



- (51) International Patent Classification:
G06F 17/30 (2006.01)
- (21) International Application Number:
PCT/CN2011/077615
- (22) International Filing Date:
26 July 2011 (26.07.2011)
- (25) Filing Language: English
- (26) Publication Language: English
- (71) Applicant (for all designated States except US): **EMPIRE TECHNOLOGY DEVELOPMENT LLC** [US/US];
2711 Centerville Road, Suite 400, Wilmington, Delaware 19808 (US).
- (72) Inventor; and
- (75) Inventor/Applicant (for US only): **FANG, Jun** [CN/CN];
No. 127, Youyi XiLu, Xi'an, Shaanxi 710072 (CN).
- (74) Agent: **UNITALEN ATTORNEYS AT LAW**; 7th Floor,
Scitech Place, No.22, Jian Guo Men Wai Ave., Chao Yang
District, Beijing 100004 (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, 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, IS, JP, KE, KG, KM, 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, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG,
SK, SL, SM, ST, SV, SY, TH, TJ, TM, 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, SD, SL, SZ, TZ, UG,
ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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,

[Continued on next page]

(54) Title: METHOD AND SYSTEM FOR RETRIEVING INFORMATION FROM SEMANTIC DATABASE

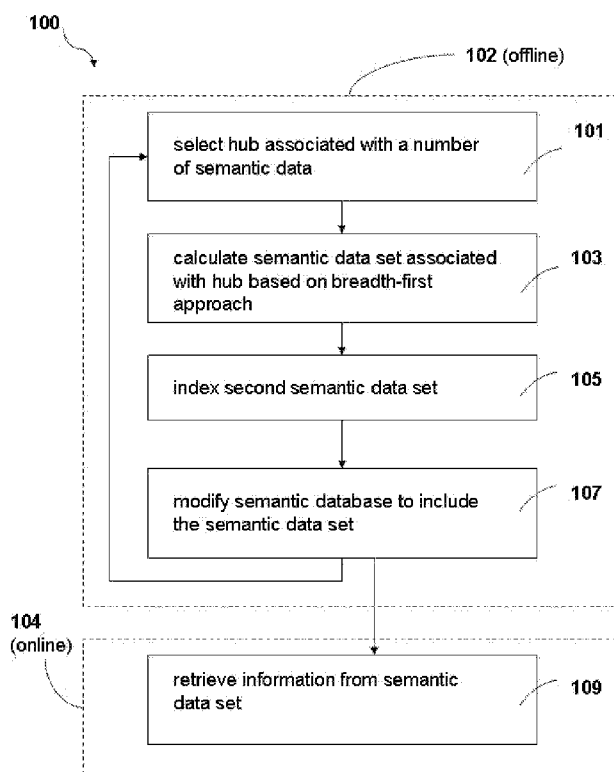


FIG. 1

(57) Abstract: Embodiments of the present disclosure set forth a method for retrieving information from a semantic database having a plurality of semantic data. The method includes selecting a hub that is resolved with a number of semantic data based on a resolution rule, when the number of the semantic data is greater than a threshold in an offline environment; calculating a semantic data set that is associated with the hub based on a searching approach and the resolution rule in the offline environment; indexing the semantic data set in the offline environment; modifying the semantic database to include the semantic data set in the offline environment; and retrieving information from the semantic data set in an online environment.



MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, **Published:**
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, — *with international search report (Art. 21(3))*
GW, ML, MR, NE, SN, TD, TG).

- 1 -

METHOD AND SYSTEM FOR RETRIEVING INFORMATION FROM SEMANTIC DATABASE

TECHNICAL FIELD

[001] The present disclosure relates generally to a semantic web and more specifically to
5 techniques for retrieving information from semantic databases.

BACKGROUND OF THE DISCLOSURE

[002] A semantic web includes various semantic data. The semantic web extends the
network of hyperlinked human-readable web pages by inserting machine-readable metadata
about pages, and how the pages are related to each other, enabling a machine to access
10 the semantic web more intelligently and perform tasks independently.

SUMMARY

[003] Some embodiments of the present disclosure may generally relate to methods for
retrieving information from a semantic database having a plurality of semantic data. One
example method may include selecting a hub that is resolved with a number of semantic
15 data based on a resolution rule, when the number of the semantic data is greater than a
threshold in an offline environment; calculating a semantic data set that is associated with
the hub based on a searching approach and the resolution rule in the offline environment;
indexing the semantic data set in the offline environment; modifying the semantic database
to include the semantic data set in the offline environment; and retrieving information from
20 the semantic data set in an online environment.

[004] Other embodiments of the present disclosure may generally relate to computer-
readable media containing instructions for retrieving information from a semantic database
having a plurality of semantic data. One example computer-readable media may contain
instructions, which when executed by a computing device, causes the computing device to
25 select a hub that is resolved with a number of semantic data based on a resolution rule,
when the number of the semantic data is greater than a threshold in an offline environment;
calculate a semantic data set that is associated with the hub based on a searching
approach and the resolution rule in the offline environment; index the semantic data set in
the offline environment; modify the semantic database to include the semantic data set in
30 the offline environment; and retrieve information from the semantic data set in an online
environment.

- 2 -

[005] Additional embodiments of the present disclosure may generally relate to computing devices configured to retrieve information from a semantic database having a plurality of semantic data. One example computing device may contain a processing unit. The processing unit is configured to select a hub that is resolved with a number of semantic data
5 based on a resolution rule, when the number of the semantic data is greater than a threshold in an offline environment; calculate a semantic data set that is associated with the hub based on a searching approach and the resolution rule in the offline environment; index the semantic data set in the offline environment; modify the semantic database to include the semantic data set in the offline environment; and retrieve information from the semantic
10 data set in an online environment.

[006] The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative aspects, embodiments, and features described above, further aspects, embodiments, and features will become apparent by reference to the drawings and the following detailed description.

15 BRIEF DESCRIPTION OF THE DRAWINGS

[007] The foregoing and other features of the present disclosure will become more fully apparent from the following description and appended claims, taken in conjunction with the accompanying drawings. Understanding that these drawings depict only several
20 embodiments in accordance with the disclosure and are therefore not to be considered limiting of its scope, the disclosure will be described with additional specificity and detail through use of the accompanying drawings.

[008] FIG. 1 is a flow chart of an illustrative embodiment of a method for retrieving information from a semantic database;

FIG. 2 shows a block diagram of an illustrative embodiment illustrating a computer
25 program product that is arranged for retrieving information from a semantic database; and

FIG. 3 shows a block diagram of an illustrative embodiment of a computing device that is arranged for retrieving information from a semantic database, all arranged in accordance with at least some embodiments of the present disclosure.

DETAILED DESCRIPTION

[009] In the following detailed description, reference is made to the accompanying drawings, which form a part hereof. In the drawings, similar symbols typically identify similar components, unless context dictates otherwise. The illustrative embodiments described in the detailed description, drawings, and claims are not meant to be limiting. Other
5 embodiments may be utilized, and other changes may be made, without departing from the spirit or scope of the subject matter presented here. It will be readily understood that the aspects of the present disclosure, as generally described herein, and illustrated in the figures, can be arranged, substituted, combined, and designed in a wide variety of different configurations, all of which are explicitly contemplated and make part of this disclosure.

10 [0010] This disclosure is drawn, *inter alia*, to methods, apparatus, computer programs, and systems of retrieving information from a semantic database having a plurality of semantic data sets.

[0011] In this disclosure, a resolution rule exists between two semantic data if they are resolved. A semantic data can be semantically translated to a First-order Logic (FOL)
15 formula, which is constructed by atomic symbols and operators using a set of rules. The operators may include, without limitation, negation $\neg$, conjunction $\wedge$, disjunction $\vee$, existential qualifier $\exists$, universal qualifier $\forall$, and etc. A FOL formula can be transformed into a standard formula if all negation operations in it are close to atomic symbols. Two
20 standard FOL formulas can be resolved if there exists an atomic symbol in the first standard formula and the negation of the atomic symbol in the second standard formula. The two standard formulas have a resolvent. Two semantic data can be resolved if their corresponding standard FOL formulas are resolved. Two semantic data are semantically equivalent if the two semantic data have the same interpretation. In this disclosure, a “hub” generally refers to a semantic data that is associated with a semantic data set based on a
25 resolution rule.

[0012] In the disclosure, a semantic database is analogized to a scale-free network. Any one of the semantic data in the semantic database may correspond to a node in the scale-free network. When two semantic data are resolved, an undirected network link may be deemed to form between the two semantic data. In a scale-free network, only few nodes
30 have a large number of links to other nodes, which are typically called hub nodes. Similarly, only few semantic data in the semantic database is resolved with a large number of semantic data. The semantic data which is resolved with a number of semantic data more than a predetermined threshold is defined as a “hub” in the semantic database.

[0013] In some embodiments, retrieving information from a semantic database having a plurality of semantic data may include offline processing operations and online processing operations. The offline processing operations may be continuously carried out in the background to compile the semantic database. In response to a query, the online
5 processing operations may include performing a reasoning task to the query and retrieving information from the compiled semantic database based on the result of the reasoning task.

[0014] The offline processing operations may include constructing a network in the semantic database and selecting hubs in the semantic database. As set forth above, a semantic data in the semantic database may be analogized to a node in a scale-free
10 network. A hub in the semantic database generally refers to a semantic data which is resolved with a number of semantic data, where the number of the semantic data resolved with the hub is greater than a predetermined threshold.

[0015] Degree of the semantic data is defined as the number of links of the semantic data in the analog scale-free network. Selection of a semantic data to be a hub in the semantic
15 database is based on the degree of the semantic data. If the degree of the semantic data is greater than the average degree of all calculated semantic data in the semantic database to a threshold (e.g., a specified times of the average degree), then the semantic data is selected as a hub. In a very big or even infinite semantic database, calculation of the average degree can be gradually approximated, depends on the computation ability and
20 time.

[0016] The offline processing operations further includes calculating a semantic data set that is associated with a hub selected as set forth above. The calculation may be based on a breadth-first method. In the first level of the breadth-first method, for a specific hub, the calculation may include calculating resolvents of the hub and the semantic data which
25 directly links to the hub based on a resolution rule. The directly linked semantic data is selected as a part of the semantic data set when it can be resolved with the hub. In the level k greater than 1, the calculation may further include calculating resolvents of the semantic data set selected in the level k-1 and the semantic data which is within a distance of k from the hub based on the resolution rule. When a semantic data within a distance of k
30 from the hub is resolved with any one of semantic data of the semantic data set, the semantic data within a distance of k from the hub is selected as a part of the semantic data set. The number of k should be greater than a specified times of the average path length between the analog nodes in the analog scale-free network. The average path length may

be gradually approximated in the very big or infinite semantic database. The calculation of the semantic data set may continually execute in the background, until there are no new implicit semantic data in the current level or the level is greater than a specified value. The implicit semantic data is the semantic data not explicit expressed in the semantic database, while can be implicitly obtained from the semantic database based on the logic reasoning, e.g., deduction, induction, and abduction. The semantic data set may be incomplete and keep growing if the semantic database is very big or infinite. Depending on the computation ability and time, the calculation of the semantic data set may gradually examine most contents of the semantic database.

10 [0017] The offline processing operations may further include compressing the semantic data set. Two or more semantic equivalent data may be mapped to a same new semantic data to decrease the size of the semantic data set.

[0018] The offline processing operations may further include indexing the compressed semantic data set. For a semantic data set associated with a hub, the semantic data in the semantic data set is indexed to quickly retrieve semantic data associated with a given symbol.

[0019] In response to a query associated with the first semantic data, the online processing operations may include locating the semantic data set through all hubs. When the query and a hub share a common symbol, it indicates the semantic data set associated with the hub is related to the query.

[0020] Unlike the conventional approaches, the online processing operations may include retrieving information from the specific semantic data set (e.g., putting the negation of the query in the specific semantic data set), rather from all semantic data in the semantic database.

25 [0021] FIG. 1 is a flow chart of an illustrative embodiment of a method 100 for retrieving information from a semantic database. The method 100 may include one or more operations, functions, or actions illustrated by blocks 101, 103, 105, 107, and 109. Although the blocks are illustrated in a sequential order, these blocks may also be performed in parallel, and/or in a different order than those described herein. Also, the various blocks may be combined into fewer blocks, divided into additional blocks, and/or eliminated based upon the desired implementation. The method 100 may begin at block 101.

[0022] At block 101, the method 100 may include selecting a hub that is resolved with a number of semantic data based on a resolution rule. The number of the semantic data may be greater than a threshold. For example, the first hub may be a politician's name. The semantic data resolve with the hub may include the various types of information associated with the politician, such as, without limitation, the politician's height, weight, relatives' information, political party, and city of birth. The politician's height may be resolved with the politician's name, because both of them include the name of the politician. The politician's name may be further resolved with any one of the weight, relatives' information, political party, and city of birth.

[0023] The number of the semantic data associated with the politician is usually much greater than the number of semantic data set associated with a less public person. However, the number of the semantic data associated with the politician may be less than the number of the semantic data associated with an even more public figure, such as a professional basketball player. The relationships among the plurality of semantic data in the semantic database may be construed in accordance with the resolution rule among them. When a semantic data is resolved with a number of a semantic data that is greater than a threshold, the semantic data (i.e., hub) is then selected. The method continues at block 103.

[0024] At block 103, the method 100 may include calculating a semantic data set that is associated with the hub. Depending on the computation ability, a number of semantic data other than the hubs may be selected and then calculated.

[0025] A new semantic data may be resolved with a hub itself or a semantic data associated with the hub. Such new semantic data may be included in the semantic data set. Then the new semantic data may be used in the calculation of the implicit semantic data. An implicit semantic data is not originally included in the semantic database, but it can be obtained from the semantic database based on the logic reasoning, such as deduction, induction, and abduction. Alternatively, a new semantic data may not be resolved with any hub or any semantic data associated with any one of the hubs. Then the new semantic data is not a part of the semantic data set.

[0026] The selection of the new semantic data may be based on a feasible searching technique, e.g., breadth-first searching or depth-first computation. A computation time interval may be preset in accordance with the computation ability. A number of new

- 7 -

semantic data will be selected within the preset computation time interval. The method continues in step 105.

[0027] The calculation of the implicit semantic data based on the selected new semantic data is carried out by any feasible technique, for example a breadth-first calculation

5 technique. Given the hub and the selected semantic data described above, in the level 1, the implicit semantic data is calculated by utilizing the resolution rule and logic reasoning between the hub and the semantic data directly resolved with the hub. In the level k (i.e., greater than 1), the calculation of the implicit semantic data may further include calculating
10 resolvers of the semantic data set calculated in the level k-1 and the semantic data which is within a distance of k from the hub based on the resolution rule.

[0028] At block 105, the semantic data set associated with any one of hubs may be compressed by keeping only one symbol for all symbols of the same semantic interpretation. Then they are indexed by any technical feasible approaches. The method continues at
15 block 107. At block 107, the indexed semantic data set may be included in the semantic database.

[0029] Blocks 101, 103, 105, and 107 may be repeated in sequence so that the contents of the semantic database may keep growing. Blocks 101, 103, 105, and 107 may be
20 continuously carried out in an offline environment. The offline environment in the disclosure generally refers to an environment in which the semantic database may be configured to substantially operate in a disconnected state.

[0030] The semantic database may receive a query. At block 109, in response to a query, the method may include retrieving information from the indexed semantic data set associated with a specific hub in an online environment. The online environment in the
25 disclosure generally refers to an environment in which the semantic database may be configured to substantially operate in a state of connectivity and interact with another party. When both of the query and a specific hub have a symbol in common, the method may include locating the semantic data set associated with the specific hub and retrieving information in response to the query from the semantic data set.

[0031] FIG. 2 shows a block diagram illustrating a computer program product that is
30 arranged for selecting a preferred data set. The computer program product 200 may include a signal bearing medium 204, which may include one or more sets of executable

- 8 -

instructions 202 that, when executed by, for example, a processor of a computing device, may provide at least the functionality described above and illustrated in FIG. 1.

[0032] In some implementations, the signal bearing medium 204 may encompass a non-transitory computer readable medium 208, such as, but not limited to, a hard disk drive, a Compact Disc (CD), a Digital Versatile Disk (DVD), a digital tape, memory, etc. In some implementations, the signal bearing medium 204 may encompass a recordable medium 210, such as, but not limited to, memory, read/write (R/W) CDs, R/W DVDs, etc. In some implementations, the signal bearing medium 204 may encompass a communications medium 206, such as, but not limited to, a digital and/or an analog communication medium (e.g., a fiber optic cable, a waveguide, a wired communications link, a wireless communication link, etc.) The computer program product 200 may also be recorded in the non-transitory computer readable medium 208 or another similar recordable medium 210.

[0033] Fig. 3 shows a block diagram of an illustrative embodiment of a computing device that is arranged for selecting a preferred data set. In a very basic configuration 301, computing device 300 typically includes one or more processors 310 and a system memory 320. A memory bus 330 may be used for communicating between processor 310 and system memory 320.

[0034] Depending on the desired configuration, processor 310 may be of any type including but not limited to a microprocessor (μ P), a microcontroller (μ C), a digital signal processor (DSP), or any combination thereof. Processor 310 may include one more levels of caching, such as a level one cache 311 and a level two cache 312, a processor core 313, and registers 314. An example processor core 313 may include an arithmetic logic unit (ALU), a floating point unit (FPU), a digital signal processing core (DSP Core), or any combination thereof. An example memory controller 315 may also be used with processor 310, or in some implementations memory controller 315 may be an internal part of processor 310.

[0035] Depending on the desired configuration, system memory 320 may be of any type including but not limited to volatile memory (such as RAM), non-volatile memory (such as ROM, flash memory, etc.) or any combination thereof. System memory 320 may include an operating system 321, one or more applications 322, and program data 324. In some embodiments, application 322 may include a data retrieving algorithm 323 that is arranged to perform the functions as described herein including those described with respect to the steps 101 and 109 of the method 100 of Fig. 1. Program data 324 may include semantic

data sets 325 that may be useful for the operation of data retrieving algorithm 323 as will be further described below. In some embodiments, the semantic data sets 325 may include, without limitation, the first semantic data set and the second semantic data set of blocks 101 and 103 of Fig. 1, respectively. In some embodiments, application 322 may be arranged to
5 operate with program data 324 on operating system 321 such that implementations of selecting preferred data set may be provided as described herein. This described basic configuration 301 is illustrated in Fig. 3 by those components within the inner dashed line.

[0036] In some other embodiments, application 322 may include data retrieving algorithm 323 that is arranged to perform the functions as described herein including those described
10 with respect to the steps 101 and 109 of the method 100 of Fig. 1.

[0037] Computing device 300 may have additional features or functionality, and additional interfaces to facilitate communications between basic configuration 301 and any required devices and interfaces. For example, a bus/interface controller 340 may be used to facilitate communications between basic configuration 301 and one or more data storage
15 devices 350 via a storage interface bus 341. Data storage devices 350 may be removable storage devices 351, non-removable storage devices 352, or a combination thereof. Examples of removable storage and non-removable storage devices include magnetic disk devices such as flexible disk drives and hard-disk drives (HDD), optical disk drives such as compact disk (CD) drives or digital versatile disk (DVD) drives, solid state drives (SSD), and
20 tape drives to name a few. Example computer storage media may include volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, program modules, or other data.

[0038] System memory 320, removable storage devices 351 and non-removable storage
25 devices 352 are examples of computer storage media. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which may be used to store the desired information and which may be accessed by computing
30 device 300. Any such computer storage media may be part of computing device 300.

[0039] Computing device 300 may also include an interface bus 342 for facilitating communication from various interface devices (e.g., output devices 360, peripheral

interfaces 370, and communication devices 380) to basic configuration 301 via bus/interface controller 340. Example output devices 360 include a graphics processing unit 361 and an audio processing unit 362, which may be configured to communicate to various external devices such as a display or speakers via one or more A/V ports 363. Example peripheral
5 interfaces 370 include a serial interface controller 371 or a parallel interface controller 372, which may be configured to communicate with external devices such as input devices (e.g., keyboard, mouse, pen, voice input device, touch input device, etc.) or other peripheral devices (e.g., printer, scanner, etc.) via one or more I/O ports 373. An example
10 communication device 380 includes a network controller 381, which may be arranged to facilitate communications with one or more other computing devices 390 over a network communication link via one or more communication ports 382. In some embodiments, the other computing devices 390 may include other applications, which may be operated based on the results of the application 322.

[0040] The network communication link may be one example of a communication media.
15 Communication media may typically be embodied by computer readable instructions, data structures, program modules, or other data in a modulated data signal, such as a carrier wave or other transport mechanism, and may include any information delivery media. A “modulated data signal” may be a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and
20 not limitation, communication media may include wired media such as a wired network or direct-wired connection, and wireless media such as acoustic, radio frequency (RF), microwave, infrared (IR) and other wireless media. The term computer readable media as used herein may include both storage media and communication media.

[0041] Computing device 300 may be implemented as a portion of a small-form factor
25 portable (or mobile) electronic device such as a cell phone, a personal data assistant (PDA), a personal media player device, a wireless web-watch device, a personal headset device, an application specific device, or a hybrid device that include any of the above functions. Computing device 300 may also be implemented as a personal computer including both laptop computer and non-laptop computer configurations.

30 [0042] There is little distinction left between hardware and software implementations of aspects of systems; the use of hardware or software is generally (but not always, in that in certain contexts the choice between hardware and software can become significant) a design choice representing cost versus efficiency tradeoffs. There are various vehicles by

which processes and/or systems and/or other technologies described herein can be effected (e.g., hardware, software, and/or firmware), and that the preferred vehicle will vary with the context in which the processes and/or systems and/or other technologies are deployed. For example, if an implementer determines that speed and accuracy are paramount, the
5 implementer may opt for a mainly hardware and/or firmware vehicle; if flexibility is paramount, the implementer may opt for a mainly software implementation; or, yet again alternatively, the implementer may opt for some combination of hardware, software, and/or firmware.

[0043] The foregoing detailed description has set forth various embodiments of the devices
10 and/or processes via the use of block diagrams, flowcharts, and/or examples. Insofar as such block diagrams, flowcharts, and/or examples contain one or more functions and/or operations, it will be understood by those within the art that each function and/or operation within such block diagrams, flowcharts, or examples can be implemented, individually and/or collectively, by a wide range of hardware, software, firmware, or virtually any
15 combination thereof. In one embodiment, several portions of the subject matter described herein may be implemented via Application Specific Integrated Circuits (ASICs), Field Programmable Gate Arrays (FPGAs), digital signal processors (DSPs), or other integrated formats. However, those skilled in the art will recognize that some aspects of the embodiments disclosed herein, in whole or in part, can be equivalently implemented
20 in integrated circuits, as one or more computer programs running on one or more computers (e.g., as one or more programs running on one or more computer systems), as one or more programs running on one or more processors (e.g., as one or more programs running on one or more microprocessors), as firmware, or as virtually any combination thereof, and that designing the circuitry and/or writing the code for the software and or
25 firmware would be well within the skill of one of skill in the art in light of this disclosure. In addition, those skilled in the art will appreciate that the mechanisms of the subject matter described herein are capable of being distributed as a program product in a variety of forms, and that an illustrative embodiment of the subject matter described herein applies regardless of the particular type of signal bearing medium used to actually carry out the
30 distribution. Examples of a signal bearing medium include, but are not limited to, the following: a recordable type medium such as a floppy disk, a hard disk drive, a Compact Disc (CD), a Digital Versatile Disk (DVD), a digital tape, a computer memory, etc.; and a transmission type medium such as a digital and/or an analog communication medium (e.g.,

a fiber optic cable, a waveguide, a wired communications link, a wireless communication link, etc.).

[0044] Those skilled in the art will recognize that it is common within the art to describe devices and/or processes in the fashion set forth herein, and thereafter use engineering practices to integrate such described devices and/or processes into data processing systems. That is, at least a portion of the devices and/or processes described herein can be integrated into a data processing system via a reasonable amount of experimentation. Those having skill in the art will recognize that a typical data processing system generally includes one or more of a system unit housing, a video display device, a memory such as volatile and non-volatile memory, processors such as microprocessors and digital signal processors, computational entities such as operating systems, drivers, graphical user interfaces, and applications programs, one or more interaction devices, such as a touch pad or screen, and/or control systems including feedback loops and control motors (e.g., feedback for sensing position and/or velocity; control motors for moving and/or adjusting components and/or quantities). A typical data processing system may be implemented utilizing any suitable commercially available components, such as those typically found in data computing/communication and/or network computing/communication systems.

[0045] The herein described subject matter sometimes illustrates different components contained within, or connected with, different other components. It is to be understood that such depicted architectures are merely exemplary, and that in fact many other architectures can be implemented which achieve the same functionality. In a conceptual sense, any arrangement of components to achieve the same functionality is effectively "associated" such that the desired functionality is achieved. Hence, any two components herein combined to achieve a particular functionality can be seen as "associated with" each other such that the desired functionality is achieved, irrespective of architectures or intermedial components. Likewise, any two components so associated can also be viewed as being "operably connected", or "operably coupled", to each other to achieve the desired functionality, and any two components capable of being so associated can also be viewed as being "operably couplable", to each other to achieve the desired functionality. Specific examples of operably couplable include but are not limited to physically mateable and/or physically interacting components and/or wirelessly interactable and/or wirelessly interacting components and/or logically interacting and/or logically interactable components.

[0046] With respect to the use of substantially any plural and/or singular terms herein, those having skill in the art can translate from the plural to the singular and/or from the singular to the plural as is appropriate to the context and/or application. The various singular/plural permutations may be expressly set forth herein for sake of clarity.

5 [0047] It will be understood by those within the art that, in general, terms used herein, and especially in the appended claims (e.g., bodies of the appended claims) are generally intended as "open" terms (e.g., the term "including" should be interpreted as "including but not limited to," the term "having" should be interpreted as "having at least," the term "includes" should be interpreted as "includes but is not limited to," etc.). It will be further
10 understood by those within the art that if a specific number of an introduced claim recitation is intended, such an intent will be explicitly recited in the claim, and in the absence of such recitation no such intent is present. For example, as an aid to understanding, the following appended claims may contain usage of the introductory phrases "at least one" and "one or more" to introduce claim recitations. However, the use of such phrases should not be
15 construed to imply that the introduction of a claim recitation by the indefinite articles "a" or "an" limits any particular claim containing such introduced claim recitation to disclosures containing only one such recitation, even when the same claim includes the introductory phrases "one or more" or "at least one" and indefinite articles such as "a" or "an" (e.g., "a" and/or "an" should typically be interpreted to mean "at least one" or "one or more"); the
20 same holds true for the use of definite articles used to introduce claim recitations. In addition, even if a specific number of an introduced claim recitation *is* explicitly recited, those skilled in the art will recognize that such recitation should typically be interpreted to mean *at least* the recited number (e.g., the bare recitation of "two recitations," without other modifiers, typically means *at least* two recitations, or *two or more* recitations). Furthermore,
25 in those instances where a convention analogous to "at least one of A, B, and C, etc." is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (e.g., "a system having at least one of A, B, and C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). In those instances
30 where a convention analogous to "at least one of A, B, or C, etc." is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (e.g., "a system having at least one of A, B, or C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, etc.). It will be further understood by those

- 14 -

within the art that virtually any disjunctive word and/or phrase presenting two or more alternative terms, whether in the description, claims, or drawings, should be understood to contemplate the possibilities of including one of the terms, either of the terms, or both terms. For example, the phrase "A or B" will be understood to include the possibilities of "A" or "B" or "A and B."

[0048] While various aspects and embodiments have been disclosed herein, other aspects and embodiments will be apparent to those skilled in the art. The various aspects and embodiments disclosed herein are for purposes of illustration and are not intended to be limiting, with the true scope and spirit being indicated by the following claims.

We Claim:

1. A method for retrieving information from a semantic database having a plurality of semantic data, comprising:
selecting, by one or more processors, a hub that is resolved with a number of
5 semantic data based on a resolution rule, when the number of the semantic data is greater than a threshold in an offline environment;
calculating, by one or more processors, a semantic data set that is associated with the hub based on a searching approach and the resolution rule in the offline environment;
10 indexing, by one or more processors, the semantic data set in the offline environment;
modifying, by one or more processors, the semantic database to include the semantic data set in the offline environment; and
retrieving, by one or more processors, information from the semantic data set in an
15 online environment.
2. The method of claim 1, wherein the selecting, the calculating, the indexing, and the modifying are repeated in sequence.
- 20 3. The method of claim 1, wherein the calculating further comprises selecting a first semantic data into the semantic data set based on the resolution rule.
4. The method of claim 3, wherein the calculating further comprises calculating an implicit semantic data based on the resolution rule, the semantic data set and logical
25 reasoning.
5. The method of claim 1, further comprising compressing the semantic data set.
6. The method of claim 5, further comprising mapping a second semantic data and a
30 third semantic data to a fourth semantic data, wherein the second semantic data and the third semantic data are semantically equivalent.
7. The method of claim 1, wherein the retrieving further comprises locating the semantic data set by comparing a query to the hub.

8. A computer-readable storage medium encoded with computer-executable instructions for retrieving information from a semantic database having a plurality of semantic data, which when executed by a computing device, causes the computing device to:

- 5 select a hub that is resolved with a number of semantic data based on a resolution rule, when the number of the semantic data is greater than a threshold in an offline environment;
- calculate a semantic data set that is associated with the hub based on a searching approach and the resolution rule in the offline environment;
- 10 index the semantic data set in the offline environment;
- modify the semantic database to include the semantic data set in the offline environment; and
- retrieve information from the semantic data set in an online environment.
- 15 9. The computer-readable storage medium of claim 8, further containing additional instructions, which when executed by the computing device, causes the computing device to repeatedly select the hub, calculate the semantic data set, index the semantic data set, and modify the semantic database in sequence.
- 20 10. The computer-readable storage medium of claim 8, further containing additional instructions, which when executed by the computing device, causes the computing device to select a first semantic data into the semantic data set based on the resolution rule.
11. The computer-readable storage medium of claim 10, further containing additional
- 25 instructions, which when executed by the computing device, causes the computing device to calculate an implicit semantic data based on the resolution rule, the semantic data set and logical reasoning..
12. The computer-readable storage medium of claim 8, further containing additional
- 30 instructions, which when executed by the computing device, causes the computing device to compress the semantic data set.
13. The computer-readable storage medium of claim 12, further containing additional instructions, which when executed by the computing device, causes the computing device to

map a second semantic data and a third semantic data to a fourth semantic data, wherein the second semantic data and the third semantic data are semantically equivalent.

14. The computer-readable storage medium of claim 8, further containing additional
5 instructions, which when executed by the computing device, causes the computing device to locate the semantic data set by comparing a query to the hub.

15. A computing device configured to retrieve information from a semantic database having a plurality of semantic data, comprising:

10 a processing unit, wherein the processing unit is configured to select a hub that is resolved with a number of semantic data based on a resolution rule, when the number of the semantic data is greater than a threshold in an offline environment;
calculate a semantic data set that is associated with the hub based on a searching
15 approach and the resolution rule in the offline environment;
index the semantic data set in the offline environment;
modify the semantic database to include the semantic data set in the offline environment; and
retrieve information from the semantic data set in an online environment.

20

16. The computing device of claim 15, wherein the processing unit is further configured to repeatedly select the hub, calculate the semantic data set, index the semantic data set, and modify the semantic database in sequence.

25 17. The computing device of claim 15, wherein the processing unit is further configured to select a first semantic data into the semantic data set based on the resolution rule.

18. The computing device of claim 15, wherein the processing unit is further configured to calculate an implicit semantic data based on the resolution rule, the semantic data set
30 and logical reasoning.

19. The computing device of claim 15, wherein the processing unit is further configured to compress the semantic data set.

20. The computing device of claim 15, wherein the processing unit is further configured to locate the semantic data set by comparing a query to the hub.

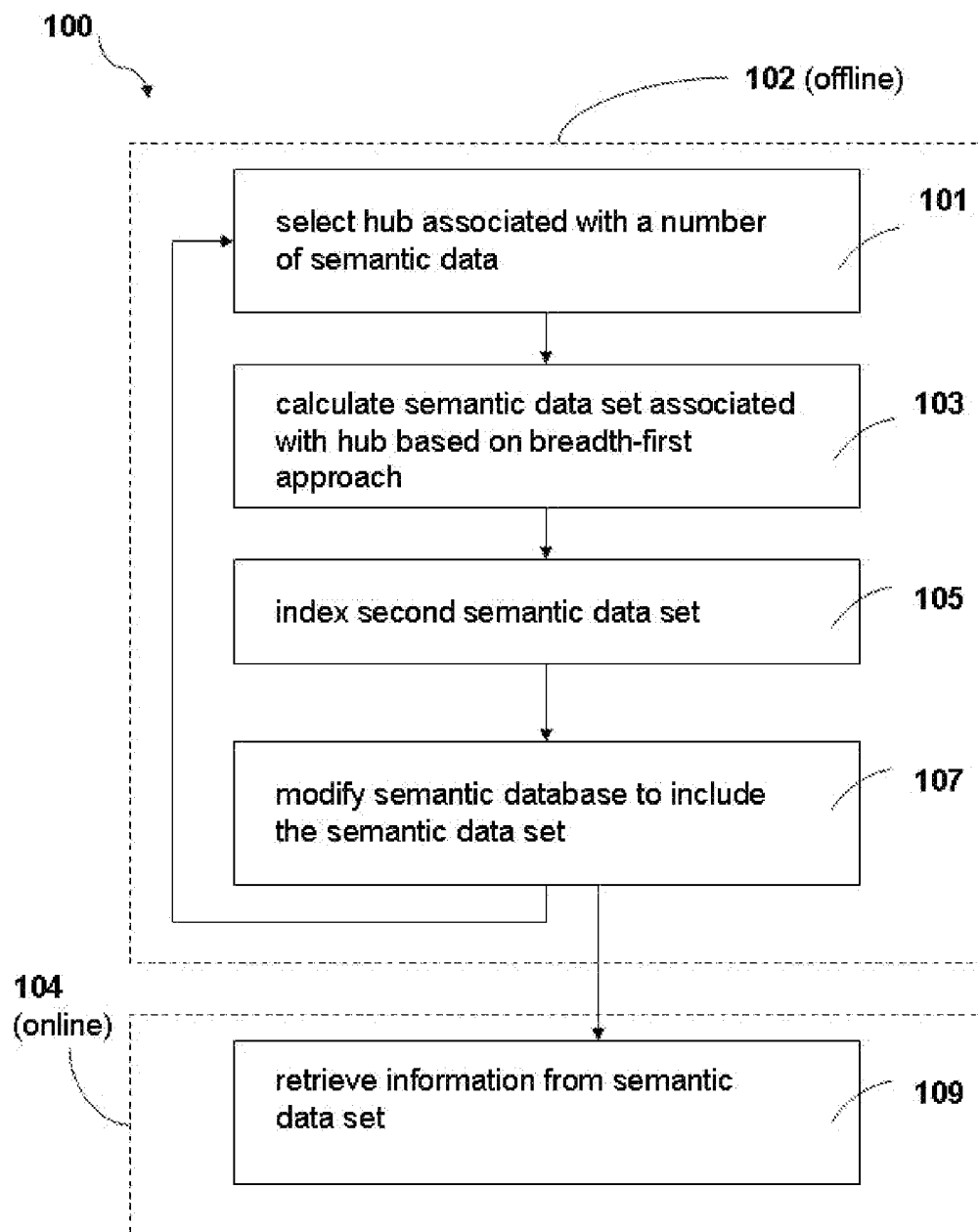


FIG. 1

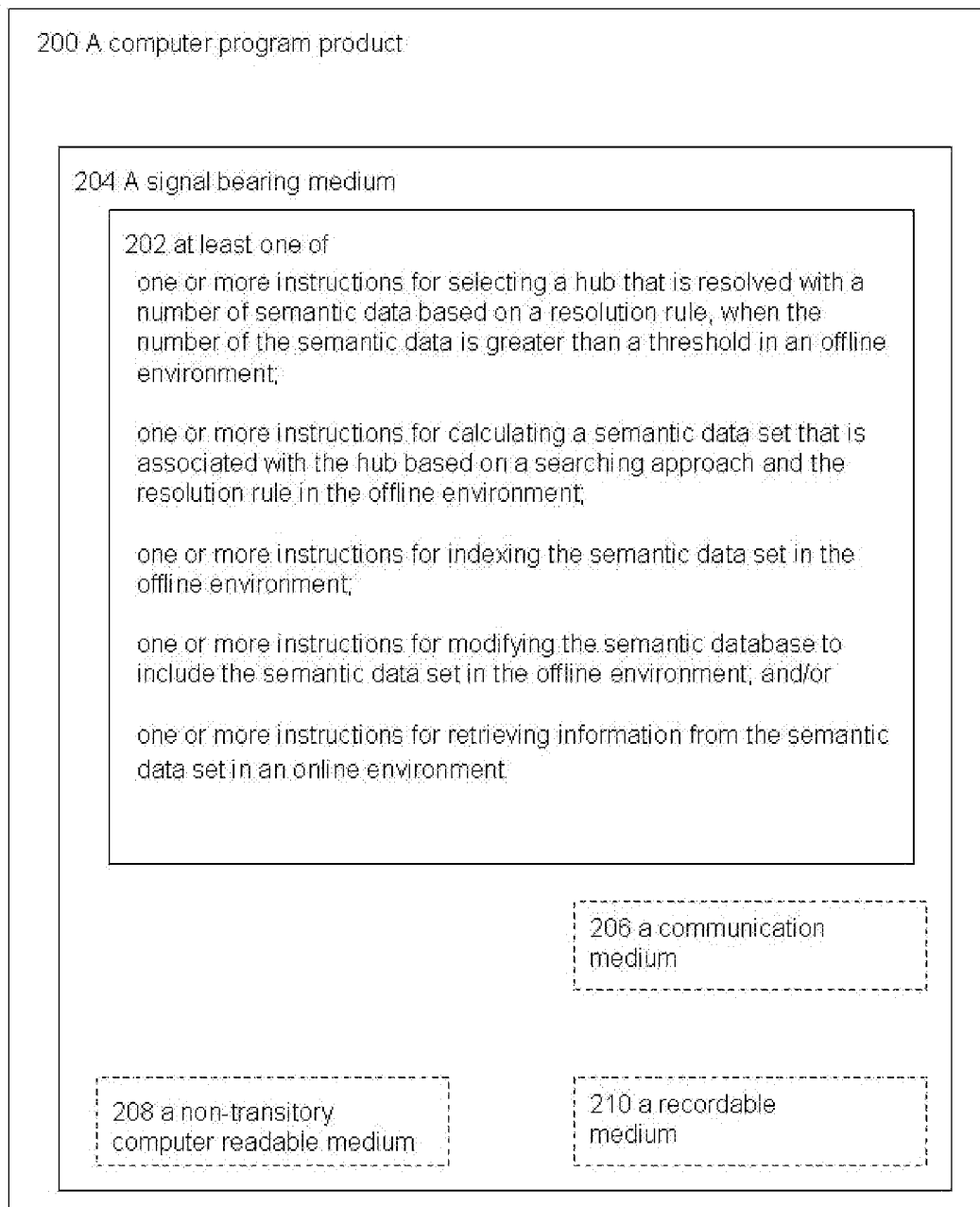


FIG. 2

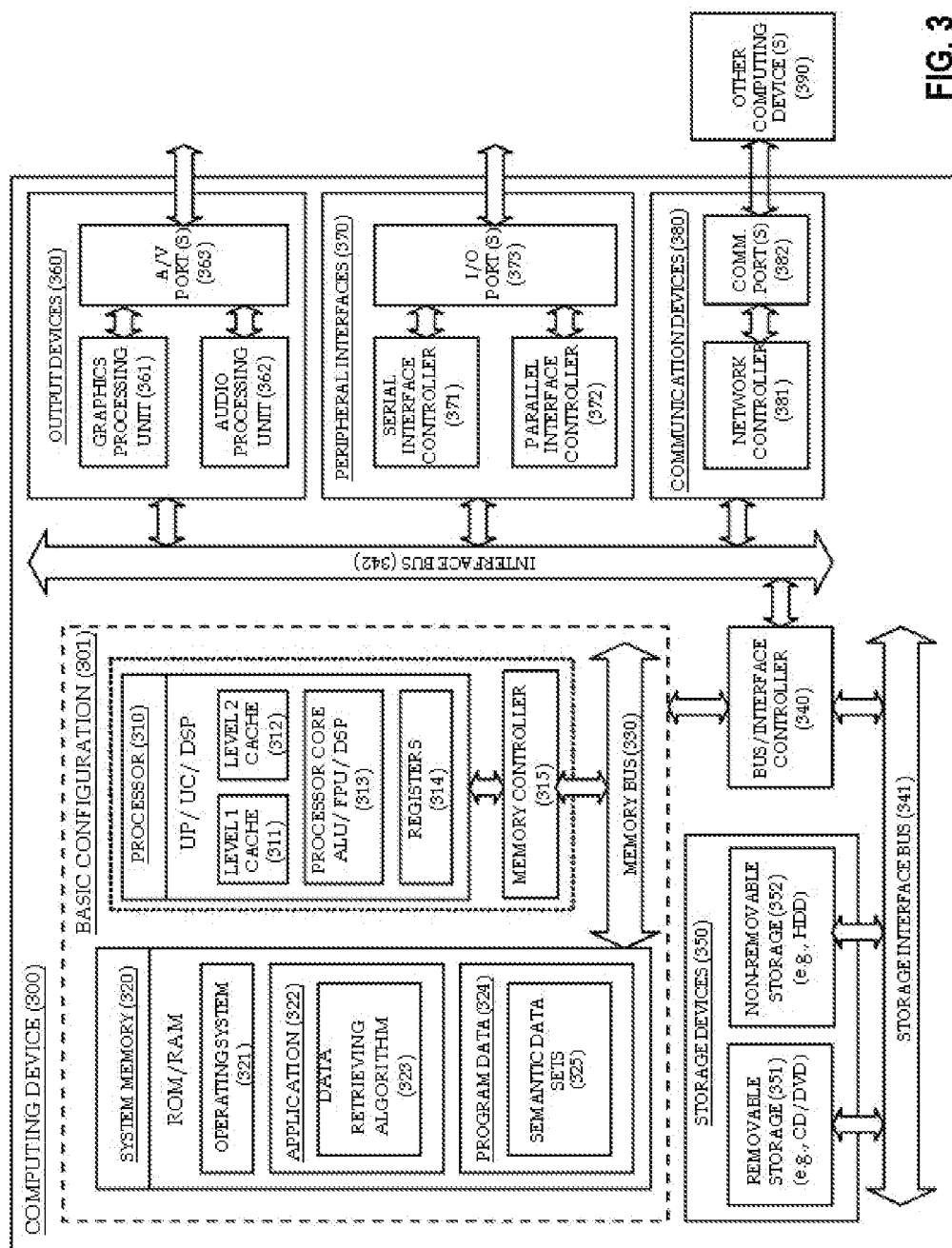


FIG. 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2011/077615

A. CLASSIFICATION OF SUBJECT MATTER

G06F17/30 (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: G06F17/-; G06N

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)

CNTXT,CPRSABS,CNABS,CNMED, DWPL,SIPOABS,CPEA, USTXT,WOTXT,EPTXT,CATXT, GOOGLE: semantic w database, semantic w net, network, relati+, resolv+, label+, index, tag, hub, hot, popular, focus, engine, set, sub w set, sub?set, threshold, compress+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US2003/0177000A1 (VERITY, INC.) 18 Sep.2003 (18.09.2003) description, paragraphs 0047-0065, figures 3-4	1-20
A	US6778970B2 (AU, Lawrence) 17 Aug. 2004 (17.08.2004) the whole document	1-20
A	US2008/0275694A1 (EXPERT SYSTEM SPA) 06 Nov. 2008 (06.11.2008) the whole document	1-20
A	US2007/0112803A1 (PETTOVELLO, Primo M.) 17 May 2007 (17.05.2007) the whole document	1-20
A	US2010/0036789A1 (COGNIVA INFORMATION SOLUTIONS INC.) 11 Feb. 2010 (11.02.2010) 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
16 Apr. 2012 (16.04.2012)Date of mailing of the international search report
03 May 2012 (03.05.2012)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
YAO, Yaqian
Telephone No. (86-10)62413461

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2011/077615

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US2003177000A1	18.09.2003	WO03079221A1	25.09.2003
		AU2003225742A1	29.09.2003
		EP1520236A1	06.04.2005
		CN1774713A	17.05.2006
US6778970B2	17.08.2004	NONE	
US2008275694	06.11.2008	NONE	
US2007112803A1	17.05.2007	NONE	
US2010036789A1	11.02.2010	WO2008074160A1	26.06.2008
		CA2673422A1	26.06.2008