



(51) International Patent Classification:
G06F 3/00 (2006.01)

(21) International Application Number:
PCT/CN2013/091197

(22) International Filing Date:
31 December 2013 (31.12.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201310246467.X 20 June 2013 (20.06.2013) CN

(71) Applicant: **TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED** [CN/CN]; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN).

(72) Inventors: **WANG, Fei**; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN). **XU, Hao**; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen,

Guangdong 518044 (CN). **ZHU, Pan**; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN). **HAN, Guiping**; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN). **LIAO, ling**; Room 403, East Block 2, SEG Park, Zhenxing Road, Futian District, Shenzhen, Guangdong 518044 (CN).

(74) Agent: **ADVANCE CHINA IP LAW OFFICE**; Room 3901, No.85 Huacheng Avenue, Tianhe District, Guangzhou, Guangdong 510623 (CN).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM,

[Continued on next page]

(54) Title: PROCESSING METHOD AND DEVICE OF THE USER INPUT INFORMATION

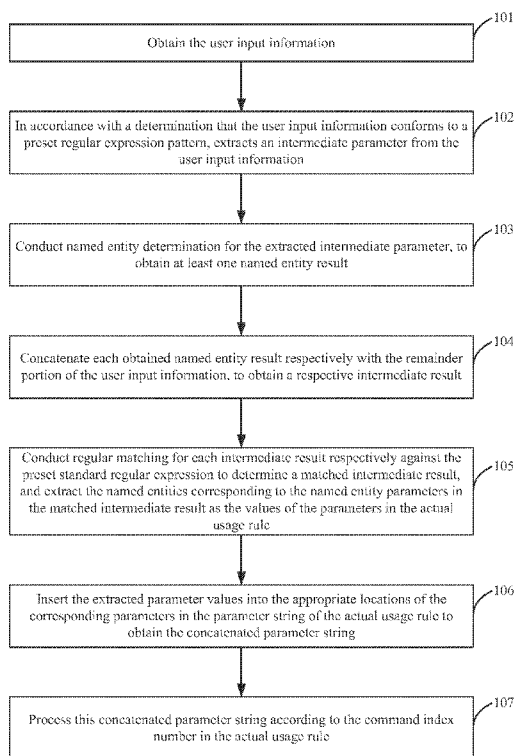


Fig. 1

(57) Abstract: A method and device for processing user input, including: storing a respective usage rule for processing a user request for a task, the respective usage rule comprising a respective rule string, a respective command index number, and a respective parameter string; performing a first transformation on the respective rule string of the respective usage rule to obtain a respective regular expression pattern; performing a second transformation on the respective rule string of the respective usage rule to obtain a respective standard regular expression; obtaining a user input that matches the respective rule string of the respective usage rule; obtaining a concatenated parameter string using the respective regular expression pattern and the respective standard regular expression; and executing a command corresponding to the respective command index number using the concatenated parameter string.



TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— with international search report (Art. 21(3))

PROCESSING METHOD AND DEVICE OF THE USER INPUT INFORMATION

RELATED APPLICATIONS

[0001] This application claims priority to Chinese Patent Application Serial No. CN201310246467.X, entitled “Method and Device of Processing User Input Information”, filed June 20, 2013, the entirety of which is incorporated herein by reference.

FIELD OF THE INVENTION

[0002] The present invention relates to the communication technical field, in particular, to a method and device for processing user input information.

BACKGROUND OF THE INVENTION

[0003] With the rapid development of communication technology, the types and functions of terminal applications are becoming more and more diversified. The terminal application for determining user intent is one of such applications. In these terminal applications, for example, a speech processing application can receive the user input information (e.g., a textual string or a speech input), identify user intent according to the user input information, and then provide corresponding service to the user according to the determined user intent. In other words, through using these terminal applications, a user can interact with a computer using a natural language to a certain extent, realizing man-machine interaction.

[0004] In conventional technology, in general, user intent may be identified by matching the obtained user input information with one or more regular rules. The so called regular rule is also referred to as the regular expression rule, which is a filter rule (or matching rule) set based on a regular expression of accepted inputs. However, a regular expression refers to a single character string which is used to describe or match a character string meeting a certain syntactic rule. When matching the user input information, in general, parameters in the user input information can be extracted at fixed positions. Therefore, for scenarios where fixed input statements are used, in general, relevant parameters can be correctly extracted and user intention can be identified.

[0005] During the research and practice of conventional technology, it is discovered that even though the conventional scheme can identify user intent well for the scenarios in which fixed statements are used, however, for the scenarios in which diverse statements may be used, the real user intent cannot be easily identified.

SUMMARY

[0006] In this disclosure, a method and device for processing user input information with the goal of identifying the correct user intent for performing an appropriate action on a computing device are disclosed. In one aspect, multiple intermediate results containing different combinations of the

named entities recognized from the user input are obtained in accordance with a regular expression pattern derived from a respective usage rule representing a task or user intent. The correct named entity parameters of the task or user intent are identified by comparing a standard regular expression and each intermediate result to identify a matched intermediate result, where the standard regular expression is also derived from the respective usage rule representing the task or user intent. The regular expression pattern and the standard regular expression are obtained from the usage rule using different transformations, such that the standard regular expression can capture the regular expression pattern when the regular expression pattern is filled with the correct named entities recognized from the user input information. By using different transformations to obtain the regular expression pattern and the standard regular expression for the same rule string of the usage rule, the matching between the user input and the rule string can be made more flexible, increasing the likelihood of a matching success.

[0007] In one aspect, a method of processing user input information includes: at a device having one or more processors and memory: storing a respective usage rule for processing a user request for a task, the respective usage rule comprising a respective rule string, a respective command index number, and a respective parameter string; performing a first transformation on the respective rule string of the respective usage rule to obtain a respective regular expression pattern; performing a second transformation on the respective rule string of the respective usage rule to obtain a respective standard regular expression; obtaining a user input that matches the respective rule string of the respective usage rule; obtaining a concatenated parameter string using the respective regular expression pattern and the respective standard regular expression; and executing a command corresponding to the respective command index number using the concatenated parameter string. In some embodiments, obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further comprises: in accordance with a determination that the user input matches the respective rule string of the respective usage rule, extracting a respective intermediate parameter from the user input in accordance with the respective regular expression pattern, the extracting leaving a reminder portion of the user input; performing named entity recognition for the respective intermediate parameter to obtain at least one named entity result; generating at least one intermediate result by concatenating the recognized named entity result with the remainder portion of the user input; performing matching analysis between the at least one intermediate result and the respective standard regular expression to determine a matched named entity result; and obtaining actual parameter values for the concatenated parameter string based on one or more respective named entities identified in accordance with the matched named entity result.

[0008] In another aspect, a device comprises one or more processors, memory, and one or more program modules stored in the memory and configured for execution by the one or more

processors. The one or more program modules include instructions for performing the method described above. In another aspect, a non-transitory computer readable storage medium having stored thereon instructions, which, when executed by a device, cause the device to perform the method described above.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The aforementioned features and advantages of the invention as well as additional features and advantages thereof will be more clearly understood hereinafter as a result of a detailed description of preferred embodiments when taken in conjunction with the drawings.

[0010] Fig. 1 is a flowchart diagram of a method of processing user input information in accordance with some embodiments;

[0011] Fig. 2 is a flowchart diagram of a method of processing user input information in accordance with some embodiments;

[0012] Fig. 3 is a flowchart diagram of a method of processing user input information in accordance with some embodiments;

[0013] Fig. 4A is a structural schematic diagram of a device for processing user input information in accordance with some embodiments;

[0014] Fig. 4B is a structural schematic diagram of a device for processing user input information in accordance with some embodiments;

[0015] Fig. 5 is a structural schematic diagram of a device in accordance with some embodiments.

[0016] Like reference numerals refer to corresponding parts throughout the several views of the drawings.

DESCRIPTION OF EMBODIMENTS

[0017] A method and device for processing user input information in accordance with some embodiments is provided with respect to Fig. 1.

[0018] In some embodiments, the device performing the processing of the user input information is a user device such as a smart phone, a tablet device, an ebook reader, an MP3 player (Moving Picture Experts Group Audio Layer III), an MP4 (Moving Picture Experts Group Audio Layer IV) player, a laptop, a desk computer and so on. In some embodiments, the device performing the processing of the user input information can be a remote server. In some embodiments, some steps of the method can be performed on a set of interacting devices, such as a client device and a server device.

[0019] In some embodiments, the method of processing the user input information includes: obtaining the user input information; when it is determined that the user input information satisfies a respective preset regular expression pattern (e.g., a regular expression pattern derived from a rule string of a particular usage rule associated with a task or user intent using a first transformation),

extracting an intermediate parameter from the user input information (e.g., extracting the intermediate parameter according to a first transformation of the rule string), performing named entity recognition on the extracted intermediate parameter, obtaining at least one named entity result from the intermediate parameter, concatenating each named entity result respectively with the remainder portion of the user input information to obtain a respective intermediate result, conducting regular expression matching for each intermediate result respectively against a preset standard regular expression (e.g., a standard regular expression derived from the rule string of the usage rule using a second transformation) to determine a matched intermediate result, obtaining the named entities corresponding to the parameters in the matched intermediate result as the actual values of the named entity parameters in the parameter string of the usage rule, inserting the actual values of the named entity parameters into the parameter string in the usage rule to obtain a concatenated parameter string, and processing this concatenated parameter string according to a command identified by the command index number in the usage rule.

[0020] As shown in Fig. 1, the detailed process can be as follows:

[0021] 101, a device obtains the user input information.

[0022] For example, the user input information can be a textual string received from a user, or a result of speech-to-text processing of a speech input received from the user. For example, a user may provide a natural language speech input using a microphone installed on a user device, and the user input information is a textual string obtained through speech-to-text processing of the natural language speech input at the user device or at a remote server. In some embodiments, the user may enter a textual string (e.g., "Find the nearest McDonald's") in a graphical user interface provided on a user device. In some embodiments, the device performing the processing of the user input information is the user device that receives the user input directly from the user. In some embodiments, the device performing the processing of the user input information is a processing device other than the user device directly capturing the user's input. In some embodiments, the device performing the processing of the user input information is a remote server.

[0023] 102, the device, in accordance with a determination that the user input information obtained by Step 101 conforms to a preset regular expression pattern (e.g., the regular expression pattern obtained by transforming the rule string of a particular usage rule), extracts an intermediate parameter from the user input information.

[0024] In some embodiments, this preset regular expression pattern is established according to an actual usage rule, where this actual usage rule includes a rule string (e.g., a string pattern representing an acceptable or recognizable user input for requesting a particular task or expressing a particular user intent, such as "Find a <business type> nearby," or "Look up the <transportation type> schedule for <Time>," etc.), a command index number, and a parameter string (e.g., a

parameter string listing the parameters needed by a command referenced by the command index number).

[0025] For example, for the task or user intent for inquiring about the trains from one city to another city, the actual usage rule can be represented concretely as:

[0026] %p28(到|去)%p28(火车|列车|车次) 10020 1|%s1|%s2|3|%s

[0027] In the above usage rule, “%p28(到|去)%p28(火车|列车|车次)” is the rule string. The rule string is an abstraction of multiple possible expressions in Chinese for saying “trains from one place to another place.” A literal English translation of the above Chinese rule string is %p28(“arrive”|“go”)%p28(expression 1 for the term “train”| expression 2 for the term “train”| expression 3 for the term “train”). The Chinese rule string reflects the correct grammar for placing the names of the start and the destination of the train in the remainder of the rule string to make up the correct expression in Chinese expression for “trains from one place to another place.” The parameter “%p28” is a placeholder for inserting the names of the starting and destination places, respectively. An analogous rule string in English would be “(trains | train schedule | schedule for trains|) from %p28 to %p28”, in accordance with some embodiments.

[0028] In the above usage rule, “10020” is the command index number for the command processing a task corresponding to looking up trains from one city to another city. For example, in some embodiments, each task that can be performed by the device has a respective command index number, and a respective command for starting a set of processing steps to perform the requested task. The processing steps take the parameters as specified by the parameter string of the usage rule.

[0029] In the above usage rule, “1|%s1|%s2|3|%s” is a parameter string for the processing steps identified by the command index number. The parameter values of the parameter string will be filled out with the actual parameter values subsequently obtained from the user input information. In this example, the parameters “1” or “3” in the above parameter string are fixed values. The parameters %s1, %s2, and %s will be replaced with corresponding actual parameter values derived from the user input information, before the parameter string is provided to the process represented by the command index number 10020. The parameter string containing the actual parameter values rather than the parameter identifiers (e.g., %s1, %s2, and %s) is referred to as a concatenated parameter string.

[0030] In some embodiments, before obtaining the user input information (i.e. Step 101), the method of processing the user input information includes:

[0031] Obtaining a predetermined actual usage rule, where this actual usage rule includes a respective rule string, a respective command index number, and a respective parameter string; and converting the rule string in this actual usage rule to a regular expression pattern, and convert the rule string in this actual usage rule to a standard regular expression. For example, in some embodiments, the device performs the following steps:

[0032] (1) To convert the rule string in this actual usage rule to a regular expression pattern:

[0033] The device converts the content between the first parameter and the last parameter in this rule string to a first identifier, where the first identifier is used to indicate a position for extracting a parameter. For example, this first identifier can be "%s1", specifically, in some embodiments.

[0034] In some embodiments, the identifier "%pxx" in a rule string is a named entity placeholder, which is also called a named entity position parameter or a parameter in the rule string. Taking the rule string "%p28(到|去)%p28(火车|列车|车次)" as an example, the identifier "%p28" in the rule string of the actual usage rule is a parameter in the rule string. If the first identifier is "%s1", the device can convert the content between the first "%pxx" and the last "%pxx" in this rule string to "%s1", namely, the device can convert the content between the two "%p28" (including the two "%p28") to "%s1", as "%s1(火车|列车|车次)".

[0035] In the above example, the regular expression "%s1(火车|列车|车次)" is the regular expression pattern obtained from the rule string "%p28(到|去)%p28(火车|列车|车次)" through the first transformation described above.

[0036] (2) To convert the rule string in this actual usage rule to a standard regular expression:

[0037] The device separately adds a second identifier immediately before and immediately after each parameter of the rule string. In some embodiments, this second identifier is used to identify a parameter that can be used to match any arbitrary character string. For example, this second identifier can be ".*", specifically, in some embodiments.

[0038] For example, still taking the rule string "%p28(到|去)%p28(火车|列车|车次)" as an example. If the second identifier is ".*", the device adds ".*" in front of and after each %pxx in the rule string. In other words, the device adds ".*" before and after each "%p28" in the rule string to obtain a standard regular expression ".*%p28.*(到|去).%p28.*(火车|列车|车次)". This standard regular expression is a more general representation of the rule string.

[0039] It should be noted that, there may exist many pieces of regular expression patterns in a system. The regular expression patterns may be stored in respective vectors. In addition, the regular expression patterns with different priorities can be stored in different vectors. In some embodiments, the regular expression pattern with the higher priorities shall be processed earlier, and the regular expression patterns with the same priority level shall be processed in turn according to the preconfigured order by which they appear. In some embodiments, the Step 102 can be performed as follows:

[0040] The device determines the current regular expression pattern from many pieces of regular expression patterns according to an order of descending order of priorities, determines whether the user input information meets the current regular expression pattern; if the user input

information meets the current regular expression pattern, the device extracts an intermediate parameter from the user input information, and if the user input information does not meet the current regular expression pattern, the device moves on to process the next regular expression pattern.

[0041] In some embodiments, to extract the intermediate parameter from the user input information, the device extracts the portion of the user input information that corresponds to the first identifier in the regular expression pattern converted from the rule string. After the intermediate parameter is extracted from the user input information, the remainder portion of the user input information is stored for subsequent use.

[0042] 103, the device conducts named entity determination for the extracted intermediate parameter, to obtain at least one named entity result. For example, in some embodiments, the device performs matching for the extracted intermediate parameter against preset named entities in a named entity library, and obtains at least one named entity result matching this extracted intermediate parameter. For example, the device may recognize the intermediate parameter as one or more alternative sets of one or more named entities, each set of one or more name entities serving as a respective named entity result.

[0043] In some embodiments, a named entity is the name of a person, the name of an institution, a place name, and other names that are used to identify entities. In some embodiments, entities in a broader sense can also include numbers, dates, currency, addresses, and so on.

[0044] In some embodiments, the named entity library can be pre-established by maintainers. In some embodiments, the named entity library can also be self-built after the system collects resources from the Internet, and before the system obtains the user input information (i.e. Step 101). In some embodiments, the method can also include: building a named entity library, where this named entity library can be stored using a Double-Array Trie.

[0045] 104, the device concatenates each obtained named entity result respectively with the remainder portion of the user input information, to obtain a respective intermediate result, the mentioned remainder portion is the remainder of the user input information after the intermediate parameter has been extracted. Since the device has obtained at least one named entity result, the device will obtain an equal number of intermediate results.

[0046] 105, the device conducts regular matching (e.g., matching or filtering using the standard regular expression) for each intermediate result respectively against the preset standard regular expression (i.e., the standard regular expression derived from the usage rule that produced the regular expression pattern matching the user input information) to determine a matched intermediate result, and extracts the named entities corresponding to the named entity parameters in the matched intermediate result as the values of the parameters in the actual usage rule.

[0047] In some embodiments, the standard regular expression is derived by converting the rule string in the actual usage rule, e.g., as described in Step 102.

[0048] 106, the device inserts the extracted parameter values into the appropriate locations of the corresponding parameters in the parameter string of the actual usage rule to obtain the concatenated parameter string.

[0049] 107, the device processes this concatenated parameter string according to the command index number in the actual usage rule. For example, in some embodiments, the processing includes:

[0050] Calling the corresponding command word according to the command index number in the actual usage rule, and using this command word to process the concatenated parameter string.

[0051] From the above, the method includes: obtaining user input information; determining whether the user input information meets a preset regular expression pattern; if it is determined that the user input information meets the preset regular expression pattern; extracting an intermediate parameter from the user input information; conducting named entity determination for the intermediate parameter; obtaining at least one named entity result from the intermediate parameter; concatenating each named entity result respectively with the remainder of the user input information to obtain a respective intermediate result; conducting regular matching for each intermediate result respectively with a preset standard regular expression of the usage rule that produced the regular expression pattern to confirm a matched intermediate result; obtaining the named entities corresponding to the named entity parameters in the matched intermediate result as the actual parameter values for the parameter string of the actual usage rule; concatenate the actual parameter values with the parameter string of the actual usage rule to obtain the concatenated parameter string; and processing this concatenated parameter string according to the command index number in the actual usage rule to realize the purpose of identifying the user intent expressed in the user input information.

[0052] Because when matching the obtained user input information, the method described herein can determine the named entities from the extracted intermediate parameter, and use regular expressions to perform multiple rounds of extraction, matching, and concatenation on the user input information, the method can match the obtained user input information more flexibly and comprehensively, so as to fully obtain various possibilities meeting user intention. In other words, even for the scenarios with many flexible statements expressing the user intentions, this scheme can identify the user intention effectively, and accurately extract the relevant parameters for post-processing use.

[0053] According to the method described herein, the following will describe the process in detail by taking a concrete example, e.g., the scenario of inquiring about train schedules.

[0054] In this scenario, the user may say "查一下深圳到北京的火车" (the corresponding English translation is "Look up trains of Shenzhen to Beijing"), "深圳到北京有哪些车次" (the corresponding English translation is "What train times are there from Shenzhen to Beijing"), "深圳

到北京有没有卧铺" (the corresponding English translation is "Are there sleeping cars for Shenzhen to Beijing"), "查一下广州来这里的火车" (the corresponding English translation is "Look up trains from Guangzhou to here"), and so on. The core sentence pattern of these different input is "<火车站城市>到<火车站城市>(火车|车次|卧铺)" (the corresponding English translation is "<city of train station> to <city of train station> (trains | train times | sleeping cars)") or "<火车站城市>来这里的火车" (the corresponding English translation is "<city of train station> trains that come here"), among which, "火车站城市" (the corresponding English translation is "city of train station") is a named entity. Note that the corresponding English translations provided above do not reflect the correct Chinese grammar for the core sentence patterns. The sentence patterns for each type of task or user intent may be obtained by finding the common and inter-changeable parts of the user inputs, and substitute the equivalent parts such as entity names and alternative words with the same meaning with placeholders. For example, in some embodiments, "%pxx" is used as the named entity placeholder in the sentence pattern. In other words, "%pxx" is used as the parameter in this rule string ""%pxx" 到"%pxx" (火车|车次|卧铺)." In some embodiments, the named entity placeholder "%pxx" corresponding to "<火车站城市>" is "%p28". As a person of ordinary skills in the art can appreciate, different placeholders of the form "%pxx" can be used to represent different types of named entities. For example, "%p20" can represent school names, and "%p16" can represent capital names, etc.

[0055] The following will take the processing of user input information "查一下深圳到北京的火车" as an example to explain the method in more detail. As shown in Fig. 2, a method of processing the user input information can be as follows:

[0056] 201, the device obtains a preset actual usage rule.

[0057] In some embodiments, this actual usage rule can include a respective rule string, a respective command index number, and a respective parameter string or parameter string. In some embodiments, the actual usage rule is also referred to as a usage rule, because the usage rule is distilled from the patterns found in many actual user inputs. The usage rule includes a rule string, which can be a string including one or more parameters (e.g., "%p28" representing <火车站城市>) or words (e.g., "火车") in a sequence. In some embodiments, the rule string also includes interchangeable words (e.g., "(火车|车次|卧铺)"). For example, this actual usage rule can concretely be as follows:

[0058] %p28(到|去)%p28(火车|车次|卧铺) 10020 1|%s1|%s2|3|%s

[0059] Among which, "%p28(到|去)%p28(火车|车次|卧铺)" is the rule string part, "10020" is command index number, and "1|%s1|%s2|3|%s" is a parameter string.

[0060] As shown above, the parameter string includes parameters (e.g., “%s1”, “%s2”, and “%s”) that can be replaced with actual parameter values, and characters or words (e.g., “1” and “3”) that will not be changed when provided to a processing procedure identified by the command index number.

[0061] In some embodiments, the rule string is mainly used to subsequently convert to a regular expression pattern and a standard regular expression to match the original input string (i.e. the original user input information). The command index number is the command index number used to identify the command for processing the parameter string after this rule is matched; and the parameter string is mainly used for the subsequent parameter string concatenation using the actual parameter values (e.g., actual names of the named entities) obtained from the user input information.

[0062] In some embodiments, when concatenating the parts of the parameter string subsequently, the “%sn” in the parameter string will be replaced with the entity name corresponding to the “nth” named entity identified in the user input string. For example, “%s1” will be replaced with the name corresponding to the first named entity identified in the user input string, “%s2” will be replaced with the name corresponding to the second named entity identified in the user input string, “%s3” will be replaced the name corresponding to the third named entity identified in the user input string, and “%s” will be replaced with the whole original user input string (i.e. the original user input information). For example, if the original input string is "查一下深圳到北京的火车", among which, the name corresponding to the first named entity identified in the string is “深圳”, the name corresponding to the second named entity identified in the string is “北京”, and the concatenated parameter string is "1|深圳|北京|3|查一下深圳到北京的火车".

[0063] 202, the device converts the rule string in the actual usage rule to a regular expression pattern. For example, it can concretely be as follows:

[0064] The device converts the content between the first parameter and the last parameter in the rule string to a first identifier, where this first identifier is used to indicate an intermediate parameter extracted at the location of the first identifier. For example, this first identifier can concretely be "%s1".

[0065] Taking the rule string "%p28(到|去)%p28(火车|车次|卧铺)" as an example, if the first identifier is "%s1", the content between the first "%pxx" and last "%pxx" in this rule string is converted to "%s1", namely, the content between the two "%p28" is converted to "%s1", to obtain a regular expression pattern "%s1(火车|车次|卧铺)".

[0066] 203, the device also converts the rule string in the actual usage rule to a standard regular expression. For example, it can concretely be as follows:

[0067] The device separately adds a second identifier in front of and after each parameter in the rule string, where this second identifier is used to identify locations for parameters which can be

matched to any arbitrary character string. For example, this second identifier can concretely be a wildcard ".*".

[0068] For example, still taking the rule string "%p28(到|去)%p28(火车|车次|卧铺)" as an example, if the second identifier is ".*", ".*" is added before and after each parameter %pxx in the rule string, namely, ".*" is added before and after each "%p28" in the rule string, to obtain a standard regular expression “.*%p28.*(到|去).%p28.*(火车|车次|卧铺)”.

[0069] 204, the device obtains the user input information.

[0070] For example, if user inputs "查一下深圳到北京的火车" on the device, at this moment, the obtained user input information by the device is: "查一下深圳到北京的火车", this user input information is also called as original character string, or the original string for short.

[0071] 205, when the device determines that the obtained user input information conforms to the preset regular expression pattern (e.g., “%p28(到|去)%p28(火车|列车|车次) 10020 1|%s1|%s2|3|%s”), the device extracts an intermediate parameter from the user input information (e.g., “查一下深圳到北京的火车”).

[0072] For example, if the user input information obtained in Step 204 is "查一下深圳到北京的火车", and at this moment, because "查一下深圳到北京的火车" conforms to the regular expression pattern "%s1 (火车|车次|卧铺)" of the actual usage rule, the device can extract the intermediate parameter (e.g., “查一下深圳到北京的”) from this user input information: "查一下深圳到北京的火车".

[0073] 206, the device will conduct matching for the extracted intermediate parameter against the preset named entities in the named entity library, to obtain at least one named entity result matching the extracted intermediate parameter.

[0074] For example, in some embodiments, there are many kinds of named entities stored in the system, so the device needs conduct named entity extraction and verification from the intermediate parameter for each kind of named entities, so as to obtain all possible combinations of the extracted named entities as the named entity results. For example, the word "北京" (the corresponding English translation is “Beijing”) is not only a named entity "%p28" for “city of train station”, but also a named entity "%p16" of “capital”. At the same time, suppose "深圳到北京" is the name of a song, and belongs to a type of named entity "%p48" for “song.” When conducting named entity determination for the intermediate parameter "查一下深圳到北京的", the device will return three different named entity results:

[0075] (1) 查一下 %p28 到 %p28 的;

[0076] (2) 查一下 %p28 到 %p16 的;

[0077] (3) 查一下 %p48 的.

[0078] In the above named entity results, the actual names of the named entities identified in the intermediate parameter have been replaced by the corresponding identifier (e.g., %p28, %p16, or %p48) for the named entity type of the actual names, respectively.

[0079] In some embodiments, the named entity library can be pre-established by maintainers. In some embodiments, the named entity library can also be self-built after the system collects resources on the Internet. In some embodiments, the named entity library can be stored in a Double-Array Trie. The following will briefly explain the named entity matching method based on a Double-Array Trie.

[0080] First, the device collects and organizes the list of named entities that may be required in usage. Taking the scenario of inquiring about train schedules as an example, a "city of train station" named entity library, a "Capital name" named entity library and "Song name" named entity library are collected.

[0081] Second, the device stores these named entity libraries in the form of a Double-Array Trie. For example, the device optionally establishes a field-specific entity name library for each type of entity names, by using the named entities for the specific field as vocabulary, and storing the index for the storage location of relevant information (here, the relevant information of the named entities can be stored using any custom structure) of each named entity as the value at leaf nodes of the Trie. In some embodiments, each kind of named entities can correspond to a respective Trie.

[0082] In addition, for each sentence to be processed for named entity matching by the device (e.g., the intermediate parameter to be processed for named entity matching by the device), the device uses each of the different field-specific Double-Array Trie for the different types of named entities to conduct the match one by one. Still take the intermediate parameter "查一下深圳到北京" as an example, first, the "city of train station" Trie is used for the matching process, which returns two named entities "深圳" and "北京" (i.e., two cities with train stations). Then, the "capital name" Trie is used for the matching process, which returns a named entity of "北京" (i.e., the capital of China). Then, the "Song name" Trie is used for the matching process, which returns a named entity of "深圳到北京". When all the field-specific Tries are used in the matching process, the device sort out all the obtained named entities and their corresponding types:

[0083] "深圳" and "北京" → name of city of train station;

[0084] "北京" → Capital name; and

[0085] "深圳到北京" → Song name.

[0086] The device obtains possible combinations for all these different types of named entities in accordance with the entity names recognized in the original input string. Continue with the above example, the device will obtain the following combinations:

[0087] (1) 查一下<city of train station:深圳>到<city of train station:北京>的;

[0088] (2) 查一下<city of train station:深圳>到<capital name:北京>的;

[0089] (3) 查一下<song name:深圳到北京>的.

[0090] Therefore, after the named entity matching is finished, the device can obtain three named entity results:

[0091] (1) 查一下 %p28 到 %p28 的;

[0092] (2) 查一下 %p28 到 %p16 的;

[0093] (3) 查一下 %p48 的.

[0094] 207, the device concatenates each of the named entity results respectively with the remainder portion of the input string, to obtain a respective intermediate result, where the remainder portion is the remaining portion of the user input information after the intermediate parameter has been extracted.

[0095] For example, still take "查一下深圳到北京的火车" as an example, if in Step 206, the returned named entity results are:

[0096] (1) 查一下 %p28 到 %p28 的;

[0097] (2) 查一下 %p28 到 %p16 的;

[0098] (3) 查一下 %p48 的.

[0099] Thus, the intermediate results obtained by the concatenating are:

[00100] (1) 查一下 %p28 到 %p28 的火车;

[00101] (2) 查一下 %p28 到 %p16 的火车;

[00102] (3) 查一下 %p48 的火车.

[00103] 208, the device conducts regular expression matching for each intermediate result respectively against the pre-established standard regular expression to determine a matched intermediate result, and obtains the named entities corresponding to the named entity parameters in the matched intermediate result as the actual values of the named entity parameters.

[00104] For example, the device optionally uses a "regexec" function to conduct the regular expression matching for the three intermediate results obtained by Step 207 respectively against the pre-established standard regular expression, to find out that the first intermediate result "查一下 %p28 到 %p28 的火车" matches standard regular expression ".*%p28.*(到|去).%p28.*(火车|列车|车次)". Therefore, the device determines that intermediate result "查一下 %p28 到 %p28 的火车" is the matched intermediate result, and obtains the named entities corresponding to the named entity parameters %p28 in the matched intermediate result (namely, "深圳" and "北京") as the actual values of the parameters in the parameter string.

[00105] 209, the device inserts the obtained actual values of the parameters at the locations of the corresponding parameters in the parameter string of the usage rule to obtain the concatenated parameter string.

[00106] For example, if in Step 208, the obtained actual parameters are “深圳” and “北京” respectively, the device concatenates these actual parameters and the parameter string "1|%s1|%s2|3|%s" to obtain the concatenated parameter string: “1|深圳|北京|3|查一下深圳到北京的火车”, where %s1 is replaced with the first parameter value (i.e., the first named entity “深圳” obtained from the matched intermediate result), %s2 is replaced with the second parameter value (i.e., the second named entity “北京” obtained from the matched intermediate result), and %s is replaced with the original input string “查一下深圳到北京的火车”.

[00107] 210, the device calls the corresponding command word according to the command index number in the actual usage rule, and uses this command word to process the concatenated parameter string.

[00108] For example, take the following actual usage rule as an example:

[00109] %p28(到|去)%p28(火车|车次|卧铺) 10020 1|%s1|%s2|3|%s

[00110] At this moment, the device calls the command word corresponding to the command index number "10020", uses this command word to process the concatenated parameter string, namely, “1|深圳|北京|3|查一下深圳到北京的火车”.

[00111] In some embodiments, the method described above is performed on a single device or a server and a client device. In some embodiments, the steps for generating the respective regular expression pattern and the respective standard regular expression for each actual usage rule are optionally performed by a server, before the user input information is received. In some embodiments, the named entity libraries are optionally generated and stored by a server as well. In some embodiments, the steps of receiving the user input information, determining whether the user input information matches the respective regular expression pattern of a particular actual usage rule, extracting the intermediate parameter according to the respective regular expression pattern of the actual usage rule, performing the named entity extraction on the intermediate parameter, obtaining the intermediate results, and matching each of the intermediate results to the respective standard regular expression of the actual usage rule to find a matched intermediate result, obtaining the named entities of the parameters in the matched intermediate result, generating the concatenated parameter string using the parameter string in the actual usage rule and the obtained named entities; and calling the command according to the command index number to process the concatenated parameter string, are optionally allocated between a server and a client device according to different design configurations that are suitable in light of the respective processing capabilities of the server and the client device.

[00112] Fig. 3 is a flowchart diagram of another method for processing user input information based on the methods described above. As shown in Fig. 3, the method includes:

[00113] 301, the device stores a respective usage rule for processing a user request for a task, the respective usage rule comprising a respective rule string, a respective command index number, and a respective parameter string;

[00114] 302, the device performs a first transformation on the respective rule string of the respective usage rule to obtain a respective regular expression pattern;

[00115] 303, the device performs a second transformation on the respective rule string of the respective usage rule to obtain a respective standard regular expression;

[00116] 304, the device obtains a user input that matches the respective rule string of the respective usage rule; and

[00117] 305, the device obtains a concatenated parameter string using the respective regular expression pattern and the respective standard regular expression.

[00118] In some embodiments, the method further includes: 306, the device executes a command corresponding to the respective command index number using the concatenated parameter string.

[00119] In some embodiments, obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further includes: in accordance with a determination that the user input matches the respective rule string of the respective usage rule, extracting a respective intermediate parameter from the user input in accordance with the respective regular expression pattern, the extracting leaving a reminder portion of the user input; performing named entity recognition for the respective intermediate parameter to obtain at least one named entity result; generating at least one intermediate result by concatenating the recognized named entity result with the remainder portion of the user input; performing matching analysis between the at least one intermediate result and the respective standard regular expression to determine a matched named entity result; and obtaining actual parameter values for the concatenated parameter string based on one or more respective named entities identified in accordance with the matched named entity result.

[00120] In some embodiments, obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further includes: constructing the concatenated parameter string in accordance with the respective parameter string of the respective usage rule and the one or more respective named entities identified in accordance with the matched named entity result.

[00121] In some embodiments, performing the first transformation on the respective rule string of the respective usage rule to obtain the respective regular expression pattern further includes: identifying content between a first parameter and a second parameter in the respective rule string of

the respective usage rule; and generating the respective regular expression pattern for the respective usage rule by substituting the content between the first parameter and the second parameter with a first identifier.

[00122] In some embodiments, extracting the respective intermediate parameter from the user input in accordance with the respective regular expression pattern further includes: extracting the respective intermediate parameter from the user input at a location indicated by the first identifier.

[00123] In some embodiments, performing the second transformation on the respective rule string of the respective usage rule to obtain the respective standard regular expression further includes: identifying all parameters in the respective rule string of the respective usage rule; and generating the respective standard regular expression for the respective usage rule by inserting a second identifier before and after each parameter identified in the respective rule string of the respective usage rule.

[00124] In some embodiments, performing the matching analysis between the at least one intermediate result and the respective standard regular expression further includes: determining whether each of the at least one intermediate result matches the respective standard regular expression; and in accordance with a determination that a respective one of the at least one intermediate result matches the respective standard regular expression, using the respective intermediate result as the matched intermediate result.

[00125] In some embodiments, the method further includes: storing a plurality of usage rules each having a respective priority level; and processing the user input against respective regular expression patterns of the plurality of usage rules in an order according to descending priority levels of the plurality of usage rules to determine whether the user input matches the respective regular expression pattern of the respective usage rule.

[00126] It should be noted that, the aforementioned description is just the processing process of a single rule. It should be understood that, in a system, there can exist many pieces of regular expression patterns, many vectors can be used to store these regular expression patterns. The regular expression patterns with different priorities shall be stored in different vectors. The regular expression pattern with high priority shall be processed first, the regular expression patterns with the same priority shall be processed in turn according to the order they are configured to appear. The processing process of each usage rule is similar with the aforementioned process.

[00127] As described above, in some embodiments, the device obtains the user input information, determines when the user input information meets a preset regular expression pattern, extracts an intermediate parameter from the user input information, conducts named entity recognition in the intermediate parameter, obtains at least one named entity result, concatenate this named entity result respectively with the remainder of the input string, obtains at least one intermediate result, conducts regular expression matching for each intermediate result respectively

with the preset standard regular expression to determine the matched intermediate result and obtains the actual parameter values corresponding to named entity parameters in the matched intermediate result, concatenates these actual parameter values with the parameter string in the usage rule to obtain the concatenated parameter string, and at last, processes this concatenated parameter string according to the command index number in the usage rule to realize the purpose of identifying user intention. Because when matching the obtained user input information, this scheme can determine the named entities for the extracted intermediate parameter, and uses the regular expressions to extract, match and concatenate the user input information for many times, it can match the obtained user input information more flexibly and comprehensively, so as to fully obtain various possibilities meeting the user intention. Even for the scenarios with flexible statements, this scheme can also identify user intention effectively, and accurately extract relevant parameters for post-processing use.

[00128] In order to implement the methods above better, the embodiment of the present invention provides a device for processing user input information. As shown in Fig. 4A, this processing device includes information obtaining unit 401, extraction unit 402, determination unit 403, first concatenating unit 404, first processing unit 405, second concatenating unit 406 and second processing unit 407.

[00129] Information obtaining unit 401, configured to obtain the user input information; for example, this user input information can be the input speech information or text information received from a user.

[00130] Extraction unit 402, configured to determine when this user input information conforms to a preset regular expression pattern, and extract an intermediate parameter from the user input information;

[00131] Among which, this regular expression pattern is established according to a usage rule, this usage rule can include a respective rule string, a respective command number and a respective parameter string. For example, for the task or user intent for inquiring about the trains from one city to another city, the actual usage rule can be represented concretely as:

[00132] %p28(到|去)%p28(火车|列车|车次) 10020 1|%s1|%s2|3|%s

[00133] In the above usage rule, “%p28(到|去)%p28(火车|列车|车次)” is the rule string, “10020” is the command index number, and “1|%s1|%s2|3|%s” is a parameter string.

[00134] Determination unit 403, configured to conduct named entity determination for the extracted intermediate parameter, to obtain at least one named entity result;

[00135] First concatenating unit 404, configured to concatenate each obtained named entity result respectively with the remainder portion of the user input information, to obtain a respective intermediate result, the mentioned remainder portion is the remainder of the user input information after the intermediate parameter has been extracted;

- [00136] First processing unit 405, configured to conduct regular matching for each intermediate result respectively against the preset standard regular expression to determine a matched intermediate result, and extract the named entities corresponding to the named entity parameters in the matched intermediate result as the values of the parameters in the actual usage rule;
- [00137] In some embodiments, the standard regular expression is derived by converting the rule string in the actual usage rule.
- [00138] Second concatenating unit 406, configured to insert the extracted parameter values into the appropriate locations of the corresponding parameters in the parameter string of the actual usage rule to obtain the concatenated parameter string;
- [00139] Second processing unit 407, configured to process this concatenated parameter string according to the command index number in the actual usage rule.
- [00140] Optionally, prior to obtaining the user input information, the preset actual usage rule can also be obtained, allowing the device to convert the rule string in this actual usage rule to a regular expression pattern, and to convert the rule string in the actual usage rule to a standard regular expression. That is, as shown in Fig. 4B, the processing device can also include rule obtaining unit 408, first converting unit 409 and second converting unit 410;
- [00141] Rule obtaining unit 408, configured to obtain a preset actual usage rule.
- [00142] In some embodiments, this actual usage rule can include a respective rule string, a respective command index number, and a respective parameter string or parameter string.
- [00143] First converting unit 409, configured to convert the rule string in the actual usage rule to a regular expression pattern;
- [00144] Second converting unit 410, configured to convert the rule string in the actual usage rule to a standard regular expression.
- [00145] For example, first converting unit 409, configured specifically to convert the content between the first parameter and the last parameter in the rule string to a first identifier, where this first identifier is used to indicate an intermediate parameter extracted at the location of the first identifier. For example, this first identifier can concretely be "%s1".
- [00146] Second converting unit 410, configured to separately add a second identifier in front of and after each parameter in the rule string, where this second identifier is used to identify locations for parameters which can be matched to any arbitrary character string. For example, this second identifier can concretely be a wildcard ".*".
- [00147] In some embodiments, determination unit 403, configured specifically to conduct matching for the extracted intermediate parameter against the preset named entities in the named entity library, to obtain at least one named entity result matching the extracted intermediate parameter.

[00148] In some embodiments, a named entity is the name of a person, the name of an institution, a place name, and other names that are used to identify entities. In some embodiments, entities in a broader sense can also include numbers, dates, currency, addresses, and so on.

[00149] In some embodiments, the named entity library can be pre-established by maintainers. In some embodiments, the named entity library can also be self-built after the system collects resources from the Internet. That is, as shown in Fig. 4B, the device of processing input information can also includes establishment unit 411;

[00150] Establishment unit 411, configured to build a named entity library, where this named entity library can be stored using a Double-Array Trie.

[00151] In some embodiments, second processing unit 407, configured specifically to call the corresponding command word according to the command index number in the actual usage rule, and use this command word to process the concatenated parameter string.

[00152] It should be understood that, in a system, there can exist many pieces of regular expression patterns, many vectors can be used to store these regular expression patterns. The regular expression patterns with different priorities shall be stored in different vectors. The regular expression pattern with high priority shall be processed first, the regular expression patterns with the same priority shall be processed in turn according to the order they are configured to appear:

[00153] Extraction unit 402 is configured to determine the current regular expression pattern from many pieces of regular expression patterns according to an order of descending order of priorities, determine whether the user input information meets the current regular expression pattern; if the user input information meets the current regular expression pattern, extract an intermediate parameter from the user input information, and if the user input information does not meet the current regular expression pattern, move on to process the next regular expression pattern.

[00154] In specific implementation, each of units above can be implemented as independent entity, or as the same one or several entities through arbitrary combination.

[00155] The processing device can be integrated concretely in devices such as smart phones, tablets, e-readers, MP3 players, MP4 players, laptop portable computers, desktop computers, and so on.

[00156] From the above, information obtaining unit 401 of the processing device is used to obtain the user input information, after determining that the user input information meets a preset regular expression pattern, extracting unit 402 is used to extract an intermediate parameter from the input information of this user, and determination unit 303 is used for named entity determination to obtain at least one named entity result, after which, first concatenating unit 304 concatenates each named entity result respectively with the remainder of the input information to obtain a respective intermediate result, followed by regular expression matching for each intermediate result respectively with a preset standard regular expression by first processing unit 305 to confirm the

matched intermediate result. Based on the matched intermediate result, second concatenating unit 306 extracts the actual parameter values corresponding to the named entity parameters in the matched intermediate result and concatenates the actual parameters with the parameter string of the actual usage rule to obtain a concatenated parameter string. At last, second processing unit 307 is used to process this concatenated parameter string according to the command index number in the actual usage rule to realize the purpose of identifying user intention. Because when matching the obtained user input information, this scheme can determine the named entities for the extracted intermediate parameter, and use the regular expression pattern to extract, match, and concatenate the user input information for multiple times, it can match the obtained user input information more flexibly and comprehensively, so as to fully obtain various possibilities meeting the user intention. In other words, even for the scenarios with flexible statements, this scheme can also identify user intention effectively, and accurately extract relevant parameters for post-processing use.

[00157] Fig. 5 is a structural diagram of an exemplary device 500 for processing user input information in accordance with some embodiments. The device 500 can perform some or all of the steps described above in accordance with various embodiments. The device 500 includes radio frequency (RF, Radio Frequency) circuit 501, memory 502 comprising of one or more computer-readable storage media, input unit 503, display unit 504, sensor 505, audio circuit 506, wireless fidelity (WiFi, Wireless Fidelity) module 507, processor 508 including one or more than one processing cores, power supply 509 and other parts. In actual embodiments, a device of processing user input information may include fewer or more components than those shown in Fig. 5. Some components shown in Fig. 5 may be combined or divided into sub-components.

[00158] RF circuit 501 is configured to send and receive information, or to send and receive signal during calling, in particular, to receive downlink information of base station, followed by handing it over to one or more processors 508 for processing; in addition, sending involved uplink data to base station. Generally, RF circuit 501 including but not limited to antenna, at least one amplifier, tuner, one or more oscillators, subscriber identity module (SIM, Subscriber Identity Module) card, transceiver, coupler, low-noise amplifier (LNA, Low Noise Amplifier) and duplexer, etc. In addition, RF circuit 501 can communicate through wireless communication, network and other devices. The mentioned wireless communication may use any communication standard or protocol, including but not limited to Global System of Mobile communication (GSM, Global System of Mobile communication), General Packet Radio Service (GPRS, General Packet Radio Service), Code Division Multiple Access (CDMA, Code Division Multiple Access), Wideband Code Division Multiple Access (WCDMA, Wideband Code Division Multiple Access), Long Term Evolution (LTE, Long Term Evolution), E-mail, Short Messaging Service (SMS, Short Messaging Service), etc.

[00159] Memory 502 is also configured to store software program and module, and processor 508 then performs various function applications and data processing by running software program and module stored in memory 502. Memory 502 may mainly include stored program area and stored data area, among which, stored program area stores operation system, application program required for at least one function (such as functions described with respect to Figs. 1-4B), etc.; stored data area stores data created according to the use of the device (such as audio data, phone book, etc.) and so on. In addition, memory 502 may include high-speed random access memory, also include non-volatile memory, such as at least one disk storage device, flash memory devices, or other volatile solid state memory elements. Accordingly, memory 502 may also include memory controller, configured to provide processor 508 and input unit 503 with the access to memory 502.

[00160] Input unit 503, configured to receive the inputting of numeric or character information, and generate the signal input of keyboard, mouse, joystick, optical or trackball related to user settings and function control. Specifically, in one specific embodiment, input unit 503 includes touch-sensitive surface and other input devices. Input unit 503 also includes other input devices besides touch-sensitive surface.

[00161] Display unit 504 is configured to display the user input information or information provided for user and various graphical user interfaces of terminal, and these graphical user interface consisting of graph, text, icon, video and random combination of them.

[00162] The device also optionally includes one or more sensors 505, such as light sensor, motion sensor and other sensors.

[00163] Audio circuit 506, speaker and microphone provide audio interface between user and terminal. Audio circuit 506 can transmit the received electric signal from conversion of audio data to speaker, where electric signal is converted and output as acoustical signal; on the other hand, microphone converts the collected acoustical signal into electric signal, which is received and converted as audio data by audio circuit 506, after which, audio data is output from processor 508 for processing, followed by being sent to another terminal through RF circuit 501, or audio data can be output to memory 502 for further processing. Audio circuit 506 may also include earplug jack to provide communication between peripheral headset and the device.

[00164] WiFi module 507, which provides users with wireless access to broadband Internet.

[00165] Processor 508, acting as the control center of terminal, using various interfaces and circuits for connection of each part of the entire mobile phone, by operating or performing software program and/or module stored in memory 502, and calling the data stored in memory 502, performs various functions of terminal and processes data to realize the overall operation of the device 500. Optionally, processor 508 can include one or more processing cores, preferably, the aforementioned application processor and modem processor can be integrated in processor 508, among which, application processor mainly deals with the application processor operating system, user interface

and application program and so on, while modem processor mainly deals with the wireless communication.

[00166] Device 500 also includes power supply 509 used for the power supply of various parts (such as battery), preferably, power supply can be connected logically with processor 508 by power supply management system, through which, charging, discharging, power management and other functions can be realized. Power supply 509 can also includes one or more DC or AC power supplies, recharging system, power failure detection circuit, power converter or inverter, power supply status indicators, and any components.

[00167] In some embodiments, specifically, processor 508 in device 500, according to the following commands, loads the executable files corresponding to the processes of one or more application programs to memory 502, and processor 508 is used for running the application programs stored in memory 502 to realize various functions: including the functions described with respect to Figs. 1-4B.

[00168] Although some of the various drawings illustrate a number of logical stages in a particular order, stages that are not order dependent may be reordered and other stages may be combined or broken out. While some reordering or other groupings are specifically mentioned, others will be obvious to those of ordinary skill in the art and so do not present an exhaustive list of alternatives. Moreover, it should be recognized that the stages could be implemented in hardware, firmware, software or any combination thereof.

[00169] The foregoing description, for purpose of explanation, has been described with reference to specific embodiments. However, the illustrative discussions above are not intended to be exhaustive or to limit the invention to the precise forms disclosed. Many modifications and variations are possible in view of the above teachings. The embodiments were chosen and described in order to best explain the principles of the invention and its practical applications, to thereby enable others skilled in the art to best utilize the invention and various embodiments with various modifications as are suited to the particular use contemplated.

What is claimed is:

Claims

1. A method of processing user input, comprising:

at a device having one or more processors and memory:

storing a respective usage rule for processing a user request for a task, the respective usage rule comprising a respective rule string, a respective command index number, and a respective parameter string;

performing a first transformation on the respective rule string of the respective usage rule to obtain a respective regular expression pattern;

performing a second transformation on the respective rule string of the respective usage rule to obtain a respective standard regular expression;

obtaining a user input that matches the respective rule string of the respective usage rule; and

obtaining a concatenated parameter string using the respective regular expression pattern and the respective standard regular expression.

2. The method of claim 1, further comprising:

executing a command corresponding to the respective command index number using the concatenated parameter string.

3. The method of claim 1, wherein obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further comprises:

in accordance with a determination that the user input matches the respective rule string of the respective usage rule, extracting a respective intermediate parameter from the user input in accordance with the respective regular expression pattern, the extracting leaving a reminder portion of the user input;

performing named entity recognition for the respective intermediate parameter to obtain at least one named entity result;

generating at least one intermediate result by concatenating the recognized named entity result with the remainder portion of the user input;

performing matching analysis between the at least one intermediate result and the respective standard regular expression to determine a matched named entity result; and

obtaining actual parameter values for the concatenated parameter string based on one or more respective named entities identified in accordance with the matched named entity result.

4. The method of claim 3, wherein obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further comprises:

constructing the concatenated parameter string in accordance with the respective parameter string of the respective usage rule and the one or more respective named entities identified in accordance with the matched named entity result.

5. The method of claim 3, wherein performing the first transformation on the respective rule string of the respective usage rule to obtain the respective regular expression pattern further comprises:

identifying content between a first parameter and a second parameter in the respective rule string of the respective usage rule; and

generating the respective regular expression pattern for the respective usage rule by substituting the content between the first parameter and the second parameter with a first identifier.

6. The method of claim 5, wherein extracting the respective intermediate parameter from the user input in accordance with the respective regular expression pattern further comprises:

extracting the respective intermediate parameter from the user input at a location indicated by the first identifier.

7. The method of claim 3, wherein performing the second transformation on the respective rule string of the respective usage rule to obtain the respective standard regular expression further comprises:

identifying all parameters in the respective rule string of the respective usage rule; and

generating the respective standard regular expression for the respective usage rule by inserting a second identifier before and after each parameter identified in the respective rule string of the respective usage rule.

8. The method of claim 7, wherein performing the matching analysis between the at least one intermediate result and the respective standard regular expression further comprises:

determining whether each of the at least one intermediate result matches the respective standard regular expression; and

in accordance with a determination that a respective one of the at least one intermediate result matches the respective standard regular expression, using the respective intermediate result as the matched intermediate result.

9. The method of claim 1, further comprising:

storing a plurality of usage rules each having a respective priority level; and

processing the user input against respective regular expression patterns of the plurality of usage rules in an order according to descending priority levels of the plurality of usage rules to determine whether the user input matches the respective regular expression pattern of the respective usage rule.

10. A system of processing user input, comprising:

one or more processors; and

memory having instructions stored thereon, the instructions, when executed by the one or more processors, cause the processors to perform operations comprising:

storing a respective usage rule for processing a user request for a task, the respective usage rule comprising a respective rule string, a respective command index number, and a respective parameter string;

performing a first transformation on the respective rule string of the respective usage rule to obtain a respective regular expression pattern;

performing a second transformation on the respective rule string of the respective usage rule to obtain a respective standard regular expression;

obtaining a user input that matches the respective rule string of the respective usage rule; and

obtaining a concatenated parameter string using the respective regular expression pattern and the respective standard regular expression.

11. The system of claim 10, wherein the operations further comprise:

executing a command corresponding to the respective command index number using the concatenated parameter string.

12. The system of claim 10, wherein obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further comprises:

in accordance with a determination that the user input matches the respective rule string of the respective usage rule, extracting a respective intermediate parameter from the user input in accordance with the respective regular expression pattern, the extracting leaving a reminder portion of the user input;

performing named entity recognition for the respective intermediate parameter to obtain at least one named entity result;

generating at least one intermediate result by concatenating the recognized named entity result with the remainder portion of the user input;

performing matching analysis between the at least one intermediate result and the respective standard regular expression to determine a matched named entity result; and

obtaining actual parameter values for the concatenated parameter string based on one or more respective named entities identified in accordance with the matched named entity result.

13. The system of claim 12, wherein obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further comprises:

constructing the concatenated parameter string in accordance with the respective parameter string of the respective usage rule and the one or more respective named entities identified in accordance with the matched named entity result.

14. The system of claim 12, wherein performing the first transformation on the respective rule string of the respective usage rule to obtain the respective regular expression pattern further comprises:

identifying content between a first parameter and a second parameter in the respective rule string of the respective usage rule; and

generating the respective regular expression pattern for the respective usage rule by

substituting the content between the first parameter and the second parameter with a first identifier.

15. The system of claim 14, wherein extracting the respective intermediate parameter from the user input in accordance with the respective regular expression pattern further comprises:

extracting the respective intermediate parameter from the user input at a location indicated by the first identifier.

16. The system of claim 12, wherein performing the second transformation on the respective rule string of the respective usage rule to obtain the respective standard regular expression further comprises:

identifying all parameters in the respective rule string of the respective usage rule; and

generating the respective standard regular expression for the respective usage rule by inserting a second identifier before and after each parameter identified in the respective rule string of the respective usage rule.

17. The system of claim 16, wherein performing the matching analysis between the at least one intermediate result and the respective standard regular expression further comprises:

determining whether each of the at least one intermediate result matches the respective standard regular expression; and

in accordance with a determination that a respective one of the at least one intermediate result matches the respective standard regular expression, using the respective intermediate result as the matched intermediate result.

18. The system of claim 10, wherein the operations further comprise:

storing a plurality of usage rules each having a respective priority level; and

processing the user input against respective regular expression patterns of the plurality of usage rules in an order according to descending priority levels of the plurality of usage rules to determine whether the user input matches the respective regular expression pattern of the respective usage rule.

19. A non-transitory computer-readable medium having instructions stored thereon, the instructions, when executed by one or more processors, cause the processors to perform operations comprising:

storing a respective usage rule for processing a user request for a task, the respective usage rule comprising a respective rule string, a respective command index number, and a respective parameter string;

performing a first transformation on the respective rule string of the respective usage rule to obtain a respective regular expression pattern;

performing a second transformation on the respective rule string of the respective usage rule to obtain a respective standard regular expression;

obtaining a user input that matches the respective rule string of the respective usage rule; and

obtaining a concatenated parameter string using the respective regular expression pattern and the respective standard regular expression.

20. The computer-readable medium of claim 19, wherein obtaining the concatenated parameter string using the respective regular expression pattern and the respective standard regular expression further comprises:

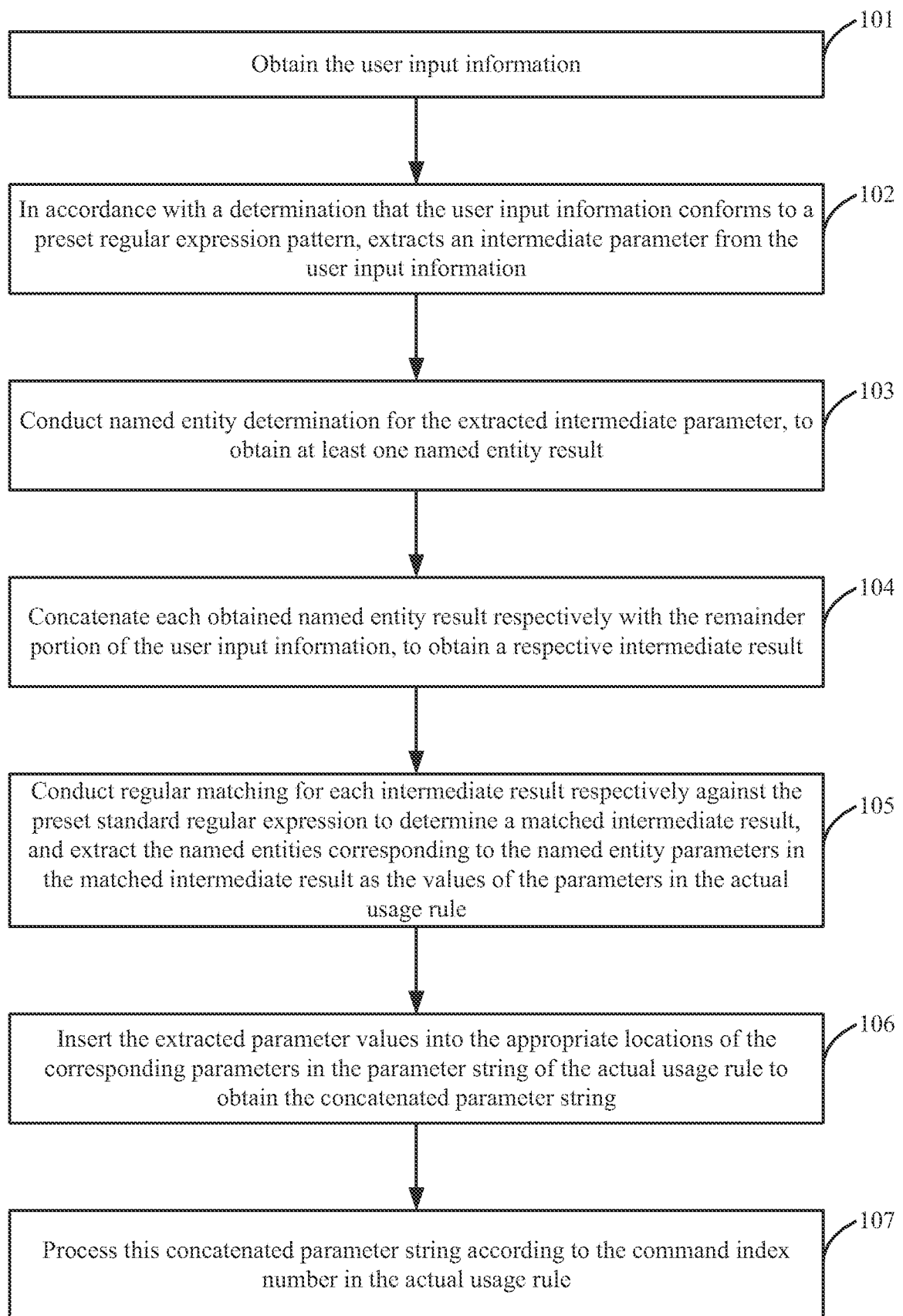
in accordance with a determination that the user input matches the respective rule string of the respective usage rule, extracting a respective intermediate parameter from the user input in accordance with the respective regular expression pattern, the extracting leaving a reminder portion of the user input;

performing named entity recognition for the respective intermediate parameter to obtain at least one named entity result;

generating at least one intermediate result by concatenating the recognized named entity result with the remainder portion of the user input;

performing matching analysis between the at least one intermediate result and the respective standard regular expression to determine a matched named entity result; and

obtaining actual parameter values for the concatenated parameter string based on one or more respective named entities identified in accordance with the matched named entity result.

**Fig. 1**

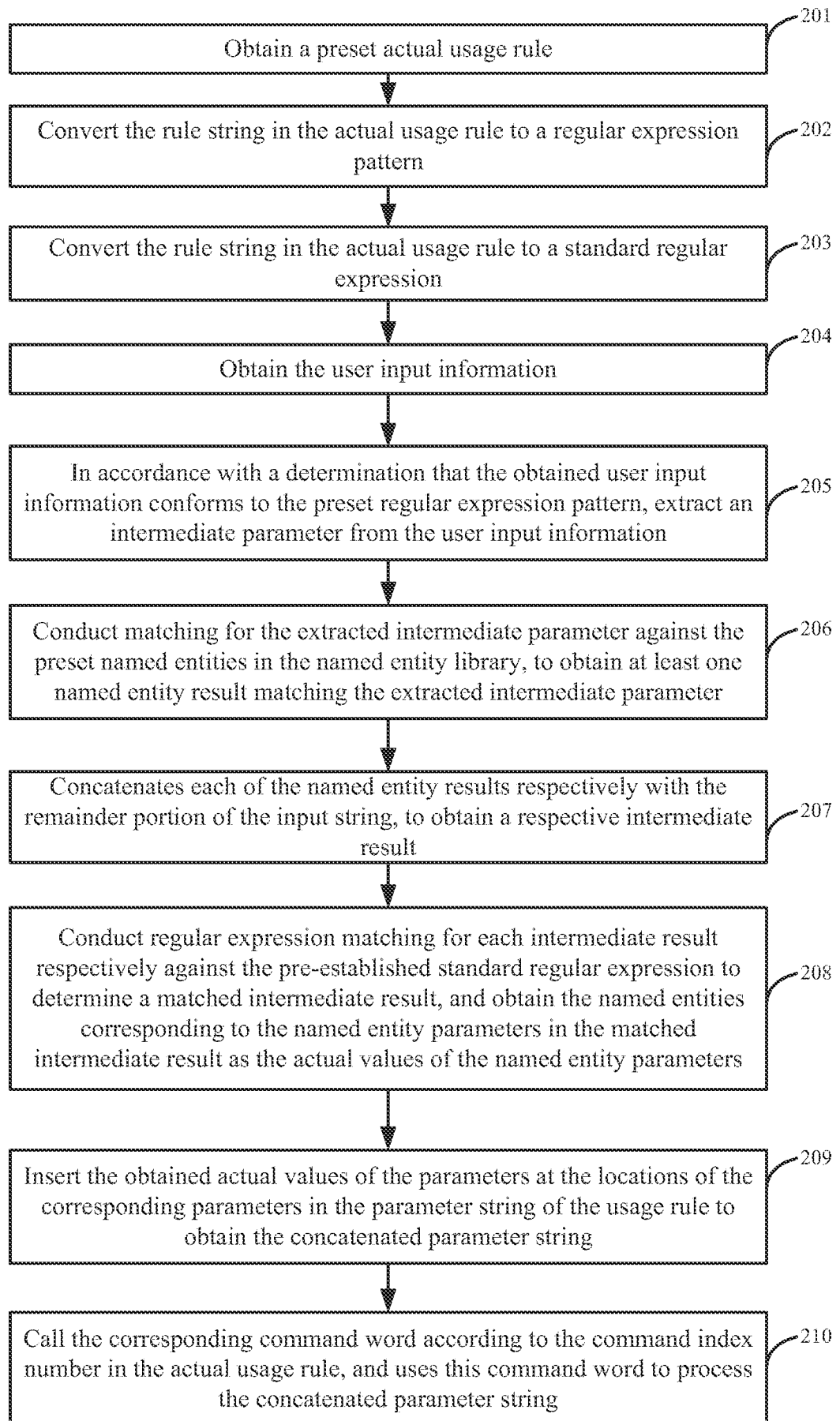
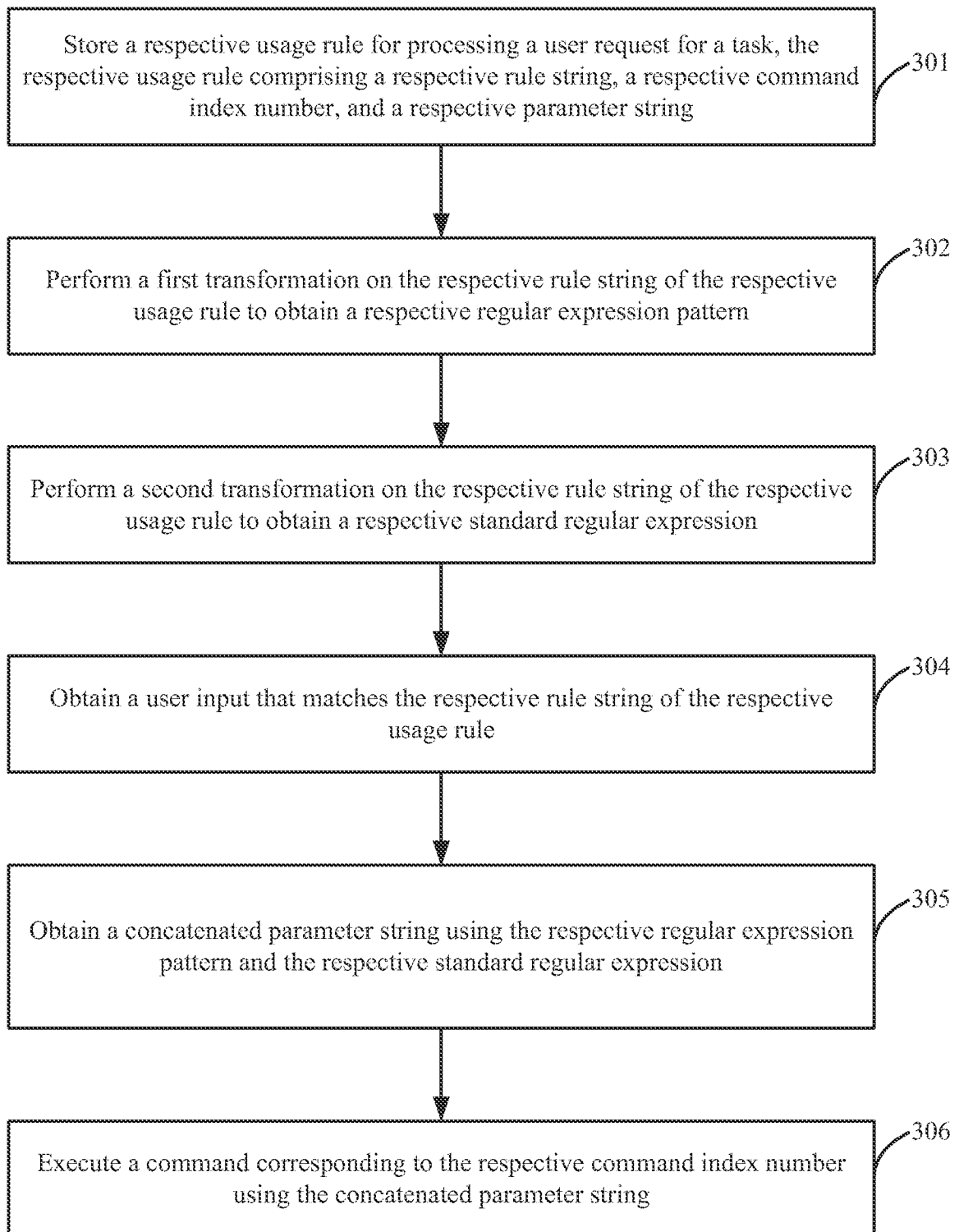
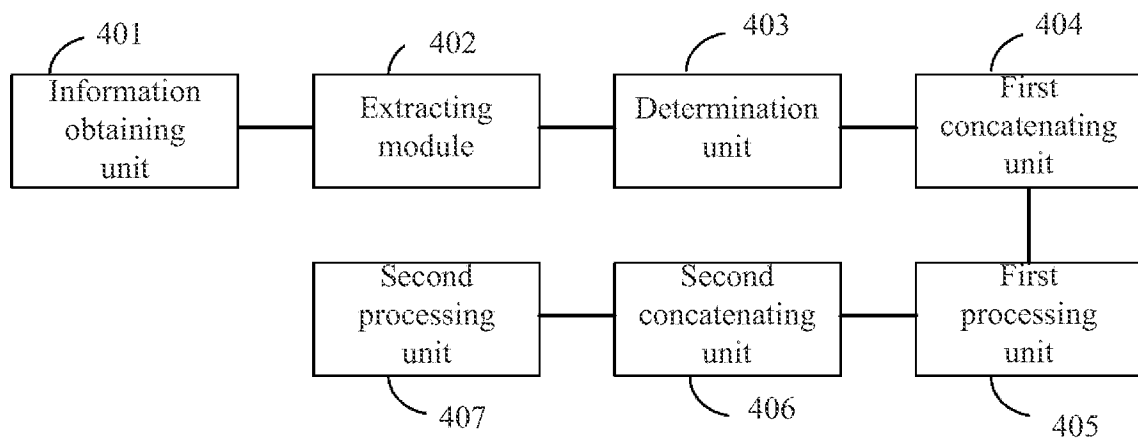
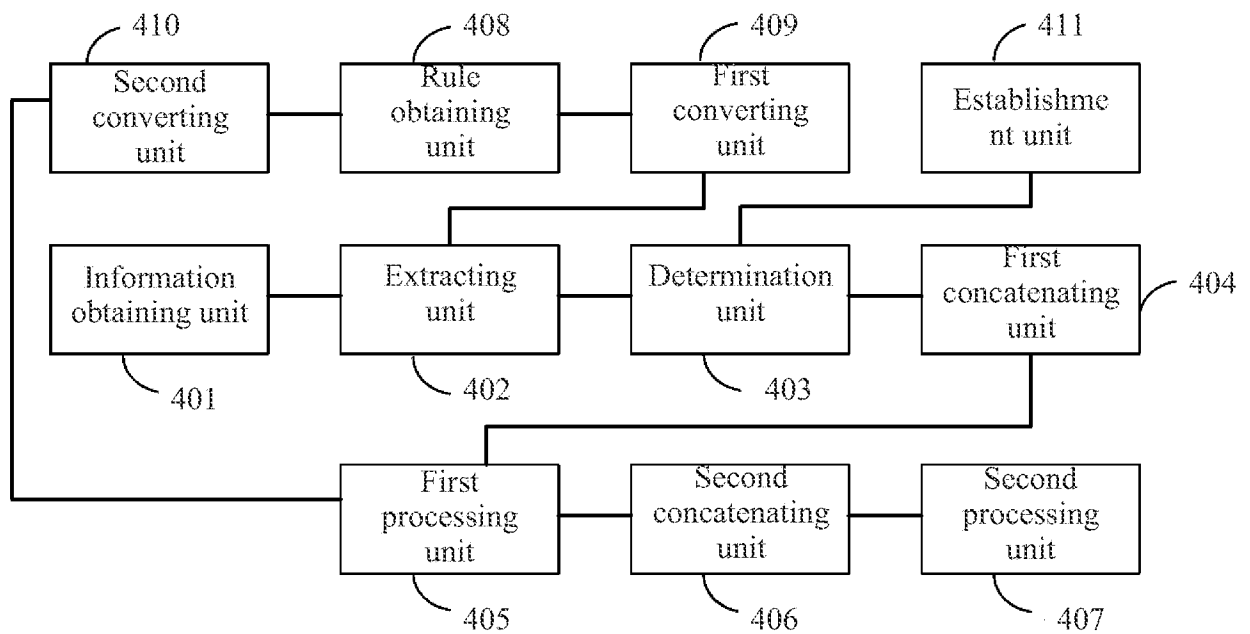
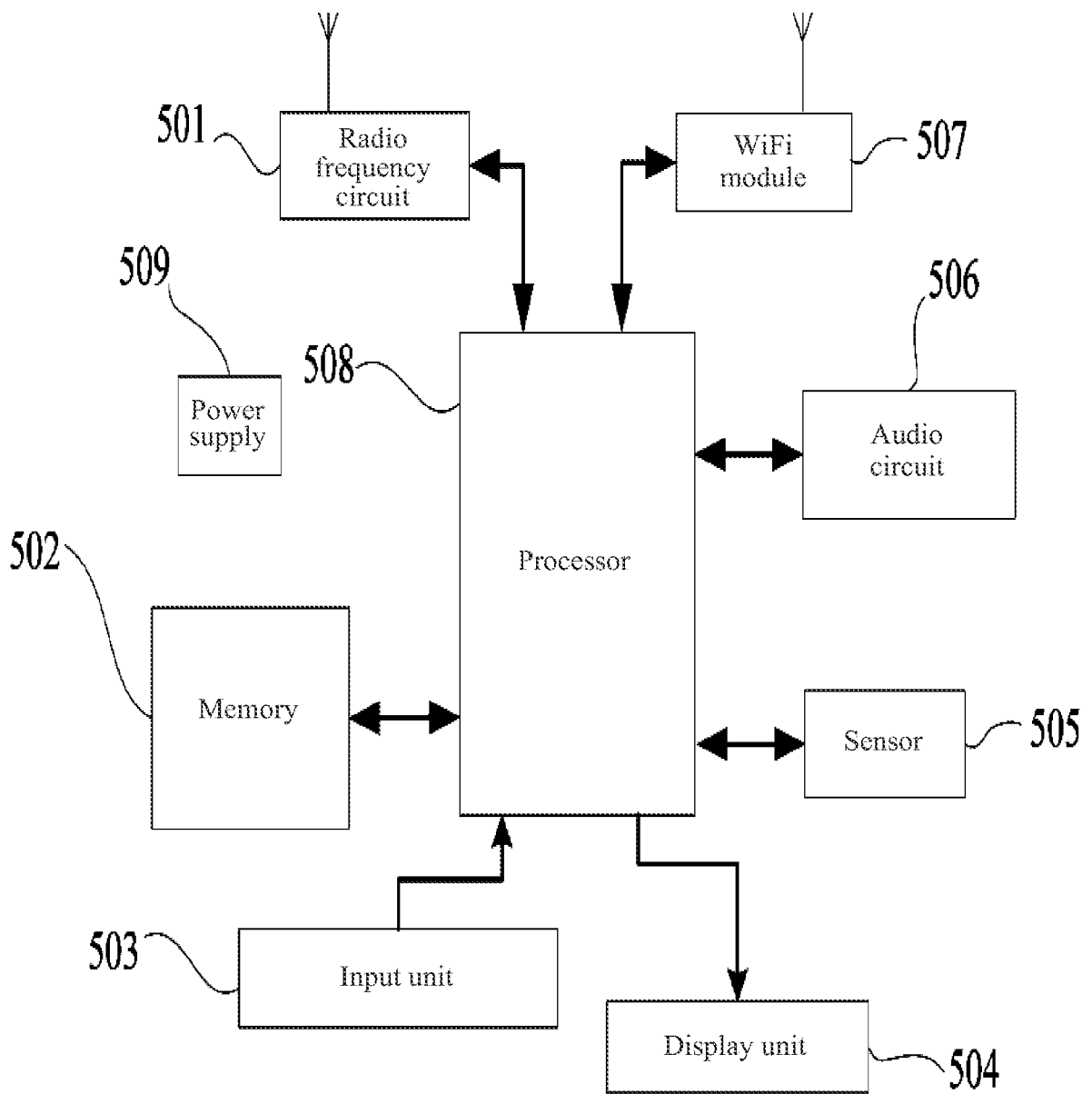


Fig. 2

**Fig. 3**

**Fig. 4A****Fig. 4B**

**Fig. 5**

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/091197

A. CLASSIFICATION OF SUBJECT MATTER

G06F 3/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F 3/-; G06F 17/-

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, CNKI, WPI, EPODOC: regular, expression, parameter, abstract, string, character, rule, standard, match, extract, identify

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2011153641 A1 (AT&T INTELLECTUAL PROPERTY I LP) 23 June 2011 (23.06.2011) the whole document	1-20
A	US 2012310648 A1 (FUJITSU LTD) 06 December 2012 (06.12.2012) the whole document	1-20
A	CN 101360088 A (HUAWEI TECHNOLOGIES CO., LTD.) 04 February 2009 (04.02.2009) the whole document	1-20

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
20 February 2014 (20.02.2014)Date of mailing of the international search report
27 Mar. 2014 (27.03.2014)Name and mailing address of the ISA/CN
The State Intellectual Property Office, the P.R.China
6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China
100088
Facsimile No. 86-10-62019451Authorized officer
YU, Chen
Telephone No. (86-10)62411716

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2013/091197

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 2011153641 A1	23.06.2011	None	
US 2012310648 A1	06.12.2012	JP 2012252594 A	20.12.2012
CN 101360088 A	04.02.2009	EP 2184687 B1	04.12.2013
		WO 2009015603 A1	05.02.2009
		US 8413124 B2	02.04.2013
		CN 101360088 B	14.09.2011
		EP 2184687 A1	12.05.2010
		US 2010131935 A1	27.05.2010