



(51) **International Patent Classification:**  
**G06F 17/27** (2006.01) **G06F 17/20** (2006.01)  
**G06F 17/00** (2006.01)

(74) **Agent:** **MIRANDAH, Patrick**; Patrick Mirandah Co. (Malaysia) SDN. BHD., Suite 3B-19-3, Plaza Sentral, Jalan Stesen Sentral 5, 50470 Kuala Lumpur (MY).

(21) **International Application Number:**  
PCT/MY2010/000084

(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.

(22) **International Filing Date:**  
25 May 2010 (25.05.2010)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**  
PI 20092121 25 May 2009 (25.05.2009) MY

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

(72) **Inventors; and**

(75) **Inventors/Applicants** (*for US only*): **SADANANDAN, Arun, Anand** [MY/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **ONN, Kow Weng** [MY/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **ZADEH, Mohammad, Reza, Beik** [IR/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **LUKOSE, Dickson** [MY/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY).

(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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

#### Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

[Continued on next page]

(54) **Title:** A METHOD AND SYSTEM FOR EXTENDABLE SEMANTIC QUERY INTERPRETATION

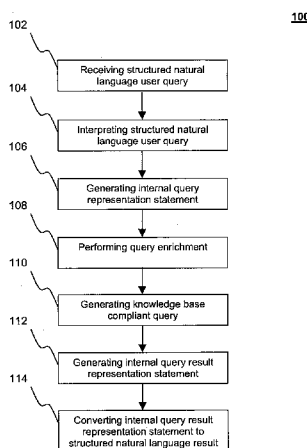


FIG. 1

(57) **Abstract:** A method (100) and a system (200) for an extendable semantic query interpretation, the system (200) comprises an intelligent word sense (202), a semantic query interpreter (206), a query transformer (208), a query enricher (210) and a natural language generator (218). The intelligent word sense (202) comprises means for receiving a structured natural language user query (102). The semantic query interpreter (206) comprises means for interpreting the structured natural language user query (104). The query transformer (208) comprises means for generating from the structured natural language user query, an internal query representation statement (106). The query enricher (210) comprises means for performing query enrichment (108) to generate at least one enriched internal query representation statement, generating from the at least one enriched internal query representation statement, a knowledge base compliant query (110) to provide for searching at least one query result from an ontology knowledge base (220) and generating at least one internal query result representation statement (112). The natural language generator (218) comprises means for converting the at least one internal query result representation statement to a structured natural language result (114).



---

**Published:**

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

## **A METHOD AND SYSTEM FOR EXTENDABLE SEMANTIC QUERY INTERPRETATION**

### **FIELD OF INVENTION**

5

The present invention relates to a method and system for extendable semantic query interpretation.

### **BACKGROUND ART**

10

The use of semantic query engines is becoming increasingly accepted as the preferred query processing means to provide meaning-based queries and retrieval of documents from a semantic based Web. Resources in the semantic based Web are semantically annotated in Resource Description Framework (RDF) triple format to provide for ontology knowledge bases.

15

However, to query an Resource Description Framework (RDF) triple or ontology knowledge base, a user is required to have knowledge of an SQL-like language such as SPARQL or an Artificial Intelligence (AI) language such as PROLOG or LISP. While this is considered an acceptable option for developers and the technically inclined, general users need not be required to have such in-depth knowledge to perform a query on the semantic based web.

20

Although SPARQL or AI language query format has been well defined to query ontology knowledge bases with its powerful inference engine, the query layer remains as the building block for developing semantic technology applications. This is because a user is

25

required to know exactly how and where the information is stored in the knowledge base in order to perform a query on the same. This hinders the efforts of providing a simple and practical approach of performing a query on the semantic based web.

5. Furthermore, there has been a rising need to address the issue of meaningful query retrieval. In many occasions, it is observed that users have different ways of constructing a particular query that is required to yield identical query results. It is therefore essential to understand the purpose of a query, accurately interpret the same and provide the precise query results.

**SUMMARY OF INVENTION**

In one embodiment of the present invention is a method for extendable semantic query interpretation. The method comprises receiving a structured natural language user query, interpreting the structured natural language user query, generating from the structured natural language user query, an internal query representation statement, performing query enrichment to generate at least one enriched internal query representation statement, generating from the at least one enriched internal query representation statement, a knowledge base compliant query to provide for searching at least one query result from an ontology knowledge base, generating at least one internal query result representation statement and converting the at least one internal query result representation statement to a structured natural language result.

In another embodiment of the present invention is a system for extendable semantic query interpretation. The system comprises an intelligent word sense, a semantic query interpreter, a query transformer, a query enricher and a natural language generator. The intelligent word sense comprises means for receiving a structured natural language user query. The semantic query interpreter comprises means for interpreting the structured natural language user query. The query transformer comprises means for generating from the structured natural language user query, an internal query representation statement. The query enricher comprises means for performing query enrichment to generate at least one enriched internal query representation statement, generating from the at least one enriched internal query representation statement, a knowledge base compliant query to provide for searching at least one query result from an ontology knowledge base and generating at least one internal query result representation statement. The natural language generator comprises means for converting the at least

one internal query result representation statement to a structured natural language result.

The present invention consists of several features and a combination of parts hereinafter  
5 fully described and illustrated in the accompanying drawings, it being understood that  
various changes in the details may be made without departing from the scope of the  
invention or sacrificing any of the advantages of the present invention.

**BRIEF DESCRIPTION OF THE ACCOMPANYING DRAWINGS**

To further clarify various aspects of some embodiments of the present invention, a more particular description of the invention will be rendered by references to specific  
5 embodiments thereof, which are illustrated, in the appended drawings. It is appreciated that these drawings depict only typical embodiments of the invention and are therefore not to be considered limiting of its scope. The invention will be described and explained with additional specificity and detail through the accompanying drawings in which:

10 FIG. 1 illustrates a flowchart of a method for an extendable semantic query interpretation.

FIG. 2 illustrates a block diagram of a system for an extendable semantic query interpretation.

**DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

The present invention relates to a method and system for extendable semantic query interpretation. Hereinafter, this specification will describe the present invention according to the preferred embodiments of the present invention. However, it is to be understood that limiting the description to the preferred embodiments of the invention is merely to facilitate discussion of the present invention and it is envisioned that those skilled in the art may devise various modifications and equivalents without departing from the scope of the appended claims.

The embodiments of the present invention relate to a method and system for extendable semantic query interpretation. The present invention provides an extendable semantic query interpretation to enable a user to query information stored in ontology knowledge bases, which are essentially knowledge bases in Resource Description Framework (RDF) format, without having in-depth understanding of the underlying knowledge base.

The embodiments of the present invention simplifies the process of generating valid query statements by creating knowledge base compliant query statements from the user's query which is received in the form of a structured natural language query.

Further, the embodiments of the present invention also translate the query results obtained from the ontology knowledge base in the Resource Description Framework (RDF) format into a structured natural language result. Hence, this substantially eliminates the necessity for the user to explicitly understand the structure of the knowledge base to perform a query on the knowledge base. Additionally, the user is shielded from the complexity of the query processing while having the liberty of constructing complex queries.

The embodiments of the present invention addresses the issue of meaningful query retrieval by employing query enrichment techniques. With these techniques, the user's query terms residing in the form of a structured natural language query are expanded by making reference to a lexical database such as Wordnet. Additionally, different word senses, synonyms, antonyms and the like are used to expand the user's query to enable comprehensive and meaningful query processing.

Reference is now collectively being made to FIGs. 1 and 2. FIG. 1 illustrates a flowchart of the method (100) for an extendable semantic query interpretation. FIG. 2 illustrates a block diagram of the system (200) for an extendable semantic query interpretation.

The method (100) for the extendable semantic query interpretation comprises receiving a structured natural language user query (102), interpreting the structured natural language user query (104), generating from the structured natural language user query, an internal query representation statement (106), performing query enrichment (108) to generate at least one enriched internal query representation statement, generating from the at least one enriched internal query representation statement, a knowledge base compliant query (110) to provide for searching at least one query result from an ontology knowledge base, generating at least one internal query result representation statement (112) and converting the at least one internal query result representation statement to a structured natural language result (114).

The system (200) for the extendable semantic query interpretation comprises an intelligent word sense (202), a semantic query interpreter (206), a query transformer (208), a query enricher (210) and a natural language generator (218). These

components operate in tandem with a parser (204), a query module (214), a lexical database (212) and a knowledge base (220) as a semantic query engine for the purpose of providing the extendable semantic query interpretation while the user performs a query on the knowledge base (220).

5

The intelligent word sense (202) is the interface module between the user and the semantic query engine. It has the purpose of preprocessing the query received from the user for the ontology knowledge base query. The intelligent word sense (202) receives the user's query in the form of a structured natural language user query (102).

10

In some embodiments of the present invention, the intelligent word sense (202) provides a plurality of auto-completion suggestions for the terms of the structured natural language user query as the same is being received from the user. The auto-completion suggestions are obtained from a plurality of terms available in the ontology knowledge base (220). The plurality of terms is defined by a list of concepts and relations in the ontology knowledge base (220).

15

This feature of auto-completion suggestions assists the user while typing the query statement in the structured natural language form by providing auto-completion suggestions from terms available in the ontology knowledge base (220) as an option to be adopted by the user. This allows the user to provide a query statement without having knowledge of the variety of complex terms available in the ontology knowledge base (220). Further more, this feature provides the basis of efficient and accurate query statements where the occurrences typographical errors are reduced.

20

25

The query statement received as the structured natural language user query by the intelligent word sense (202) is then forwarded to the semantic query interpreter (206). In this module, the structured natural language user query (104) is interpreted before further processing by the query transformer (208).

5

The semantic query interpreter (206) comprises means for interpreting the structured natural language user query (104), which includes two steps. Firstly, using a parser (204) the structured natural language user query is parsed to identify a Part-of-Speech (POS) information of each term in the query statement. The parser (204) may comprise  
10 of a lexical parser. Upon parsing the structured natural language user query, the Part-of-Speech (POS) information is identified using a word-level tagging lookup.

The second step comprises identifying a type of query by matching the parsed structured natural language user query to a set of configurable query templates. From  
15 the parsed structured natural language user query, the sub-phrases are chunked together such that by matching the same with a set of configurable query templates, the type of query may be identified. The set of configurable query templates are a predefined set of possible identified query types.

20 The complexity of the structured natural language user query is determined at this stage that is whether the structured natural language user query is a union, intersection or negation type of query. The complexity of the user's query may be enhanced by the extensiveness of the set of configurable query templates.

25 For example, where the structured natural language user query is as follows:  
"What are the symptoms of essential hypertension",

parsing the structured natural language user query would yield as follows:

"What/**WP** are/**VBP** the/**DT** symptoms/**NNS** of/**IN** essential/**JJ** hypertension/**NN**"

From the parsed structured natural language user query, the sub-phrases such as "the  
5 symptoms" and "essential hypertension" are chunked together and then matched to the  
set of configurable query templates. Hence, in this example, the parsed structured  
natural language user query is classified as being the type of query as follows:

"What/**WP** is/**VBZ NP1 IN NP2**",

where **NP1** would refer to "the symptoms" and **NP2** would refer to "essential  
10 hypertension".

Once the type of query is identified to be within a particular template, the information  
pertaining to the parsed structured natural language user query and the type of query is  
forwarded to the query transformer (208) for further processing.

15 The query transformer (208) comprises a Subject-Predicate-Object transformer (208).  
The Subject-Predicate-Object transformer (208) comprises means for generating from  
the structured natural language user query, an internal query representation statement  
(106). The internal query representation statement is generated in the form of Resource  
20 Description Framework (RDF) triple format.

The Subject-Predicate-Object transformer (208) receives the parsed structured natural  
language user query and the type of query. Based on the Part-of-Speech (POS)  
information from the parsed structured natural language user query and the type of  
25 query of the same, the Subject-Predicate-Object transformer (208) extracts a subject  
part, a predicate part and an object part of the structured natural language user query.

The subject part, the predicate part and the object part are used to generate at least one internal query representation statement in the Resource Description Framework (RDF) triple format. The subject part refers to the root term of the query statement and the predicate part refers to the term of the query statement connecting the subject part to the object part. Since Resource Description Framework (RDF) stores triples of data in the ontology knowledge base (220), a query may be performed on either one of the subject part, predicate part or the object part.

10 The Subject-Predicate-Object transformer (208) then tags an argument of the at least one internal query representation statement by tagging the subject part, the predicate part or the object part for searching the at least one query result from the ontology knowledge base (220). The tag used for tagging the argument comprises a symbol normally a question mark, "?".

15

For example, based on the parsed structured natural language user query as follows:

"What/**WP** are/**VBP** the/**DT** symptoms/**NNS** of/**IN** essential/**JJ** hypertension/**NN**"

and

the type of query as follows:

20 "What/**WP** is/**VBZ** **NP1** **IN** **NP2**",

where **NP1** would refer to "the symptoms" and **NP2** would refer to "essential hypertension",

obtained from the semantic query interpreter (206), the Subject-Predicate-Object transformer (208) generates the internal query representation statement comprising a

25 Subject-Predicate-Object string as follows:

"essential hypertension; symptoms; ?What"

where “?What” is the tagged argument.

The internal query representation statement in the Resource Description Framework (RDF) triple format comprising a Subject-Predicate-Object string is then forwarded to the query enricher (210). The query enricher (210) comprises means for performing query enrichment (108) to generate at least one enriched internal query representation statement, generating from the at least one enriched internal query representation statement, a knowledge base compliant query (110) to provide for searching at least one query result from an ontology knowledge base (220) and generating at least one internal query result representation statement (112).

The query enricher (210) performs query enrichment (108) to generate the at least one enriched internal query representation statement using syntactic matching and semantic matching using a lexical database (212). Query enrichment enables comprehensive and meaningful query processing.

Syntactic matching determines the degree or score of matching between terms in the query statement and terms in the ontology knowledge base (220) using syntactic distance techniques such as Tf-Idf, Jaro Winkler and the like. If the matching score is higher than a predetermined threshold score value, which is manually set, the terms are considered matched and vice-versa. In the case where the terms are matched, the original term in the query statement will be substituted with the term identified from the ontology knowledge base (220).

Semantic matching identifies terms in the ontology knowledge base (220) that have similar meaning with terms in the internal query representation statement. Semantic

matching may have several level of matching including synonym matching, associative matching, "kind-of" matching and "part-of" matching. In the embodiments of the present invention, semantic matching is achieved by matching the terms in the internal query representation statement with terms in a lexical database (212) such as Wordnet from which synsets of a term are retrieved and utilized to generate the at least one enriched internal query representation statement. The synsets comprise nouns, verbs, adjectives and adverbs that are grouped into sets of cognitive synonyms, each expressing a distinct concept.

10 For example, query enrichment performed on the internal query representation statement as follows:

"essential hypertension; symptoms; ?What"

using syntactic matching and semantic matching, would possibly yield in three enriched internal query representation statement as follows:

15 "essential hypertension; **has\_Symptoms**; ?What"

"**High\_Blood\_Pressure**; has\_Symptoms; ?What"

"High\_Blood\_Pressure; **has\_Indication**; ?What"

The query enricher (210) then generates from the at least one enriched internal query representation statement, the knowledge base compliant query (110) to provide for searching the at least one query result from the ontology knowledge base (220). This begins with the query enricher (210) combining the at least one enriched internal query representation statement generated from query enrichment (108). The at least one enriched internal query representation statement is combined to form a valid query statement.

20

25

Using the valid query statement and based on the properties of the structured natural language user query, the query enricher (210) automatically identifies which knowledge base compliant query statement is to be generated, either a SPARQL query statement or a PROLOG query statement. This knowledge base compliant query statement is used  
5 for searching the at least one query result from the ontology knowledge base (220).

The query module (214) searches the at least one query result from the ontology knowledge base (220). The query module (214), specifically a Resource Description Framework (RDF) query module receives the knowledge base compliant query  
10 statement from the query enricher (210) in the form of Resource Description Framework (RDF) triple format for SPARQL query and PROLOG statement for PROLOG query. The same is then validated to generate a validated query statement. The query module (214) then uses the validated query statement to search for the at least one query result in the ontology knowledge base (220) residing in an appropriate knowledge base server. The  
15 query module (214) then returns the at least one query result from the ontology knowledge base (220) to the query enricher (210).

For example, where the internal query representation statement is as follows:

“essential hypertension; symptoms; ?What”,  
20 the query module (214) returns “headache”, “visual\_disturbance” and “vomiting” as the at least one query result obtained from the ontology knowledge base (220).

The query enricher (210) finally generates the at least one internal query result representation statement (112) by integrating the at least one query result obtained from  
25 the query module (214) with the internal query representation statement generated by the Subject-Predicate-Object transformer (208). This is performed by an answer modifier

module of the query enricher (210). The answer modifier module replaces the tagged argument of the internal query representation statement with the at least one query result. The answer modifier then tags the at least one query result to generate the at least one internal query result representation statement. The tag used for tagging the at least one query result comprises a symbol normally an exclamation question mark, "!".

For example, using the internal query representation statement as follows:

"essential hypertension; symptoms; ?What",

and the at least one query result as follows:

10 "headache", "visual\_disturbance" and "vomiting"

the query enricher (210) generates three internal query result representation statement as follows:

"essential hypertension; symptoms; !headache"

"essential hypertension; symptoms; !visual\_disturbance"

15 "essential hypertension; symptoms; !vomiting".

The final step for the extendable semantic query interpretation comprises the natural language generator (218) converting the at least one internal query result representation statement to a structured natural language result (114). The natural language generator (218) generates the structured natural language result using the structured natural language user query and the at least one internal query result representation statement. More specifically, the natural language generator (218) utilizes the Part-of-Speech (POS) information from the parsed structured natural language user query to generate the structured natural language result by considering linguistic features such as tense, part of speech, vowels, consonants and the like.

The at least one internal query result is provided to the user as the structured natural language result instead of merely a list of query results. Based on the structured natural language user query and type of query, the structured natural language result may comprise of a plurality of structured natural language result.

5

For example, the structured natural language result for the structured natural language query "What are the symptoms of essential hypertension?", is as follows:

"The symptoms of essential hypertension are:

Headache

10 Visual\_disturbance

Vomiting".

In another embodiment of the present invention, in the case where a plurality of internal query representation statements is generated, multiple arguments will be tagged. For this embodiment, the structured natural language result will be either a union or intersection of the at least one internal query result obtained from the query module (214). The union or intersection operation is dependant on the type of query as identified by the semantic query interpreter (206).

15

20 Additionally, by using PROLOG statement for PROLOG query, it is possible to create rules that may infer new connections or relationships between concepts in the ontology knowledge base (220). These new connections or relationships may be used for searching the at least one query result from the ontology knowledge base (220). For example, in a given ontology knowledge base that has a "hasParent" relationship, it is possible to infer other relationships such as 'grandparent' and 'sibling'.

25

**CLAIMS**

1. A method (100) for an extendable semantic query interpretation, the method (100) comprises

5           receiving a structured natural language user query (102);  
          interpreting the structured natural language user query (104);  
          generating from the structured natural language user query, an internal  
          query representation statement (106);  
          performing query enrichment (108) to generate at least one enriched  
10          internal query representation statement;  
          generating from the at least one enriched internal query representation  
          statement, a knowledge base compliant query (110) to provide for  
          searching at least one query result from an ontology knowledge base;  
          generating at least one internal query result representation statement  
15          (112); and  
          converting the at least one internal query result representation statement  
          to a structured natural language result (114).

- 20          2. The method (100) according to claim 1, wherein the internal query representation  
          statement and the at least one internal query result representation statement  
          comprises Resource Description Framework (RDF) triple format.

- 25          3. The method (100) according to claim 1, wherein receiving the structured natural  
          language user query (102) further comprises providing from a plurality of terms in  
          the ontology knowledge base, a plurality of auto-completion suggestions for at  
          least one term of the structured natural language user query.

4. The method (100) according to claim 1, wherein interpreting the structured natural language user query (104) further comprises

5 parsing the structured natural language user query to identify a Part-of-Speech (POS) information; and

identifying a type of query by matching the parsed structured natural language user query to a set of configurable query templates.

5. The method (100) according to claim 1 and 4, wherein generating from the structured natural language user query, the internal query representation statement (106) further comprises

10 extracting a subject part, a predicate part and an object part based on the Part-of-Speech (POS) information and the type of query to generate at least one internal query representation statement; and

15 tagging an argument of the at least one internal query representation statement by tagging either the subject part, the predicate part or the object part for searching the at least one query result from the ontology knowledge base.

- 20 6. The method (100) according to claim 1, wherein performing query enrichment (108) to generate the at least one enriched internal query representation statement query further comprises

syntactic matching; and

semantic matching using a lexical database.

7. The method (100) according to claim 1, wherein generating from the at least one enriched internal query representation statement, the knowledge base compliant query (110) to provide for searching the at least one query result from the ontology knowledge base further comprises

5 combining the at least one enriched internal query representation statement to form a valid query statement; and  
generating from the valid query statement, a SPARQL query statement or a PROLOG query statement.

10 8. The method (100) according to claim 1 and 5, wherein generating the at least one internal query result representation statement (112) further comprises

replacing the tagged argument of the internal query representation statement with the at least one query result; and  
tagging the at least one query result to generate the at least one internal  
15 query result representation statement.

9. The method (100) according to claim 1, wherein converting the at least one internal query result representation statement to the structured natural language result (114) further comprises generating the structured natural language result  
20 using the structured natural language user query and the at least one internal query result representation statement.

10. A system (200) for an extendable semantic query interpretation, the system (200) comprises

25 an intelligent word sense (202);  
a semantic query interpreter (206);

a Subject-Predicate-Object transformer (208);

a query enricher (210); and

a natural language generator (218);

characterized in that

5           the intelligent word sense (202) comprises means for receiving a structured natural language user query;

the semantic query interpreter (206) comprises means for interpreting the structured natural language user query;

10           the query transformer (208) comprises means for generating from the structured natural language user query, an internal query representation statement;

the query enricher (210) comprises means for

performing query enrichment to generate at least one enriched internal query representation statement;

15           generating from the at least one enriched internal query representation statement, a knowledge base compliant query to provide for searching at least one query result from an ontology knowledge base (220); and

20           generating at least one internal query result representation statement;

the natural language generator (218) comprises means for converting the at least one internal query result representation statement to a structured natural language result.

11. The system (200) according to claim 10, wherein the internal query representation statement and the at least one internal query result representation statement comprises Resource Description Framework (RDF) triple format.
- 5 12. The system (200) according to claim 10, wherein the intelligent word sense (202) further comprises means for providing from a plurality of terms in the ontology knowledge base (220), a plurality of auto-completion suggestions for at least one term of the structured natural language user query.
- 10 13. The system (200) according to claim 10, wherein the semantic query interpreter (206) further comprises means for
- parsing the structured natural language user query to identify a Part-of-Speech (POS) information using a parser (204); and
- identifying a type of query by matching the parsed structured natural language user query to a set of configurable query templates.
- 15
14. The system (200) according to claim 10 and 13, wherein the query transformer (208) comprises a Subject-Predicate-Object transformer (208) characterized in that the Subject-Predicate-Object transformer (208) comprises means for
- 20       extracting a subject part, a predicate part and an object part based on the Part-of-Speech (POS) information and the type of query to generate at least one internal query representation statement; and
- tagging an argument of the at least one internal query representation statement by tagging either the subject part, the predicate part or the
- 25       object part for searching the at least one query result from the ontology knowledge base (220).

15. The system (200) according to claim 10 and 14, wherein the query enricher (210)

further comprises means for

syntactic matching;

5 semantic matching using a lexical database (212);

combining the at least one enriched internal query representation  
statement to form a valid query statement;

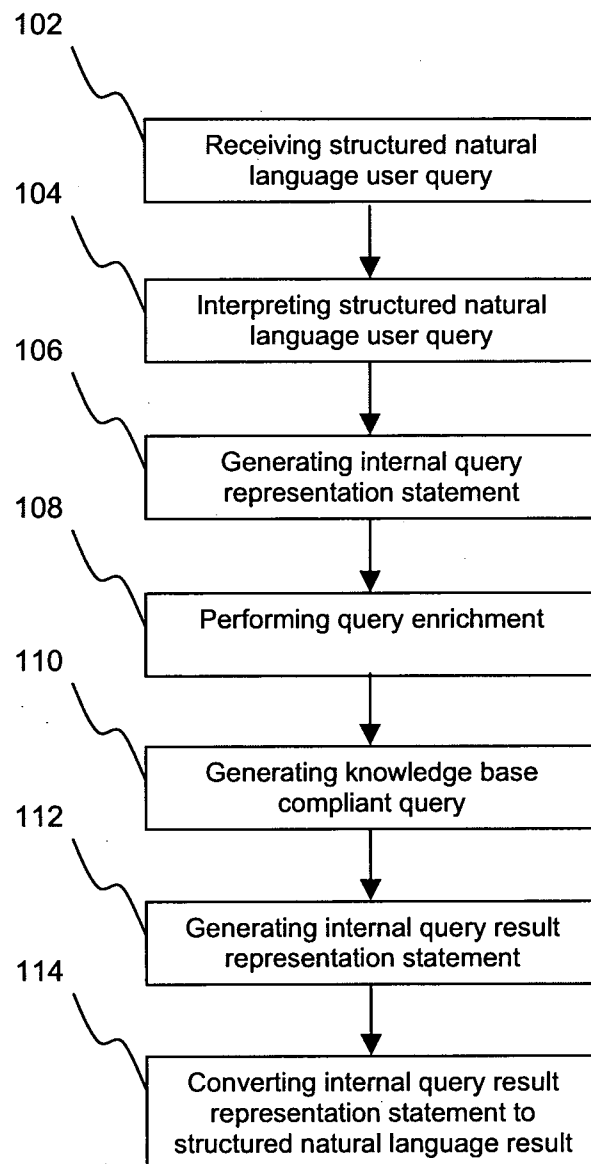
generating from the valid query statement, a SPARQL query statement or  
a PROLOG query statement;

10 replacing the tagged argument of the internal query representation  
statement with the at least one query result; and

tagging the at least one query result to generate the at least one internal  
query result representation statement.

15 16. The system (200) according to claim 10, wherein the natural language generator  
(218) further comprises means for generating the structured natural language  
result using the structured natural language user query and the at least one  
internal query result representation statement.

20 17. The system (200) according to claim 10, wherein the system (200) further  
comprises a query module (214) for searching the at least one query result from  
the ontology knowledge base (220).

**100****FIG. 1**

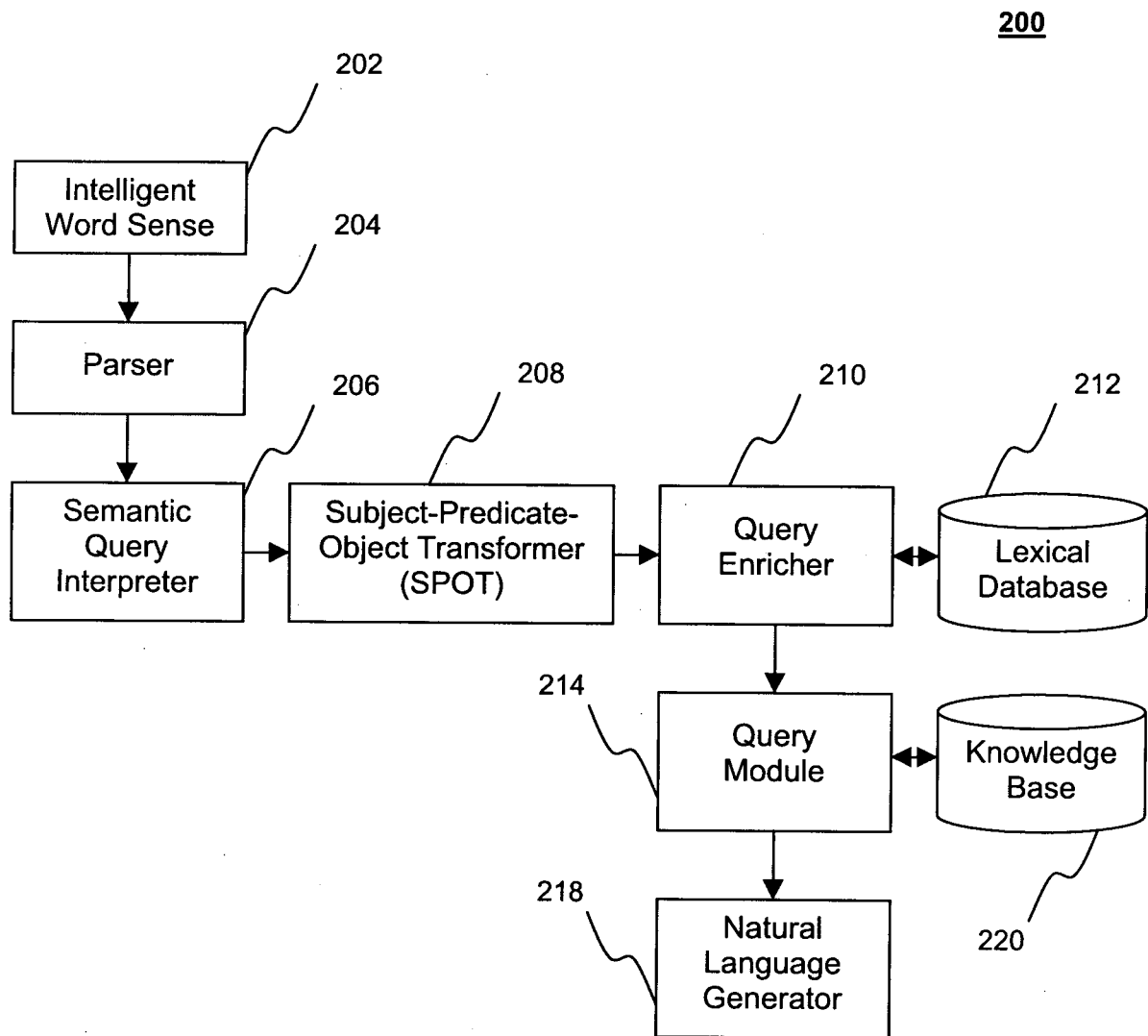


FIG. 2

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/MY2010/000084

## A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

G06F 17/27 (2006.01) G06F 17/00 (2006.01) G06F 17/20 (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)

G06F

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, Google Patents, Google Scholar, ESP@CE, USPTO: semantic, query, interpretation, ontology, rdf, natural, language and other similar terms

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                          | Relevant to claim No. |
|-----------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| X         | LIANG J. et al, 'Ontology-Based Natural Language Query processing for the Biological Domain', Proceedings of the BioNLP Workshop on Linking Natural Language Processing and Biology at HLT-NAACL 2006, pages 9-16 [retrieved on 07 July 2010] Retrieved from Internet <URL: | 1-3, 6-7, 9-12, 16-17 |
| Y         | <a href="http://www.aclweb.org/anthology/W/W06/W06-3302.pdf">http://www.aclweb.org/anthology/W/W06/W06-3302.pdf</a><br>(see the abstract, section 2, 2.2-2.3 and 2.5, figures 1-3)                                                                                          | 4-5, 13-14            |
| X         | US 5197005 (SHWARTZ et al) 23 March 1993<br>(see the abstract)                                                                                                                                                                                                              | 1-3, 7, 9             |
| X         | US 5963940 A (LIDDY et al) 05 October 1999<br>(see the abstract, column 2 lines 48-49, figure 3, column 8 lines 37-38)                                                                                                                                                      | 1-3, 7, 9             |
| Y         |                                                                                                                                                                                                                                                                             | 4                     |
| Y         | US 2003/0069880 A1 (HARRISON et al) 10 April 2002<br>(see the abstract, paragraph [0029])                                                                                                                                                                                   | 4-5, 13-14            |

☒ 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.  
06 August 2010

Date of mailing of the international search report

10 AUG 2010

Name and mailing address of the ISA/AU  
AUSTRALIAN PATENT OFFICE  
PO BOX 200, WODEN ACT 2606, AUSTRALIA  
E-mail address: pct@ipaustalia.gov.au  
Facsimile No. +61 2 6283 7999

Authorized officer  
**ANDREW MUNRO**  
AUSTRALIAN PATENT OFFICE  
(ISO 9001 Quality Certified Service)  
Telephone No : +61 3 9935 9620

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/MY2010/000084

| C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |                                                                                                                                                                                                                                                                                                                                                                                                                                           |                       |
|-------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Category*                                             | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                                                                                                                                        | Relevant to claim No. |
| A                                                     | WO 2008/100849 A2 (CYCORP, INC) 21 August 2008<br>(see whole document)                                                                                                                                                                                                                                                                                                                                                                    |                       |
| A                                                     | US 2007/0073680 A1 (KAWAMURA et al) 29 march 2007<br>(see whole document)                                                                                                                                                                                                                                                                                                                                                                 |                       |
| A                                                     | US 2005/0289124 A1 (KAISER et al) 29 December 2005<br>(see whole document)                                                                                                                                                                                                                                                                                                                                                                |                       |
| A                                                     | US 2006/0036512 A1 (MACCARTHY et al) 16 February 2006<br>(see whole document)                                                                                                                                                                                                                                                                                                                                                             |                       |
| A                                                     | TOMASSEN S. et al, 'Document Space Adapted Ontology: Application in Query Enrichment', 11th International Conference on Applications of Natural Language to Information Systems, NLDB 2006, vol. 3999, pages 46-57 [retrieved on 07 July 2010] Retrieved from Internet <URL: <a href="http://folk.ntnu.no/steint/publications.php">http://folk.ntnu.no/steint/publications.php</a> ><br>(see whole document)                              |                       |
| A                                                     | US 2002/0152202 A1 (PERRO et al) 17 October 2002<br>(see whole document)                                                                                                                                                                                                                                                                                                                                                                  |                       |
| A                                                     | US 2007/0016563 A1 (OMOIGUI) 18 January 2007<br>(see whole document)                                                                                                                                                                                                                                                                                                                                                                      |                       |
| A                                                     | BROEKSTRA J. et al, 'Sesame: A Generic Architecture for Storing and Querying RDF and RDF Schema', Administrator Nederland b.v., 01 October 2001 [retrieved on 07 July 2010] Retrieved from Internet <URL: <a href="http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.9588&amp;rep=rep1&amp;type=pdf">http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.103.9588&amp;rep=rep1&amp;type=pdf</a> ><br>(see whole document) |                       |
| A                                                     | US 2006/0100995 A1 (ALBORNOZ et al) 11 May 2006<br>(see whole document)                                                                                                                                                                                                                                                                                                                                                                   |                       |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/MY2010/000084

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                                                                                                                        | 5197005    | NONE           |                      |            |                |
| US                                                                                                                        | 5963940    | EP 0845102     | US                   | 5861241    | US 5996402     |
|                                                                                                                           |            | US 6026388     | WO                   | 9708528    |                |
| US                                                                                                                        | 2003069880 | US 7403938     | US                   | 2008263019 |                |
| WO                                                                                                                        | 2008100849 | CN 101681348   | EP                   | 2109831    | US 2008109212  |
|                                                                                                                           |            | US 7734623     |                      |            |                |
| US                                                                                                                        | 2007073680 | JP 2007094775  |                      |            |                |
| US                                                                                                                        | 2005289124 | US 7720674     |                      |            |                |
| US                                                                                                                        | 2006036512 | AU 2002334759  | CA                   | 2467735    | EP 1573438     |
|                                                                                                                           |            | WO 2004031884  |                      |            |                |
| US                                                                                                                        | 2002152202 | WO 03019421    |                      |            |                |
| US                                                                                                                        | 2007016563 | AU 2004213986  | BR                   | 0210589    | BR P10407451   |
|                                                                                                                           |            | CA 2555280     | CN                   | 1647070    | CN 1774814     |
|                                                                                                                           |            | CN 1853180     | EP                   | 1410258    | EP 1597768     |
|                                                                                                                           |            | EP 1599811     | EP                   | 1749259    | EP 1889233     |
|                                                                                                                           |            | EP 1952266     | EP                   | 1952286    | KR 20050098958 |
|                                                                                                                           |            | KR 20060004909 | MX                   | PA03011976 | MX PA05008670  |
|                                                                                                                           |            | SG 160209      | US                   | 2004161868 | US 6946715     |
|                                                                                                                           |            | US 2004232314  | US                   | 7407830    | US 2003126136  |
|                                                                                                                           |            | US 2004230572  | US                   | 2006026113 | US 2007038610  |
|                                                                                                                           |            | US 2007081197  | US                   | 2007260580 | US 2008016036  |
|                                                                                                                           |            | US 2008147716  | US                   | 2008147788 | US 2008162498  |
|                                                                                                                           |            | US 2008288456  | US                   | 2009050944 | US 2009254510  |
|                                                                                                                           |            | US 2010070448  | WO                   | 03001413   | WO 2004075299  |
|                                                                                                                           |            | WO 2004075466  | WO                   | 2005103883 | WO 2005111789  |
|                                                                                                                           |            | WO 2006124952  | WO                   | 2007044865 | WO 2007059536  |
|                                                                                                                           |            | WO 2008014499  |                      |            |                |
| US                                                                                                                        | 2006100995 | CN 1783124     | US                   | 7478105    | US 2009132528  |
| Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001. |            |                |                      |            |                |
| END OF ANNEX                                                                                                              |            |                |                      |            |                |