(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 July 2011 (21.07.2011)

PCT

(10) International Publication Number
WO 2011/087358 A2

(51) International Patent Classification: Not classified

(21) International Application Number: PCT/MY2010/000297

(22) International Filing Date: 25 November 2010 (25.11.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PI 2010000219 15 January 2010 (15.01.2010) 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*): **MOHAMED, Hassan** [MY/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **AB. RAHMAN, Suhaimi** [MY/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **ANAND, Arun** [IN/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **DICKSON, Lukose** [AU/MY]; Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY).

(74) Agent: **DAMODHARAN, Ramakrishna**; Kass International Sdn. Bhd., Suite 8 - 7 - 2 Menara Mutiara Bangsar, JaIan Liku Off JaIan Riong, Bangsar, 57000 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:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: A SEMANTIC INTERPRETER SYSTEM AND METHODS THEREOF

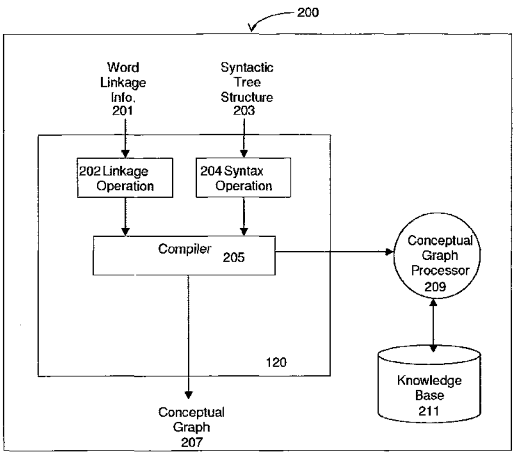

Figure 2

(57) **Abstract**: A semantic interpretation system (200) based on word linkage information (201) and contextual meaning of a plurality of natural language sentences from a syntactic tree structure (203) is provided, the system (200) includes a semantic interpretation processor (120), which further includes a compiler (205) connectable to a linkage operation module (202) and a syntax operation module (204), a conceptual graph processor (209) connectable to the semantic interpretation processor (120) and a means to access a knowledge base (211), connectable to the conceptual graph processor (209), wherein an output of the semantic interpretation processor (120) is at least one conceptual graph and a plurality of inputs to the system (200) further include word linkage information (201) and syntactic tree structure (203).

WO 2011/087358 A2

A SEMANTIC INTERPRETER SYSTEM AND METHODS THEREOF

FIELD OF INVENTION

The present invention relates to a semantic interpretation system based on word linkage information and contextual meaning of a plurality of natural language sentences from a syntactic tree structure, and methods thereof.

BACKGROUND OF INVENTION

Most of the current natural language processing (NLP) based application system has a 'semantic gap' between a syntactic structure to a semantic structure or meaning of natural language sentences. Therefore, software developers need to extract a 'meaning' of the natural language sentences, just as how humans understand the language. The meaning needs to be understood by computers in order to do further processing for reasoning and making intelligent decisions.

U.S. 2007/0100604 A1 describes a method of performing semantic analysis to interpret a linguistic structure which is done by getting an input from linguistic analysis component which are Unspecified Discourse Representation Structure (UDRS), Syntax Parse Tree, Logical Form, Tokenized String and Set of Entities. The inputs are then passed through to the semantic analysis component. An entity-and-relation model of a non-linguistic domain (application schema) is produced. However, the document relies on linguistic analysis for interpretation. This method may not accurately depict intent of the user as it focuses more in the syntactic structure of language.

U.S. 6292767 describes a method for building and running natural language understanding. The document describes taking a sentence as an input and returning a representation of possible meanings of the sentence as output using a run-time interpreter. In this document, the user is required to specify semantics of the application. The runtime interpreter makes use of this information along with a grammar file to do phrase matching. However, the user shoulders the burden of specifying semantics of natural language sentences that are being used, which adds an additional step to the process and further delays progress.

Therefore, there is a need for a solution to utilize human language understanding as a basis for developing intelligent applications and processing natural language sentences.

SUMMARY OF INVENTION

Accordingly there is provided a semantic interpretation system based on word linkage information and contextual meaning of a plurality of natural language sentences from a syntactic tree structure, the system includes a semantic interpretation processor, which further includes a compiler connectable to a linkage operation module and a syntax operation module, a conceptual graph processor connectable to the semantic interpretation processor and a means to access a knowledge base connectable to the conceptual graph processor, wherein an output of the semantic interpretation processor is at least one conceptual graph and a plurality of inputs to the system further includes word linkage information and syntactic tree structure.

There is also provided a method of semantic interpretation based on word linkage information and contextual meaning of a plurality of natural language sentences from a syntactic tree structure, the method includes the steps of requesting a service from a semantic interpretation processor to get the contextual meaning of the plurality of sentences, receiving the syntactic tree structure and web linkage information of a natural language sentence, executing a plurality of independent semantic interpreter related algorithms, manipulating a plurality of conceptual graphs in a knowledge base, sending a reference of resulting conceptual graph back to origin of the request and producing a semantic representation of the natural language sentence.

There is further provided a method of semantic interpretation based on word linkage information and contextual meaning of a plurality of natural language sentences, the method includes the steps of requesting a service from a semantic interpretation processor to get the contextual meaning of the plurality of sentences, receiving the web linkage information of a natural language sentence, executing a plurality of independent semantic interpreter related algorithms, manipulating a plurality of conceptual graphs in a knowledge base, sending a reference of resulting conceptual graph back to origin of the request and producing a semantic representation of the natural language sentence.

There is further provided a method of semantic interpretation based contextual meaning of a plurality of natural language sentences from a syntactic tree structure, the method includes the steps of requesting a service from a semantic interpretation processor to get the contextual meaning of the plurality of sentences, receiving the syntactic tree structure of a natural language sentence, executing a plurality of independent semantic interpreter related algorithms, manipulating a plurality of conceptual graphs in a knowledge base, sending a reference of resulting conceptual graph back to origin of the request and producing a semantic representation of the natural language sentence.

The present invention consists of several novel features and a combination of parts hereinafter fully described and illustrated in the accompanying description and 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 DRAWINGS

The present invention will be fully understood from the detailed description given herein below and the accompanying drawings which are given by way of illustration only, and thus are not limitative of the present invention, wherein:

Figure 1 is a flowchart illustrating a preferred embodiment of a method of semantic interpretation; and

Figure 2 is a block diagram illustrating architecture of a preferred embodiment of a semantic interpretation system.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention relates to a semantic interpretation system based on word linkage information and contextual meaning of a plurality of natural language sentences from a syntactic tree structure, and methods thereof. 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 following detailed description of the preferred embodiments will now be described in accordance with the attached drawings, either individually or in combination.

The present invention provides a semantic interpretation system (200) based on word linkage information (201) and contextual meaning of a plurality of natural language sentences from a syntactic tree structure (203) as seen in Figure 2. The system (200) includes a semantic interpretation processor (120), which further includes a compiler (205) connectable to a linkage operation module (202) and a syntax operation module (204), a conceptual graph processor (209) connectable to the semantic interpretation processor (120) and a means to access a knowledge base (211), connectable to the conceptual graph processor (209), wherein an output of the semantic interpretation processor (120) is at least one conceptual graph and a plurality of inputs to the system (200) further include word linkage information (201) and syntactic tree structure (203).

The compiler (205) is a main processing centre for the system (200) where semantic interpreter related algorithms are executed.

An example of the knowledge base (211) is a conceptual graph knowledge base. The knowledge base (211) represents hierarchical structure and relations between concepts and relations in natural language. The knowledge base (211) can be extended to incorporate various domains such as medical, engineering and computing. The present embodiment of the invention includes a means to access a knowledge base (211) populated for general as well as a medical domain. The conceptual graph knowledge base is constructed of concept type hierarchy, relation type hierarchy, type definitions, schemas, prototypes and instances. Those skilled in the art would appreciate that the above mentioned concepts form an integral part of conceptual graph theory.

As seen in Figure 1, a method of semantic interpretation based on word linkage information (201) and contextual meaning of a plurality of natural language sentences from a syntactic tree structure (203) is described. The method includes the steps of requesting a service from a semantic interpretation processor (120) to get the contextual meaning of the plurality of sentences, receiving the syntactic tree structure (203) and web linkage information (201) of a natural language sentence. Further, a plurality of independent semantic interpreter related algorithms are executed. A plurality of conceptual graphs is manipulated in a knowledge base (211) and a reference of resulting conceptual graph is sent back to origin of the request. Finally, a semantic representation of the natural language sentence is produced.

An example of the method is described further herein as seen in Figure 1. A domain application (110) requests a service to the semantic interpretation processor (120) to get a meaning representation of a natural language sentence. This is done by inputting a syntactic tree structure (203) of natural language sentence to the semantic interpretation processor (120). The semantic interpretation processor (120) processes the input and executes a predetermined number of independent semantic interpreter algorithms. In this example, three independent semantic interpreter methods are executed. Then, a request for a conceptual graph is sent to a conceptual graph processor (209) which manipulates the conceptual graph in the knowledge base (211). A reference to the resultant conceptual graph is then sent back to the semantic interpretation processor (120) and forwarded to the domain application (110).

In this example, the semantic interpretation is performed by combining functionality of the syntactic tree structure (203) algorithm and the word linkage information (201) algorithm. In certain cases, the syntactic tree structure (203) produces incomplete information about syntax of a sentence. For example, this occurs when multiple words are grouped together.

Syntactic tree structure (203):

[S [NP [NP The purpose NP] [PP of [NP this study NP] PP] NP] [VP was [S [VP to [VP determine [NP [NP the prevalence NP] [PP of [NP high blood pressure NP] PP] NP] [PP in [NP school children NP] PP] VP] VP] S] VP] . S]

Here, the words "high blood pressure" and "school children" are grouped together. In such cases, the word linkage information (201) algorithm is applied to the grouped words that form a block of data. This allows the system (200) to identify a required method to be used to generate the conceptual graph for the block of data.

Word linkage information (201):

[(LEFT-WALL)(the)(purpose.n)(of)(this.d)(study.n)(was.v)(to)(determine.v)(the)(prevalence.n)(of)(high.n)(blood.n)(pressure.n)(in)(school.n)(children.n)(.)]

[[0 18 4 (Xp)][0 2 1 (Wd)][1 2 0 (D*u)][2 6 2 (Ss)][2 3 0 (Mp)][3 5 1 (Jp)][4 5 0 (D*u)]

[6 7 0 (TO)][7 8 0 (I)][8 15 3 (MVp)][8 10 1 (Os)][9 10 0 (Dmu)][10 11 0 (Mp)][11 14 2 (Jp)]

[12 14 1 (AN)][13 14 0 (AN)][15 17 1 (Jp)][16 17 0 (AN)]]

[0 10]

A method of semantic interpretation based on word linkage information (201) and contextual meaning of a plurality of natural language sentences without using a syntactic tree structure (203) is also described. The method includes the steps of requesting a service from a semantic interpretation processor (120) to get the contextual meaning of the plurality of sentences, receiving the web linkage information (201) of a natural language sentence, executing a plurality of independent semantic interpreter related methods, manipulating a plurality of conceptual graphs in a knowledge base (211), sending a reference of resulting conceptual graph back to origin of the request and producing a semantic representation of the natural language sentence.

To illustrate the method, consider a sentence "*The cushions on these chairs need replacing*". The word linkage information (201) received by the semantic interpretation processor (120) through a Linkage Operation module (202) is shown below:

[(LEFT-WALL)(the)(cushions.n)(on)(these)(chairs.n)(need.v)(replacing.g)(.)]

[[0 8 (Xp)][0 2 (Wd)][1 2 (Dmc)][2 6 (Sp)][2 3 (Mp)][3 5 (Jp)][4 5 (Dmc)][6 7 (Os)]]

By analyzing the word linkage information (201), we construct a potential relation and establish rules between words such as:

a) The word *the* and *cushions* are linked by *Dmc* which means *connect determiners to noun*, hence the conceptual graph would be

[cushions: *].

b) The word *cushions* and *need* are linked by *Sp* which means *connects plural nouns to plural verb forms*, hence the conceptual graph would be

[cushions] ← (agent) ← [need]

c) The word *cushions*, *chairs* and *on* are linked by *Mp* and *Jp* respectively which *means connects prepositions to verbs (Mp)* and *connects prepositions to their objects (Jp).*

[cushions] ← (on) ← [chairs]

d) The word *these* and *chairs* are linked by *Dmc* which means *connect determiners to noun*, hence the conceptual graph would be

[chairs: *].

e) The word *need* and *replacing* are linked by *Os* which means *connects transitive verbs to direct or indirect object*, hence the conceptual graph would be

[need] ← (action) ← [replacing]

Relations **(agent)** in (b) and **(action)** in (e) will be verified by a schema defined in the knowledge base (211). If the relations are complying with the schema, interpretation of a complete sentence is formed by joining a canonical graph to form a large graph that represents the entire sentence. In this example the large conceptual graph is as below:

[cushions: *] –

← (agent) ← [need] ← (action) ← [replacing]

← (on) ← [chairs]

A method of semantic interpretation based contextual meaning of a plurality of natural language sentences from a syntactic tree structure (203) without using word linkage information (201) is described. The method includes the steps of requesting a service from a semantic interpretation processor (120) to get the contextual meaning of the plurality of sentences, receiving the syntactic tree structure (203) of a natural language sentence, executing a plurality of independent semantic interpreter related algorithms, manipulating a plurality of conceptual graphs in a knowledge base (211), sending a reference of resulting conceptual graph back to origin of the request and producing a semantic representation of the natural language sentence.

The system (200) receives the syntactic tree structure (203) of a natural language sentence as an input from the domain application (110) and passes it to the semantic interpretation processor (120). The semantic interpretation processor (120) constructs a conceptual graph operation list including a list of graph operations such as, but not limited to, creating graphs and maximal join. This is achieved by applying a semantic interpretation algorithm. The conceptual graph operation list is sent to the conceptual graph processor (209) for performing graph operations. A reference of the resultant conceptual graph is received by the semantic interpretation processor (120), thereby producing a semantic representation of the natural language sentence that is received as input in the form of the syntactic tree structure (203). The resultant graph is stored in the knowledge base (211) and the reference is returned to the domain application (110).

Example:

Given the syntactic tree structure (203) for a sentence "john opens the door with a key", the syntactic tree structure (203) and conceptual graph for the sentence are shown below:

Syntactic Tree Structure (203):

[S [NP john NP] [VP opens [NP the door NP] [PP with [NP a key NP] PP] VP] S]

The conceptual graph generated will be of the following form or equivalent:

$$
\begin{array}{c}
\text{[PERSON: John]} \leftarrow \text{(AGNT)} \leftarrow \text{[OPEN]} \rightarrow \text{(OBJ)} \rightarrow \text{[DOOR]} \\
\downarrow \\
\text{(INST)} \\
\downarrow \\
\text{[KEY]}
\end{array}
$$

Therefore, any ambiguities that rise with one method may be solved by combining usage of the methods to improve accuracy of representation of conceptual graphs.

The present invention is more logic-based and ontology based compared to semantic analysis of prior art documents, wherein the present invention is more focused on recognizing intended concepts as opposed to purely translating from syntactics to semantics, without manually having to specify any semantics.

The described methods and system (200) can be applied, but not restricted to, for systems that are domain specific applications which require interpretation of natural language to understand actual meaning. The domain application that uses the system (200) will be more intelligent in terms of meaning based reasoning. A successful implementation of the system (200) would generate conceptual graphs representing the meaning of natural language sentences. Therefore, the described system (200) and methods describe generating conceptual graphs representing syntactical meanings of sentences.

CLAIMS

1. A semantic interpretation system (200) based on word linkage information (201) and contextual meaning of a plurality of natural language sentences from a syntactic tree structure (203), the system (200) includes:

   a semantic interpretation processor (120), which further includes a compiler (205) connectable to a linkage operation module (202) and a syntax operation module (204);

   a conceptual graph processor (209) connectable to the semantic interpretation processor (120); and

   a means to access a knowledge base (211), connectable to the conceptual graph processor (209),

   wherein an output of the semantic interpretation processor (120) is at least one conceptual graph and a plurality of inputs to the system (200) further includes word linkage information (201) and syntactic tree structure (203).

2. The system (200) as claimed in claim 1, wherein the knowledge base (211) is a conceptual graph knowledge base.

3. The system (200) as claimed in claim 1, wherein the knowledge base (211) represents hierarchical structure and relations between concepts and relations in natural language.

4. The system (200) as claimed in claim 1, wherein the compiler (205) is a main processing centre for the system (200) where semantic interpreter related methods are executed.

5. A method of semantic interpretation based on word linkage information (201) and contextual meaning of a plurality of natural language sentences from a syntactic tree structure (203), the method includes the steps of:
    i. requesting a service from a semantic interpretation processor (120) to get the contextual meaning of the plurality of sentences,
    ii. receiving the syntactic tree structure (203) and web linkage information (201) of a natural language sentence;
    iii. executing a plurality of independent semantic interpreter related algorithms;
    iv. manipulating a plurality of conceptual graphs in a knowledge base (140);
    v. sending a reference of resulting conceptual graph back to origin of the request; and
    vi. producing a semantic representation of the natural language sentence.

6. A method of semantic interpretation based on word linkage information (201) and contextual meaning of a plurality of natural language sentences, the method includes the steps of:
    i. requesting a service from a semantic interpretation processor (120) to get the contextual meaning of the plurality of sentences,
    ii. receiving the web linkage information (201) of a natural language sentence;
    iii. executing a plurality of independent semantic interpreter related algorithms;
    iv. manipulating a plurality of conceptual graphs in a knowledge base (211);

v. sending a reference of resulting conceptual graph back to origin of the request; and

vi. producing a semantic representation of the natural language sentence.

7. A method of semantic interpretation based contextual meaning of a plurality of natural language sentences from a syntactic tree structure (203), the method includes the steps of:

i. requesting a service from a semantic interpretation processor (120) to get the contextual meaning of the plurality of sentences,

ii. receiving the syntactic tree structure (203) of a natural language sentence;

iii. executing a plurality of independent semantic interpreter related algorithms;

iv. manipulating a plurality of conceptual graphs in a knowledge base (211);

v. sending a reference of resulting conceptual graph back to origin of the request; and

vi. producing a semantic representation of the natural language sentence.

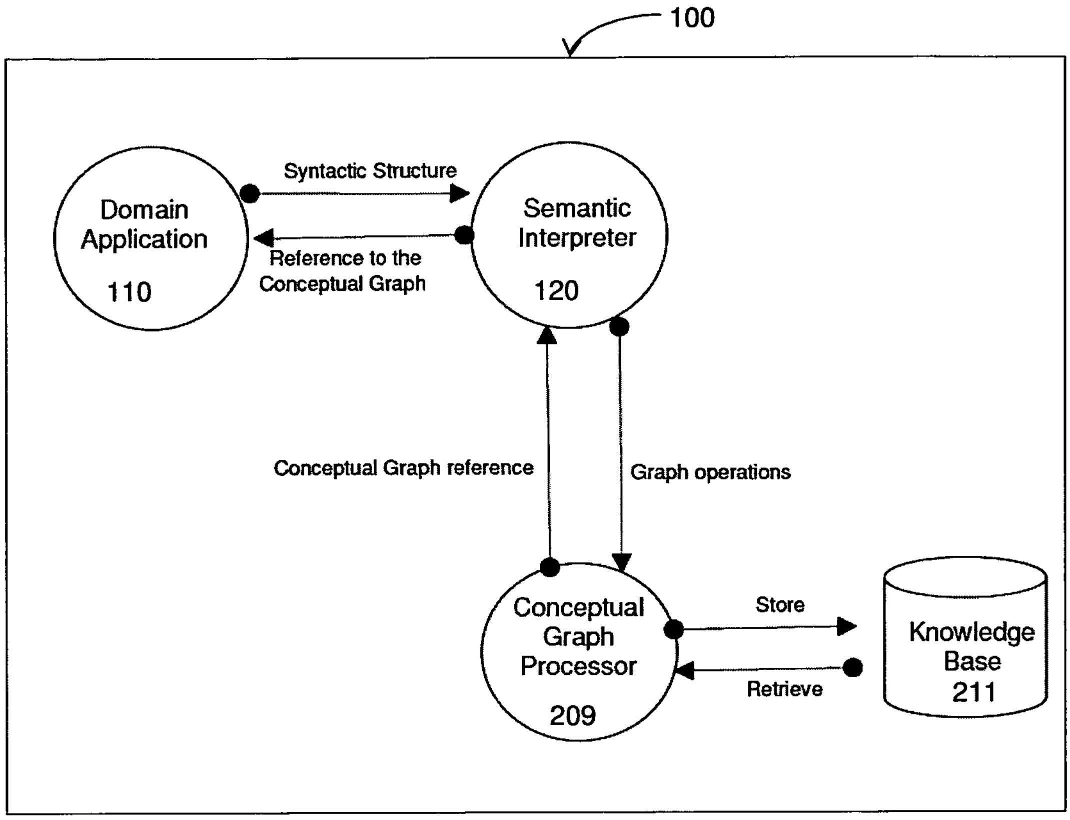
100
Domain Application
110
Syntactic Structure
Reference to the Conceptual Graph
Semantic Interpreter
120
Conceptual Graph reference
Graph operations
Conceptual Graph Processor
209
Store
Retrieve
Knowledge Base
211

Figure 1

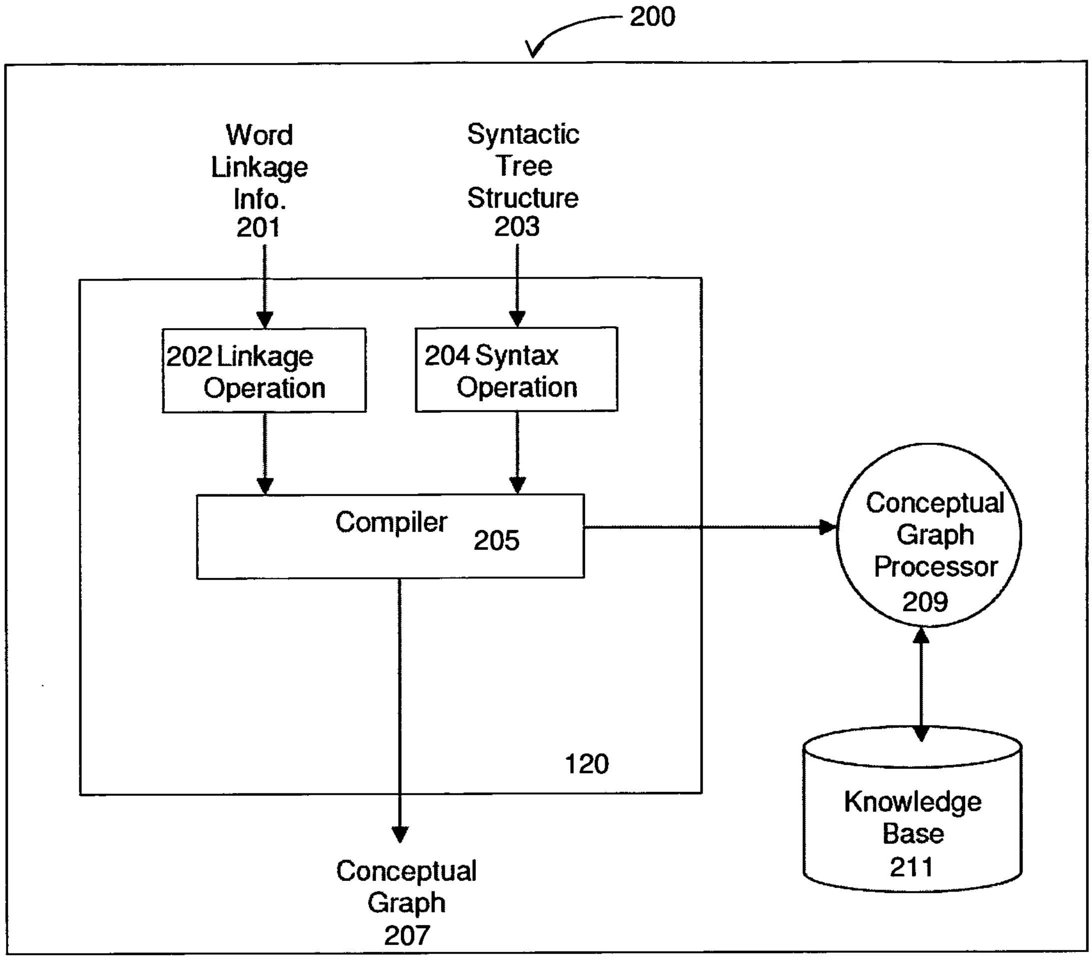
200
Word Linkage Info. 201
Syntactic Tree Structure 203
202 Linkage Operation
204 Syntax Operation
Compiler 205
Conceptual Graph Processor 209
Knowledge Base 211
120
Conceptual Graph 207

Figure 2