



(51) International Patent Classification:
G06N 5/00 (2006.01) *G06F 17/30* (2006.01)

(21) International Application Number:
PCT/MY2010/000216

(22) International Filing Date:
20 October 2010 (20.10.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PI20094418 21 October 2009 (21.10.2009) MY

(71) Applicant (for all designated States except US): **MIMOS BERHAD** [MY/MY]; Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **KOW, Wen Onn** [MY/MY]; Mimos Berhad, Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY). **BEIKZADEH, Modammad Reza** [IR/MY]; Mimos Berhad, Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY). **LUKOSE, Dickson** [AU/MY]; Mimos Berhad, Technology Park of Malaysia, Bukit Jalil, 57000 Kuala Lumpur (MY).

(74) Agent: **MOHAN, K.**; Adastra Intellectual Property Sdn Bhd, A-28-10, Menara UOA Bangsar Utama 1, No. 5, Jalan Bangsar Utama 1, 59000 Kuala Lumpur (MY).

(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, 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, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: METHOD AND SYSTEM FOR USE IN RETRIEVAL OF KNOWLEDGE OR INFORMATION USING SEMANTIC LINKS

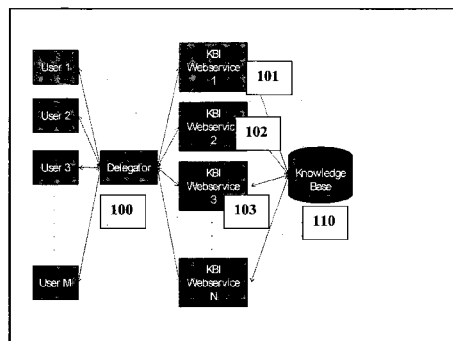


FIG 1

(57) Abstract: The present invention discloses a method and system for use in retrieval of knowledge or information, whereby the essential steps include the user creating and selecting a focus concept based on the desired information to be retrieved; delegating and sending a request for fetching of data based on said focus concept; fetching a first level of data in relation to said focus concept; providing subsequent concepts related to the focus concept in a queue for fetching of data; fetching data in relation to the subsequent concepts and the focus concept concurrently; storing all fetched data; replacing previous fetched data with a recent fetched data; displaying all fetched data; updating stored data. It is further noted that data for all concepts is fetched based on the semantic links of the respective concepts. The system allows connection by multiple users due to its ability to send multiple requests for data fetching concurrently.

**METHOD AND SYSTEM FOR USE IN RETRIEVAL OF KNOWLEDGE OR
INFORMATION USING SEMANTIC LINKS**

FIELD OF INVENTION

5

The preferred embodiments of the present invention direct to a method and system for use in obtaining or retrieval of knowledge or information, and more particularly to retrieving knowledge or information concurrently while allowing multiple users to maintain interaction within the system based on the acquired knowledge provided therein, using semantic links.

10

BACKGROUND OF INVENTION

15

20

25

The use of semantic knowledge in retrieval of information from computer-based data content is becoming rather prevalent in recent years. Coupled with the gradual and in -depth integration into the technology, newly revolutionized classes of technological capabilities in regards to the use of semantic knowledge and tasks related thereto have been fruitfully created. Generally, employment of semantics ranges from as representation for language or knowledge free text based on accessible knowledge base, transaction processes within a distributed computer system, to assessments of brain related impairments of a patient for the medical field.

Another prominent deployment in relation to semantics and information retrieval is the development of semantic web, whereby semantics play a major role in ensuring efficient operation when users explore the services contained by the web. This is achieved by way of providing a language or text understandable by the computers or machines so as to meet the requests by users while exploring the web. The formal tools and data interchange formats which may be involved with respect to this purpose may include XML, Resource Description Framework (RDF), N-Triples, Web Ontology Language (OWL) and RDF Schema.

At present, while the use of semantic knowledge in various applications particularly for information retrieval is highly recognized, a simplistic and speedy method or approach for such purpose has yet to be predominantly developed, thus making information retrieval a daunting task to users in certain cases.

The above circumstance may become overwhelming when dealing with gigantic knowledge bases especially with the introduction of Web 3.0. As it is rather apparent that loading all data which is in the form of triples from the vast knowledge base into a memory device would be impractical, a method for fetching knowledge or information therefore a more preferred option and more effective approach

as compared to downloading them. Nevertheless, fetching data becomes more complex with the presence of multiple users accessing the same knowledge. Presuming that one concept may entail hundreds or thousands of triples, at least one or if
5 not all users may end up waiting for a substantial period of time before he can view/manipulate the retrieved data.

The present invention therefore is focused on creating and thus providing an effective method and system which can be
10 used to resolve the above-discussed predicaments.

The primary object of the present invention is to provide a method and system for use in retrieval of information by way of semantic links in fetching data related to a concept
15 created by users (data fetched is in the form of triples in this context), from a form of knowledge storage base concurrently while allowing the user to still interact within the system based on the existing knowledge in memory.

20 It is a further object of the present invention to provide a method and system for use in retrieving information or knowledge based on semantic links, whereby the new knowledge or information being fetched is subsequently updated on display screen of the machine upon received from the server.

It is yet another object of the present invention to provide a method and system for use in retrieving information based on semantic links, whereby said fetching can be performed concurrently by a plurality of users.

5

It is another object of the present invention to provide a system and method for use in retrieving information based on semantic links whereby there are a plurality of servers are required in order to support concurrent fetching of
10 information and for ensuring speedy process.

Further objects and advantages of the present invention may become apparent upon referring to the preferred embodiments of the present invention as shown in the accompanying drawings
15 and as described in the following description.

SUMMARY OF INVENTION

In one aspect, the present invention relates to a method for
20 use in retrieving information from a knowledge base comprising the steps of: creating and selecting a focus concept based on the desired information to be retrieved; delegating and sending a request for fetching of data based on said focus concept; fetching a first level of data in relation to said
25 focus concept;

providing subsequent concepts related to the focus concept in a queue for fetching of data; fetching data in relation to the subsequent concepts and the focus concept concurrently; storing all fetched data; replacing previous fetched data with
5 a recent fetched data; displaying all fetched data; updating stored data; wherein all data is fetched based on the semantic links of the created focus concept.

In another aspect, the present invention provides a system for
10 use in retrieval of data from a knowledge base comprising: a plurality of data fetchers (101, 102, 103) configured to fetch data based on semantic links of a concept from the knowledge base (110); said data fetchers (101, 102, 103) are further configured to fetch data concurrently from the knowledge base
15 (110); a delegator (100) configured for delegating and thus sending data fetching requests to the plurality of data fetchers (101, 102, 103); said delegator (100) is further configured to send fetching requests to the plurality of data fetchers (101, 102, 103) concurrently.

20

BRIEF DESCRIPTION OF THE DRAWINGS

In order to provide a comprehensive understanding of the nature of the invention, reference should be made to the
25 accompanying drawings and the subsequent description. Briefly, the drawings are:

FIG 1 shows the overall structure of the system in accordance with a preferred embodiment of the present invention;

FIG 2 shows the steps involved with respect to a server when the user created a concept for the method in accordance with a preferred embodiment of the present invention;

FIG 3 shows the components involved in one preferred embodiment of the present invention;

FIG 4 shows a flowchart containing the essential steps involved when a user selects a concept as a focus concept in accordance with a preferred embodiment of the present invention;

FIG 5 shows the algorithm involved when the user selects a concept as a focus concept in accordance with another preferred embodiment of the present invention; and

FIG 6 shows the fetching of data process for the second level with respect to a plurality of linked concepts.

DETAILED DESCRIPTION OF THE PRESENT INVENTION

In line with the above summary, the following description of a number of specific and alternative embodiments is provided to

understand the inventive features of the present invention. It shall be apparent to one skilled in the art, however that this invention may be practiced without such specific details. Some of the details may not be described at length so as not to obscure the invention. For ease of reference, common reference numerals will be used throughout the figures when referring to the same or similar features common to the figures.

10 **FIG 1** provides the overall structure of the system in accordance with the preferred embodiments of the present invention. According to the preferred structure as depicted in the drawing, there may be a plurality of users connected to the system. The system comprises of at least one delegator
15 (100), a plurality of knowledge base web-services (KBI) (101, 102, and 103) as the information fetchers, and a knowledge base (110) being the primary knowledge storage where information is stored and acquired. As can be seen in FIG 1, the plurality of users (20, 40, 60, U_M) is connected to the
20 delegator (100) and thus indirectly connected to the plurality of knowledge web-services (101, 102, 103) and knowledge base (110) through said delegator (100).

The main purpose of a delegator (100) in this system is for
25 delegating fetch requests generated based on each user's demand with to the concept created. To serve this purpose, the

delegator (100) is configured to be able to store a list of all the knowledge base web-services provided within the system so as to ensure smooth delegation of fetch requests. The delegator (100) is further configured to delegate fetch
5 requests to the suitable and available knowledge web-services (101, 102, and 103).

It is preferred that each of the knowledge base web-services (101, 102, and 103) is configured to have independent or local
10 cache for data or triples which are fetched from the knowledge base (110). Additionally, it is understood that the KBIs may be replaced with other suitable fetchers of similar capabilities. With such feature, repeated fetch requests or any requests corresponding to information presently stored
15 within the cache, said information can be sent directly and promptly back to the user devoid of accessing the primary knowledge base (110). This therefore results to fast or accelerated response time to the user. Further, a cache management system will ensure that only a predetermined number
20 of triples are stored in the memory at a predetermined amount of time. In the event that the limit of the cache is reached, the system will automatically analyze and consequently remove triples that have been in the cache the longest and yet has not be accessed by any means.

FIG 2 shows the steps involved with respect to a server when the user created a concept for the method in accordance with a preferred embodiment of the present invention. It is understood that a concept in this context may include objects
5 or subjects related to the intended knowledge retrieval.

The first step is creation of concept list step (S200). Each of the lists thereby contains the requested concepts which are to be fetched from the primary knowledge base (110). It should
10 be noted that the information is acquired with the aid of the semantic links for the concepts requested. The list is then forwarded to the delegator (100) for delegation (S300) to the available knowledge base web-services (101, 102, and 103). In the next step (S400), each knowledge web-services receives the
15 assigned task thereby proceeds to fetch triples inclusive of subjects or objects related to said concept in the subsequent step (S500) from the primary knowledge base (110) which is accordingly displayed (S600) on the screen for each user.

20 In the event that there are multiple users connected within the system, the fetch list from each user can be sent concurrently to the delegator (100) to fetch the relevant triples. It is preferred that the system is configured in a manner such that the knowledge base web-services is able to
25 traverse within links so as to acquire or fetch more information. Upon fetched the requested triples, the users

screens are updated accordingly, whereby new subjects or objects are displayed over the former objects or subjects or on top of the list containing the former objects or subjects.

5 It is further noted that the user is allowed to proceed with requesting for new concepts or information even when the delegator (100) is in the midst of delegating requests to be fetched from knowledge base (110). The system therefore is configured to acquire further knowledge in the background even
10 when the users are in the midst of viewing the current information presented to them. Comprehendingly, the steps involved from creation of list to fetching from knowledge base (110) are repeated. In the event that a small amount of triples is required to fulfill a request, the turnaround time
15 or fetching time is shortened, thereby reducing the chance of the display screen to not be refreshed while waiting for bigger amount of triples to arrive and thus be displayed on screen. All information fetched based on users request is suitably stored in a memory device provided within the system.
20 Preferably, the number of fetching processes may not be specified for a period of time thereby allowing continuous fetching of objects or subjects. In addition to this, as the processes are allowed to run concurrently, the system does not control the sequence of processes to be terminated. It is
25 preferred that the fetching process is completed only if the user ceases sending further requests.

In another preferred embodiment, the system maintains the collected or fetched information in cache form that can be cleared whenever required, whereby older triples that were not used or accessed over a period of time may be replaced by more recent information. As the information fetched are non exhaustive, whereby for this system the more in depth of a search is preferred, the more information can be cached and thereby the user can traverse the semantic web more rapidly.

10 In view of **FIG 3**, the user is aided with the semantic browser (300) to perform the search which is configured to store a list containing the concepts in memory which have been loaded in the primary knowledge base (110). It is preferred that as the user explores more of the knowledge base (110); more
15 concepts are fetched and loaded from the knowledge base (110). A fetch manager (400) may be provided within the system to manage the collected fetch lists. A predetermined service - oriented architecture (SOA) (500) which may comprise a plurality of knowledge web-services as described in preceding
20 paragraphs; acts as an intermediate for governing concepts or information transition between the knowledge base (100) and the fetch manager (400).

FIG 4 shows a flowchart containing the essential steps
25 involved when a user selects a concept as a focus concept. In the first step, (U200), corresponding to that of flowchart as

seen in FIG 2, the user selects a concept which is recognized as the focus concept by the system for fetching of triples task. Based on said focus concept requested by the user, data or triples in relation to said concept is obtained from the base (100) and suitably displayed in the next step (U300) for review of the user. At the same time, the system proceeds to place the concepts which are related or linked to the initial focus concept placed in the queue (U400) and thereby allowing fetching of further triples related to the concepts placed in queue in the next stage (U500). The fetched triples are placed in the cache (U600) and replacements of older or outdated triples may be performed during this stage. The stored triples may be shown or displayed to user whenever desired. In the subsequent step (U700) the concepts which are linked to the next level concept are added to the queue created previously. It is preferred that the steps of U600 to U800 may be repeated subject to the depth or level of searches requested by users.

FIG 5 relates to the algorithm involved when the user selects a concept as a focus concept as described in the preceding paragraph. Upon selected the focus concept, a node of interest is chosen. The node of interest is referred herein as NODE A. From here, the system proceeds to fetch all links relevant to that of NODE A. These relevant links are referred herein as NODE B, C, D and E. Accordingly; the relevant links are stored at the other end of NODE A into the cache as shown in FIG 5.

NODE A therefore is fully loaded which is in contrast to the status of the remaining nodes B, C, D and E. These remaining nodes or concepts are pre-loaded in a manner such that the other nodes that these remaining nodes they may be connected or attached to. These nodes are then placed in the fetch queue to allow the next step to begin which will be described herein shortly.

Now referring to **FIG 6**, the fetching of triples process for nodes B, C, D and E are carried out. Upon initiated the first depth level of fetching triples based on the nodes, the nodes with respect to said first level are now fully loaded while the nodes for the second depth level which are partially loaded are suitably placed in cache.

The user may select or create a new concept apart from that provided on the screen. In the event that this occurs, the steps involved based creations of new focus concepts whether by a single user or multiple users are therefore repeated, whereby the nodes for the first level of depth of each concept is preferably placed at the top of the queue for fetching of triples.

With the method and system of the present invention as described earlier, fetching can be done in the background

thereby allowing the user to continue working while the loading is carried out.

It is apparent that the fetching is carried out in a parallel
5 manner, which means in the event that a focus concept requires a substantial amount of links, said concept will not block or halt the progress of other concepts with fewer amount of links from loading in an acceptable speed.

10 Apart from the above, the speed of loading knowledge or information can be improved significantly as the triples related to the knowledge are readily loaded into the memory allowing fast retrieval in the event that the subsequent new concept is similar or linked to the previous concepts.

15

The number of components within the system may be modified on the server side so as to improve performance as the number of concurrent users increase, however such modifications will not affect the efficiency of the system. Any number of
20 delegators, KBIs and knowledge bases can be configured to best suit the application and number of users.

From the foregoing, it would be appreciated that the present invention may be modified in light of the above teachings. It

is therefore understood that, within the scope of the appended claims, the invention may be practiced otherwise than as specifically described.

CLAIMS

1. A method for use in retrieving information from a knowledge base comprising the steps of:
- 5 creating and selecting a focus concept based on the desired information to be retrieved;
- delegating and sending a request for fetching of data based on said focus concept;
- fetching a first level of data in relation to said focus
- 10 concept;
- providing subsequent concepts related to the focus concept in a queue for fetching of data;
- fetching data in relation to the subsequent concepts and the focus concept concurrently;
- 15 storing all fetched data;
- replacing previous fetched data with a recent fetched data;
- displaying all fetched data;
- updating stored data;
- wherein all data is fetched based on the semantic links of the
- 20 created focus concept.
2. The method as claimed in Claim 1 wherein the method further comprising the step of continuously fetching data based on the first concept and the subsequent concepts.

3. The method as claimed in Claim 1 wherein the method further comprising the step of fetching the subsequent levels of data based on the first concept and subsequent concepts concurrently.

5

4. The method as claimed in Claim 1 wherein the method further comprising the step of sending requests for fetching of data in relation to the subsequent concepts concurrently.

10

5. The method as claimed in Claim 1 wherein the step of updating stored data comprising removing non-accessed fetched data.

15

6. The method as claimed in Claim 1 wherein the method further comprising the step of placing all fetched data in at least one local cache within the system to ensure fast retrieval.

20

7. The method as claimed in Claim 1 wherein the steps of fetching data based on concepts comprise traversing the links of the concept so as to obtain more in-depth data.

25

8. The method as claimed in Claim 1 wherein the step of providing subsequent concepts related to the focus concept comprises placing said subsequent concepts in queue for loading or fetching of data.

9. The method as claimed in Claim 1 wherein the step of fetching data in relation to the subsequent concepts and focus concept concurrently comprises the step of loading data based on a variety of fetch -depth levels.

5

10. A system for use in retrieval of data from a knowledge base comprising:

a plurality of data fetchers (101, 102, 103) configured to fetch data based on semantic links of a concept from the
10 knowledge base (110); said data fetchers (101, 102, 103) are further configured to fetch data concurrently from the knowledge base (110);

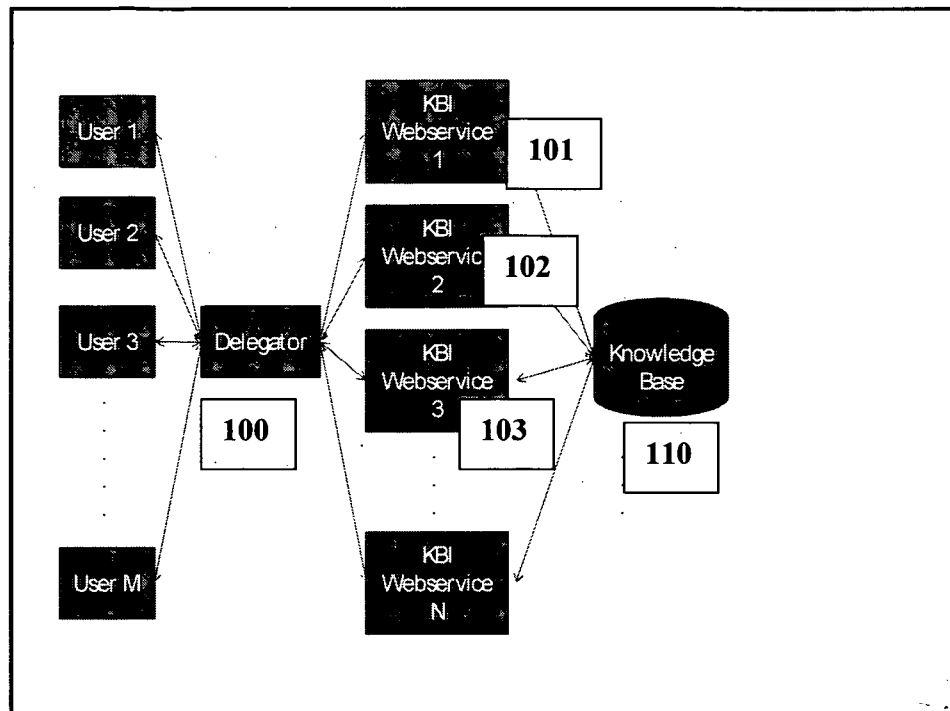
a delegator (100) configured for delegating and thus sending data fetching requests to the plurality of data fetchers (101,
15 102, 103); said delegator (100) is further configured to send fetching requests to the plurality of data fetchers (101, 102, 103) concurrently.

11. The system as claimed in Claim 10 wherein the system
20 further comprising a fetch request manager (400) configured for managing fetch requests prior to retrieval of information from the knowledge base (110).

12. The system as claimed in Claim 10 wherein the system
25 further comprising a semantic browser (300) configured for storing a list of concepts.

13. The system as claimed in Claim 10 wherein the system further comprising a service-oriented architecture (SOA) (500) configured to govern the concepts created within the system.

5 14. The system as claimed in Claim 10 wherein the system further comprising a plurality of user stations.

**FIG 1**

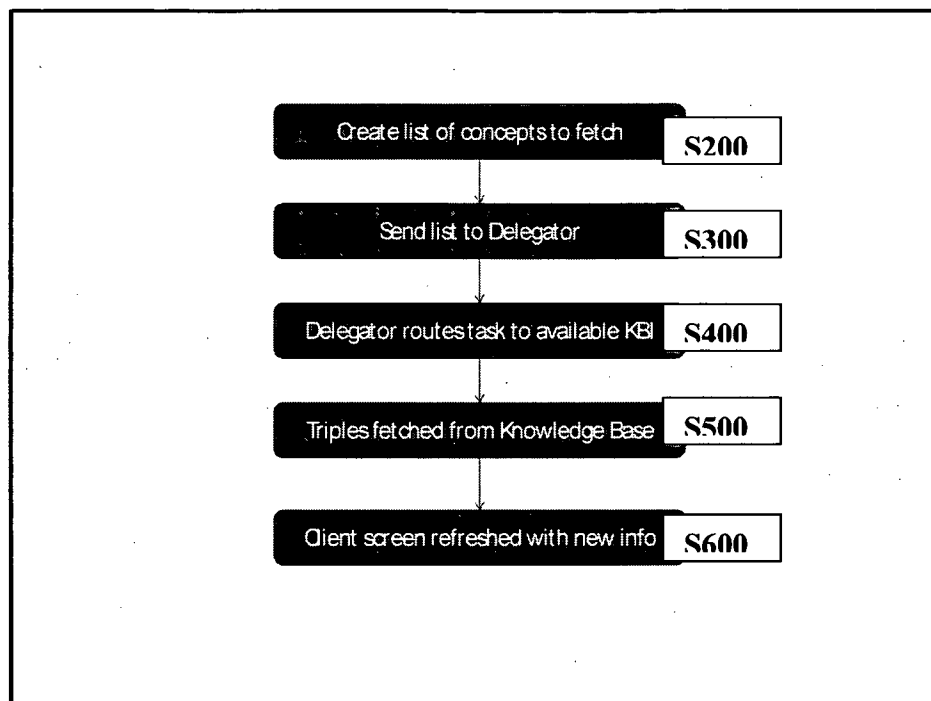
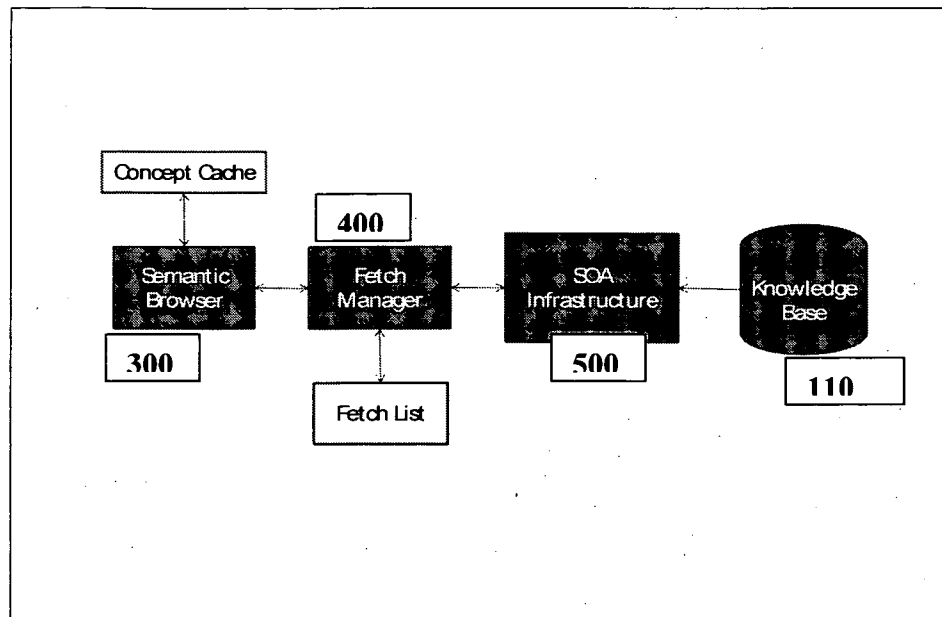
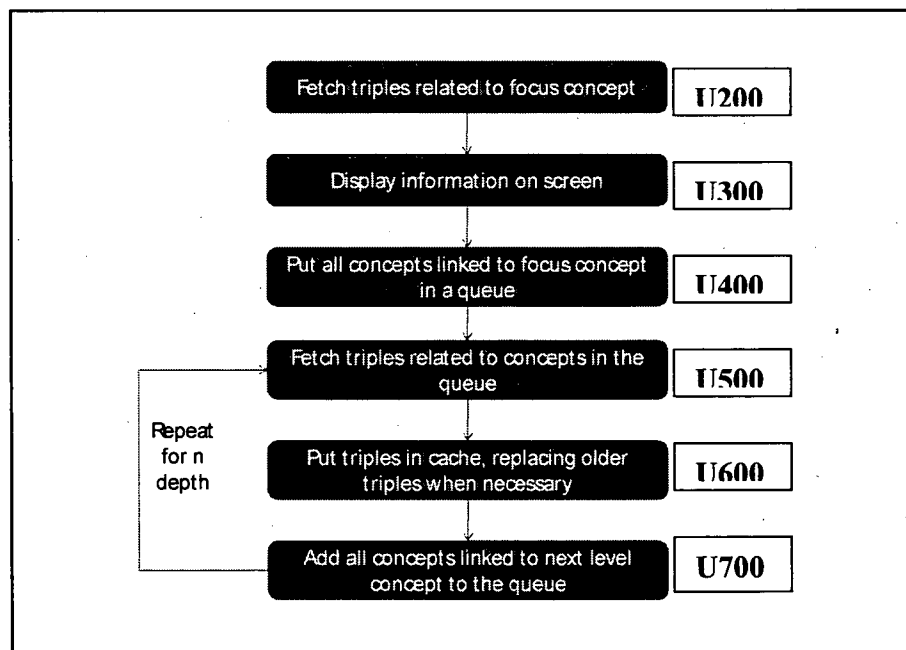


FIG 2

3/4

**FIG 3****FIG 4**

4/4

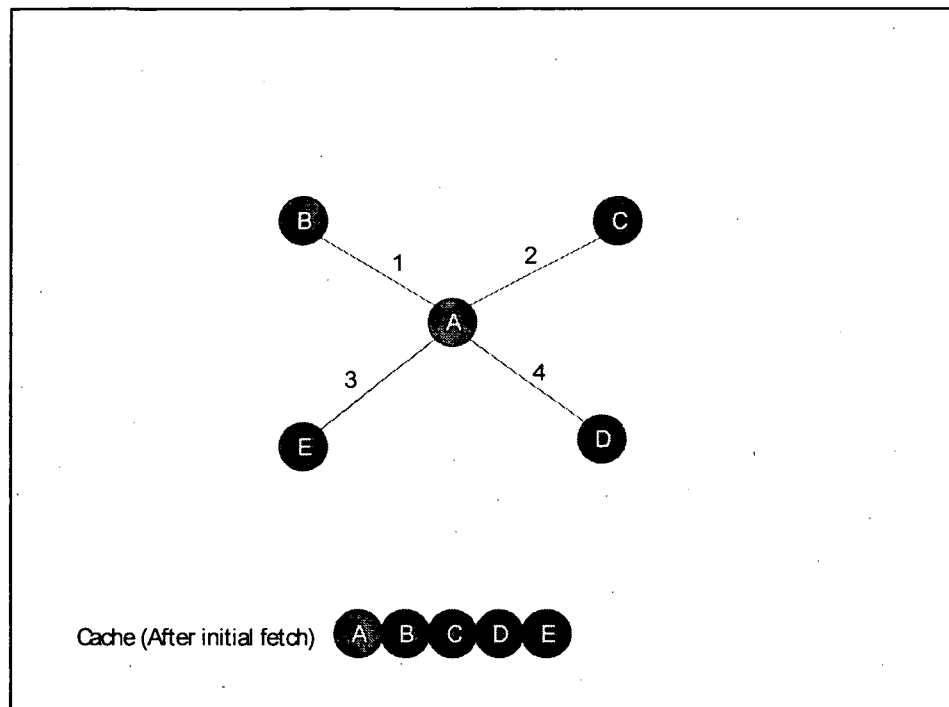


FIG 5

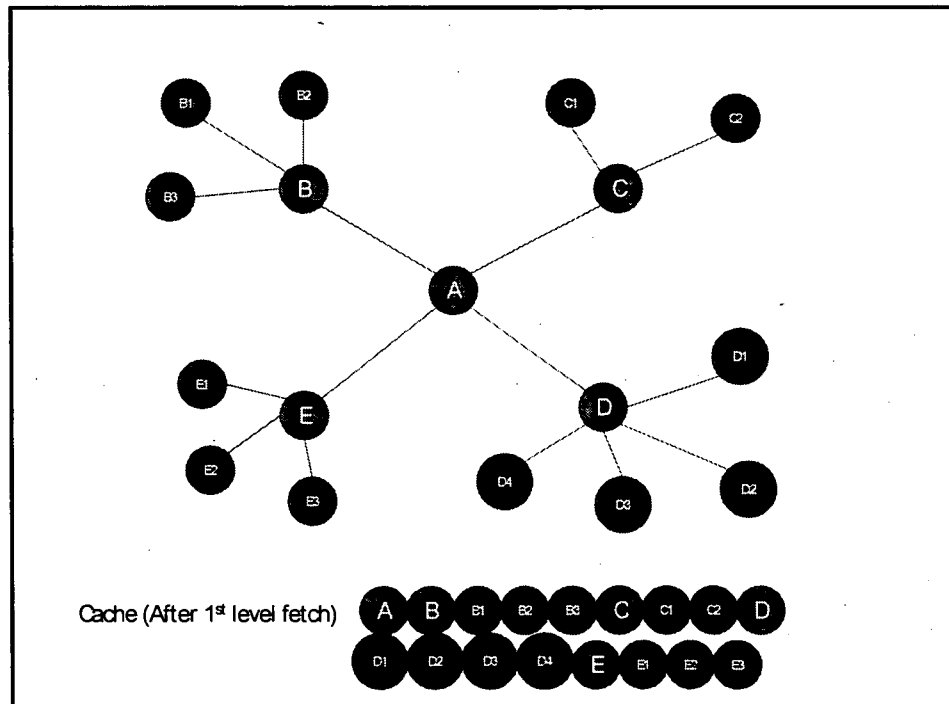


FIG 6

INTERNATIONAL SEARCH REPORT

International application No.

PCT/MY2010/000216

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

G06N 5/00 (2006.01)

G06F 17/30 (2006.01)

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)

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)

EPODOC, WPI, Google Patents:

Keywords: search engine/query, optimise, expand, concept, focus, semantic, lexicon, knowledge and like keywords

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2006/0106793 A1 (LIANG) 18 May 2006 Whole document, in particular: abstract, paragraphs [0012, 0048-0050, 0052, 0205-0213, 0314, 0319-0347, 0349-0350], figures [1, 2, 4, 14, 16]	1-14
A	WO 2009/117830 A1 (HOTGRINDS CANADA) 1 October 2009 Whole document	
A	US 2009/0112841 A1 (DEVARAKONDA <i>et al.</i>) 30 April 2009 Whole document	
A	US 2005/0154713 A1 (GLOVER <i>et al.</i>) 14 July 2005 Whole document	



Further documents are listed in the continuation of Box C



See patent family annex

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"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

25 January 2011

Date of mailing of the international search report

28 JAN 2011

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No. +61 2 6283 7999

Authorized officer

Lili Eng
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : +61 2 6283 3138

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/MY2010/000216

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member	
US	20060106793	CN 1716244	CN 1808430
		US 20050154723	US 20050160107
WO	2009117830	US 20100076984	US 2005144162
US	20090112841	NONE	US 2006047649
US	20050154713	NONE	
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			
END OF ANNEX			