sub-question and an m-th sub-answer, to obtain an m-th third sample, and setting a label of the m-th third sample as the m-th sub-question and the m-th sub-answer in the inference data, where m = 2, 3, …, N-1.

[0099] In the present embodiment, the third prompt text may be configured to instruct the large question-answering model to infer missing intermediate sub-questions and intermediate sub-answers thereof in the input content based on the input content. In the above, an actual text content of the third prompt text is not limited herein.

[0100] S1024, obtaining at least one fourth sample for each piece of the inference data using a step-by-step reasoning type construction rule based on the inference data.

[0101] In the present embodiment, the fourth subset includes all fourth samples corresponding to each piece of the inference data. That is, each piece of inference data is corresponding to at least one fourth sample in the fourth subset.

[0102] In the above, the fourth subset is obtained by a fourth construction method. The fourth construction method adopts the step-by-step reasoning type construction rule. Specifically, for each piece of reasoning sample, a process of "obtaining at least one fourth sample using a step-by-step reasoning type construction rule based on the inference data" in step S1024 may include sub-steps S10241–S10242 as follows.

[0103] S10241, concatenating a fourth prompt text, and the original question in the inference data to obtain a first fourth sample, and setting a label of the first fourth sample as each sub-question and sub-answer thereof in the inference data.

[0104] S10242, concatenating the fourth prompt text, and the original question, each sub-question preceding the n-th sub-question and sub-answers thereof in the inference data to obtain an n-th fourth sample, and setting a label of the n-th fourth sample as each sub-question subsequent to the (n-1)-th sub-question and sub-answers thereof in the inference data, where n = 2, 3, …, N.

[0105] In the present embodiment, the preset fourth prompt text may be configured to instruct the large question-answering model to infer subsequent sub-questions and sub-answers step by step based on the input content. In the above, an actual text content of the fourth prompt text is not limited herein.

[0106] In an optional example, assume that one piece of inference data includes an original question, five sub-questions (Q1–Q5) and five sub-answers (A1–A5) corresponding to the five sub-questions in a one-to-one manner. The inference data is processed by the four construction modes, respectively, thus obtaining samples as follows.

[0107] (1) By the first construction method, five first samples listed in following Table 1 can be obtained, where prompt1 represents the first prompt text.

Table 1

No.	First Sample	Label
1	prompt1+original question+Q1	A1
2	prompt1+original question+(Q1+A1)+Q2	A2
3	prompt1+original question+(Q1+A1)+(Q2+A2)+Q3	A3
4	prompt1+original question+(Q1+A1)+(Q2+A2)+(Q3+A3)+Q4	A4
5	prompt1+original question+(Q1+A1)+(Q2+A3)+(Q3+A3)+(Q4+A4)+Q5	A5

[0108] (2) By the second construction method, four second samples listed in following Table 2 can be obtained, where prompt2 represents the second prompt text.

Table 2

No.	Second Sample	Label
1	prompt2+original question+Q1	Q2
2	prompt2+original question+Q1+Q2	Q3
3	prompt2+original question+Q1+Q2+Q3	Q4
4	prompt2+original question+Q1+Q2+Q4	Q5

[0109] (3) By the third construction method, four third samples listed in following Table 3 can be obtained, where prompt3 represents the third prompt text.

Table 3

No.	Third Sample	Label
1	prompt3+original question+(Q2+A2)+(Q3+A3)+(Q4+A4)+(Q5+A5)	Q1+A1
2	prompt3+original question+(Q1+A1)+(Q3+A3)+(Q4+A4)+(Q5+A5)	Q2+A2
3	prompt3+original question+(Q1+A1)+(Q2+A2)+(Q4+A4)+(Q5+A5)	Q3+A3
4	prompt3+original question+(Q1+A1)+(Q2+A2)+(Q3+A3)+(Q5+A5)	Q4+A4

[0110] (4) By the fourth construction method, five fourth samples listed in following Table 4 can be obtained, where prompt4 represents the fourth prompt text.

Table 4

No.	Fourth Sample	Label
1	prompt4+original question	(Q1+A1)+(Q2+A2)+(Q3+A3)+(Q4+A4)+(Q5+A5)
2	prompt4+original question+(Q1+A1)	(Q2+A2)+(Q3+A3)+(Q4+A4)+(Q5+A5)
3	prompt4+original question+(Q1+A1)+(Q2+A2)	(Q3+A3)+(Q4+A4)+(Q5+A5)
4	prompt4+original question+(Q1+A1)+(Q2+A2)+(Q3+A3)	(Q4+A4)+(Q5+A5)
5	prompt4+original question+(Q1+A1)+(Q2+A2)+(Q3+A3)+(Q4+A4)	Q5+A5

[0111] The above example is only an instance, and the embodiments of the present disclosure do not limit the number of sub-questions in one piece of inference data.

[0112] Based on four subsets in the inference dataset, sub-steps of the above step S103 may include S1031 or S1032.

[0113] S1031, performing the post-training processing on the large question-answering model using the first subset, the second subset, the third subset and the fourth subset, respectively.

[0114] S1032, performing the post-training processing on the large question-answering model using the second subset, the first subset, the fourth subset and the third subset in sequence.

[0115] In conjunction with contents in the four tables of the above example, it can be seen that the first subset can be configured to guide the large question-answering model to learn to answer the input question step by step, the second subset can be configured to guide the large question-answering model to learn to logically decompose the input question, the third subset can be configured to guide the large question-answering model to learn to complement a missing intermediate process based on a step-by-step reasoning mode, and the fourth subset can be configured to guide the large question-answering model to learn to answer the input question step by step following logical decomposition of the input question.

[0116] Hence, the post-training processing can be performed on the large question-answering model using the four subsets in any sequential order. It is also feasible to perform the post-training processing on the large question-answering model using the second subset, the first subset, the fourth subset and the third subset in sequence, so that the large question-answering model is enabled to gradually develop capabilities of: logical decomposition, step-by-step answering, step-by-step answering following logical decomposition and complementing the missing intermediate process based on the step-by-step reasoning mode. Alternatively, it is also

feasible to perform the post-training processing on the large question-answering model using some of the first subset, the second subset, the third subset and the fourth subset, for example, to perform the post-training processing on the large question-answering model using the first subset, the second subset and the third subset are used in sequence, so as to enable the large question-answering model to develop the capabilities of: step-by-step answering, logical decomposition and complementing the missing intermediate process based on the step-by-step reasoning mode.

[0117] Optionally, a process of performing the post-training processing may use Supervised Fine-Tuning (SFT), Instruction Tuning, etc. Herein, a specific process of the post-training is the prior art, and is not repeated herein.

[0118] It should be noted that, an execution sequence of various steps in the above method embodiments is not limited by the drawings, and the execution sequence of various steps is based on an actual application situation.

[0119] Compared with the prior art, the embodiments of the present disclosure have the following beneficial effects.

[0120] The present disclosure constructs four subsets by four construction methods respectively to perform the post-training processing on the large question-answering model, thereby essentially enhancing the reasoning capability of the model.

[0121] In the present disclosure, the first subset is configured to guide the large question-answering model to learn to answer the input question step by step, thus enabling the large question-answering model to develop the capability of answering a latest split sub-question based on the original question and sub-questions obtained from previous splitting and sub-answers thereof.

[0122] In the present disclosure, the second subset can be configured to guide the large question-answering model to learn to logically decompose the input question, thus enabling the large question-answering model to develop capability of further splitting the original question and split sub-questions to obtain next sub-question.

[0123] In the present disclosure, the third subset can be configured to guide the large question-answering model to learn to complement the missing intermediate process based on the step-by-step reasoning mode, thus enabling the large question-answering model to develop capability of step-by-step reasoning to answer reasoning questions missing in intermediate steps.

[0124] In the present disclosure, the fourth subset can be configured to guide the large question-answering model to learn to answer step by step following logical decomposition on

the input question, thus enabling the large question-answering model to develop capability of performing complete step-by-step logical reasoning on the input question.

[0125] In order to execute corresponding steps in the above method embodiments and various possible embodiments, an implementation mode of a post-training apparatus for a large question-answering model is given below.

[0126] With reference to FIG. 3, FIG. 3 is a structural schematic view of the post-training apparatus for the large question-answering model provided by embodiments of the present disclosure. The post-training apparatus 200 for the large question-answering model includes: a data acquisition module 210, a construction module 220 and a post-training module 230.

[0127] The data acquisition module 210 is configured to acquire multiple pieces of inference data, where the inference data is obtained by logically splitting an original question and an original answer thereof.

[0128] The construction module 220 is configured to construct an inference dataset based on all the inference data.

[0129] The post-training module 230 is configured to perform post-training processing on the large question-answering model using the inference dataset, where the inference dataset is configured to guide the large question-answering model to learn to perform step-by-step reasoning on an input question.

[0130] Optionally, the inference dataset includes a first subset, a second subset, a third subset and a fourth subset. The construction module 220, when being configured to construct the inference dataset based on all the inference data, may be specifically configured to: obtain at least one first sample for each piece of the inference data using a progressive reasoning type construction rule based on the inference data, where the first subset includes all first samples corresponding to each piece of the inference data; obtain at least one second sample for each piece of the inference data using a progressive decomposition type construction rule based on the inference data, where the second subset includes all second samples corresponding to each piece of the inference data; to obtain at least one third sample for each piece of the inference data using a mid-segment extraction type construction rule based on the inference data, where the third subset includes all third samples corresponding to each piece of the inference data; and to obtain at least one fourth sample for each piece of the inference data using a step-by-step reasoning type construction rule based on the inference data, where the fourth subset includes all fourth samples corresponding to each piece of the inference data.

[0131] Optionally, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question.

[0132] The construction module 220, in a process of being configured to obtain at least one first sample using a progressive reasoning type construction rule based on the inference data, may be specifically configured to:

[0133] concatenate a first prompt text, and the original question and a first sub-question in the inference data to obtain the first one of first sample, and set a label of the first one of first sample as a first sub-answer in the inference data, where the first prompt text is configured to instruct the large question-answering model to infer an answer to a last sub-question in an input content based on the input content; and

[0134] concatenate the first prompt text, and the original question, an n-th sub-question, and each sub-question preceding the n-th sub-question and sub-answers thereof in the inference data to obtain an n-th first sample, and set a label of the n-th first sample as an n-th sub-answer in the inference data, where n = 2, 3, …, N.

[0135] Optionally, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question.

[0136] The construction module 220, in a process of being configured to obtain at least one second sample using a progressive decomposition type construction rule based on the inference data, may be specifically configured to:

[0137] concatenate a second prompt text, and the original question and the first sub-question in the inference data to obtain a first second sample, and set a label of the first second sample as a second sub-answer in the inference data, where the second prompt text is configured to instruct the large question-answering model to obtain a sub-question immediately succeeding the last sub-question in the input content by decomposing the original question in the input content based on the input content; and

[0138] concatenate the second prompt text, and the original question, an m-th sub-question and each preceding sub-question in the inference data to obtain an m-th second sample, and set a label of the m-th second sample as an (m+1)-th sub-question in the inference data, where m = 2, 3, …, N-1.

[0139] Optionally, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question.

[0140] The construction module 220, in a process of being configured to obtain at least one third sample using a mid-segment extraction type construction rule based on the inference data, may be specifically configured to:

[0141] concatenate a third prompt text, and each sub-question and sub-answer in the inference data, excluding the first sub-question and the first sub-answer, to obtain a first third sample, and set a label of the first third sample as the first sub-question and the first sub-answer in the inference data, where the third prompt text is configured to instruct the large question-answering model to infer missing intermediate sub-questions and intermediate sub-answers thereof in the input content based on the input content; and concatenate the third prompt text, and each sub-question and sub-answer in the inference data, excluding an m-th sub-question and an m-th sub-answer, to obtain an m-th third sample, and set a label of the m-th third sample as the m-th sub-question and the m-th sub-answer in the inference data, where m = 2, 3, …, N-1.

[0142] Optionally, the inference data include the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question.

[0143] The construction module 220, in a process of being configured to obtain at least one fourth sample using a mid-segment extraction type construction rule based on the inference data, may be specifically configured to:

[0144] concatenate a fourth prompt text, and the original question in the inference data to obtain a first fourth sample, and set a label of the first fourth sample as each sub-question and sub-answer thereof in the inference data, where the fourth prompt text is configured to instruct the large question-answering model to infer subsequent sub-questions and sub-answers step by step based on the input content; and

[0145] concatenate the fourth prompt text, and the original question, each sub-question preceding the n-th sub-question and sub-answers thereof in the inference data to obtain an n-th fourth sample, and set a label of the n-th fourth sample as each sub-question subsequent to the (n-1)-th sub-question and sub-answers thereof in the inference data, where n = 2, 3, …, N.

[0146] Optionally, the inference dataset includes a first subset, a second subset, a third subset and a fourth subset.

[0147] The post-training module 230, in a process of being configured to perform post-training processing on the large question-answering model using the inference dataset, may be specifically configured to:

[0148] perform the post-training processing on the large question-answering model using the first subset, the second subset, the third subset and the fourth subset, respectively, where the first subset is configured to guide the large question-answering model to learn to answer the input question step by step, the second subset is configured to guide the large question-answering model to learn to logically decompose the input question, the third subset is configured to guide the large question-answering model to learn to complement a missing intermediate process based on a step-by-step reasoning mode, and the fourth subset is configured to guide the large question-answering model to learn to answer the input question step by step following logical decomposition of the input question.

[0149] It would be clear to those skilled in the art that, for the convenience and brevity of the description, reference can be made to corresponding processes in the preceding method embodiments for specific working processes of the post-training apparatus 200 for the large question-answering model described in the above, and details will not be repeated herein.

[0150] With reference to FIG. 4, FIG. 4 is a structural schematic view of an electronic device provided by embodiments of the present disclosure. The electronic device 300 includes a processor 310, a memory 320 and a bus 330, where the processor 310 is connected to the memory 320 through the bus 330.

[0151] The memory 320 may be configured to store a software program, for example, a software program corresponding to the post-training apparatus 200 for the large question-answering model as provided by the embodiments of the present disclosure. The processor 310 executes various function applications and data processing by running the software program stored in the memory 320, so as to implement the post-training method for the large question-answering model as provided by embodiments of the present disclosure.

[0152] The memory 320 herein may be, but is not limited to, RAM (Random Access Memory), ROM (Read Only Memory), FLASH (flash memory), PROM (Programmable Read-Only Memory), EPROM (Erasable Programmable Ready-Only Memory), EEPROM (Electric Erasable Programmable Read-Only Memory), etc.

[0153] The processor 310 may be an integrated circuit chip capable of processing signals. The processor 310 may be a general-purpose processor, including CPU (Central Processing Unit), NP (Network Processor), SoC (System on Chip), etc., and also may be DSP (Digital Signal Processing), ASIC (Application Specific Integrated Circuit), FPGA (Field-Programmable Gate Array) or other programmable logic devices, discrete gates, transistor logic devices, or discrete hardware components.

[0154] It may be understood that the structure shown in FIG. 4 is merely illustrative, and the electronic device 300 may further include more or fewer components than those shown in FIG. 4, or is configured in a way different from that shown in FIG. 4. Various components shown in FIG. 4 may be realized by hardware, software or a combination thereof.

[0155] Embodiments of the present disclosure further provide a computer-readable storage medium, where the computer-readable storage medium stores a computer program, and the computer program, when executed by the processor, implements the post-training method for the large question-answering model disclosed in the above embodiments. The computer-readable storage medium may be, but is not limited to, various media that can store program codes, such as USB flash disk, mobile hard disk, ROM, RAM, PROM, EPROM, EEPROM, FLASH diskette or compact disk.

[0156] To sum up, embodiments of the present disclosure provide a post-training method and apparatus for a large question-answering model, an electronic device and a storage medium. Firstly, multiple pieces of inference data are acquired; then the inference dataset is constructed based on all the inference data; and finally the post-training processing is performed on the large question-answering model using the inference dataset. As each piece of inference data is obtained by logically splitting the original question and the original answer thereof, the inference dataset can be configured to guide the large question-answering model to learn to perform step-by-step reasoning on the input question. The present disclosure enables the large question-answering model to develop the capability of performing step-by-step reasoning on questions, thereby fundamentally improving the logical reasoning capability of the model.

[0157] The post-training method and apparatus for the large question-answering model, the related electronic device, the computer-readable storage medium, and the computer program product provided by some embodiments of the present description are introduced in detail in the above. The present description applies specific examples to illustrate the principle and embodiments of the present technical solutions, and the description of the above embodiments is only intended to help understanding the method of the technical solutions and the core idea thereof. Meanwhile, for those skilled in the art, according to the idea of the present technical solutions, there will be changes in the specific embodiments and the application range. In summary, the contents in the present description should not be construed as limitation to the present technical solutions.

INDUSTRIAL APPLICABILITY

[0158] Through the above solutions, in the process of training the large question-answering model, by acquiring multiple pieces of inference data, constructing the inference dataset based on all the inference data, and finally performing the post-training processing on the large question-answering model using the inference dataset, as each piece of inference data is obtained by logically splitting the original question and the original answer thereof, the inference dataset can be configured to guide the large question-answering model to learn to perform step-by-step reasoning on the input question, the large question-answering model is enabled to develop the capability of performing step-by-step reasoning on questions, thereby fundamentally improving the logical reasoning capability of the model.

CLAIMS

1. A post-training method for a large question-answering model, comprising steps of:

acquiring multiple pieces of inference data;

obtaining at least one first sample for each piece of the inference data using a progressive reasoning type construction rule based on the inference data, wherein a first subset comprises all first samples corresponding to each piece of the inference data;

obtaining at least one second sample for each piece of the inference data using a progressive decomposition type construction rule based on the inference data, wherein a second subset comprises all second samples corresponding to each piece of the inference data;

obtaining at least one third sample for each piece of the inference data using a mid-segment extraction type construction rule based on the inference data, wherein a third subset comprises all third samples corresponding to each piece of the inference data;

obtaining at least one fourth sample for each piece of the inference data using a step-by-step reasoning type construction rule based on the inference data, wherein a fourth subset comprises all fourth samples corresponding to each piece of the inference data; and

performing post-training processing on the large question-answering model using the inference dataset, wherein the inference dataset is configured to guide the large question-answering model to learn to perform step-by-step reasoning on an input question, wherein the inference dataset comprises: the first subset, the second subset, the third subset and the fourth subset, wherein the first subset is configured to guide the large question-answering model to learn to answer the input question step by step, the second subset is configured to guide the large question-answering model to learn to logically decompose the input question, the third subset is configured to guide the large question-answering model to learn to complement a missing intermediate process based on a step-by-step reasoning mode, and the fourth subset is configured to guide the large question-answering model to learn to answer step by step following logical decomposition on the input question,

wherein the inference data comprises the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question; and the step of obtaining at least one third sample using a mid-segment extraction type construction rule based on the inference data comprises:

concatenating a third prompt text, and each sub-question and sub-answer in the inference data, excluding a first sub-question and a first sub-answer, to obtain a first third sample, and setting a

label of the first third sample as a first sub-question and a first sub-answer in the inference data, wherein the third prompt text is configured to instruct the large question-answering model to infer missing intermediate sub-questions and intermediate sub-answers thereof in an input content based on the input content; and

concatenating the third prompt text, and each sub-question and sub-answer in the inference data, excluding an m-th sub-question and an m-th sub-answer, to obtain an m-th third sample, and setting a label of the m-th third sample as the m-th sub-question and the m-th sub-answer in the inference data, where m = 2, 3, …, N-1.

2. The method according to claim 1, wherein the step of obtaining at least one first sample using a progressive reasoning type construction rule based on the inference data comprises:

concatenating a first prompt text, and the original question and a first sub-question in the inference data to obtain a first one of first sample, and setting a label of the first one of first sample as a first sub-answer in the inference data, wherein the first prompt text is configured to instruct the large question-answering model to infer an answer to a last sub-question in an input content based on the input content; and

concatenating the first prompt text, and the original question, an n-th sub-question, and each sub-question preceding the n-th sub-question and sub-answers thereof in the inference data to obtain an n-th first sample, and setting a label of the n-th first sample as an n-th sub-answer in the inference data, wherein n = 2, 3, …, N.

3. The method according to claim 1, wherein the step of obtaining at least one second sample using a progressive decomposition type construction rule based on the inference data comprises:

concatenating a second prompt text, and the original question and a first sub-question in the inference data to obtain a first second sample, and setting a label of the first second sample as a second sub-answer in the inference data, wherein the second prompt text is configured to instruct the large question-answering model to obtain a sub-question immediately succeeding a last sub-question in an input content by decomposing the original question in the input content based on the input content; and

concatenating the second prompt text, and the original question, an m-th sub-question and each preceding sub-question in the inference data to obtain an m-th second sample, and setting a

label of the m-th second sample as an (m+1)-th sub-question in the inference data, wherein m = 2, 3, …, N-1.

4. The method according to claim 1, wherein the step of obtaining at least one fourth sample using a step-by-step reasoning type construction rule based on the inference data comprises:

concatenating a fourth prompt text, and the original question in the inference data to obtain a first fourth sample, and setting a label of the first fourth sample as each sub-question and sub-answer thereof in the inference data, wherein the fourth prompt text is configured to instruct the large question-answering model to infer subsequent sub-questions and sub-answers step by step based on an input content; and

concatenating the fourth prompt text, and the original question, each sub-question preceding an n-th sub-question and sub-answers thereof in the inference data to obtain an n-th fourth sample, and setting a label of the n-th fourth sample as each sub-question subsequent to an (n-1)-th sub-question and sub-answers thereof in the inference data, wherein n = 2, 3, …, N.

5. The method according to claim 1, wherein the step of performing post-training processing on the large question-answering model using the inference dataset comprises:

performing the post-training processing on the large question-answering model using the first subset, the second subset, the third subset and the fourth subset, respectively.

6. A post-training apparatus for a large question-answering model, comprising:

a data acquisition module, configured to acquire multiple pieces of inference data, wherein the inference data is obtained by logically splitting an original question and an original answer thereof;

a construction module, configured to:

obtain at least one first sample for each piece of the inference data using a progressive reasoning type construction rule based on the inference data, wherein a first subset comprises all first samples corresponding to each piece of the inference data;

obtain at least one second sample for each piece of the inference data using a progressive decomposition type construction rule based on the inference data, wherein a second subset comprises all second samples corresponding to each piece of the inference data;

obtain at least one third sample for each piece of the inference data using a mid-segment extraction type construction rule based on the inference data, wherein a third subset comprises all third samples corresponding to each piece of the inference data;

obtain at least one fourth sample for each piece of the inference data using a step-by-step reasoning type construction rule based on the inference data, wherein a fourth subset comprises all fourth samples corresponding to each piece of the inference data; and

a post-training module, configured to perform post-training processing on the large question-answering model using the inference dataset, wherein the inference dataset is configured to guide the large question-answering model to learn to perform step-by-step reasoning on an input question, wherein the inference dataset comprises: the first subset, the second subset, the third subset and the fourth subset, wherein the first subset is configured to guide the large question-answering model to learn to answer the input question step by step, the second subset is configured to guide the large question-answering model to learn to logically decompose the input question, the third subset is configured to guide the large question-answering model to learn to complement a missing intermediate process based on a step-by-step reasoning mode, and the fourth subset is configured to guide the large question-answering model to learn to answer step by step following logical decomposition on the input question,

wherein the inference data comprises the original question, N logically hierarchical progressive sub-questions obtained by splitting the original question and sub-answers of each sub-question; and the construction module, when configured to obtain at least one third sample using a mid-segment extraction type construction rule based on the inference data, is further configured to:

concatenate a third prompt text, and each sub-question and sub-answer in the inference data, excluding a first sub-question and a first sub-answer, to obtain a first third sample, and set a label of the first third sample as a first sub-question and a first sub-answer in the inference data, wherein the third prompt text is configured to instruct the large question-answering model to infer missing intermediate sub-questions and intermediate sub-answers thereof in an input content based on the input content; and

concatenate the third prompt text, and each sub-question and sub-answer in the inference data, excluding an m-th sub-question and an m-th sub-answer, to obtain an m-th third sample, and set a label of the m-th third sample as the m-th sub-question and the m-th sub-answer in the inference data, where m = 2, 3, …, N-1.

7. An electronic device, comprising: a memory and a processor, wherein the memory stores a software program, and when the electronic device runs, the processor executes the software program so as to implement the post-training method for the large question-answering model according to any one of claims 1–5.

8. A computer-readable storage medium, wherein the computer-readable storage medium stores a computer program, and the computer program, when executed by the processor, implements the post-training method for the large question-answering model according to any one of claims 1–5.