



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

(21) International Application Number:
PCT/MY2014/000154

(22) International Filing Date:
29 May 2014 (29.05.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PI 2013004280 27 November 2013 (27.11.2013) MY

(71) Applicant: **MIMOS BERHAD** [MY/MY]; Technology Park Malaysia, 57000 Kuala Lumpur (MY).

(72) Inventors: **Benjamin Chu Min Xian**; c/o MIMOS, Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **Liu Qiang**; c/o MIMOS, Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **Dickson Lukose**; c/o MIMOS, Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY).

(74) Agent: **YAP, Kah Hong**; Pyprus Sdn Bhd, Suit 8-02, 8th Floor, Plaza First Nationwide 161, Jalan Tun H.S. Lee 50000, 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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *with declaration under Article 17(2)(a); without abstract; title not checked by the International Searching Authority*

(54) Title: A METHOD AND SYSTEM FOR AUTOMATED WORD SENSE DISAMBIGUATION

(57) Abstract:



WO 2015/080559 A2

A Method and System for Automated Word Sense Disambiguation

Field of the Invention

[0001] The present invention relates to information processing. More specifically, the present invention relates to a system and method for automated word
5 sense disambiguation.

Background

[0002] Word Sense Disambiguation (WSD) is known to be challenging subject in the field of Natural Language Processing (NLP). The challenge arises from the lack of means to address properties of context that characterize the use of words in a given
10 sense. Further, there is also lack a standard and exhaustive inventory for word sense. It is also noted the accuracy of the current means to process disambiguation results at the final end is often questionable.

[0003] Therefore, given any text to identify the correct sense for an ambiguous word in a sentence; the main problem considered here is how to identify which sense of
15 a meaning is used in any given sentence, when the word has a number of distinct senses. When it happens, it would require substantial amounts of training examples/tagged datasets, i.e. supervised machine learning, to handle it. For example, the word “plant” may have various meaning in different contexts when it is searched through the WordNet. It is however not that these training examples and tagged datasets are still
20 not sufficient for extracting the word sense. Further, the accuracy for fine-grained sense distinctions is still rather lacking with the existing systems. Thus far, for example, the highest accuracies based on the state-of-the art approaches range from about 59.1% to

69.0%. The challenge lies on the determinations of the context length and context content. The context length refers to the size of window of text that should be taken to determine context. It can be difficult, if not impossible, to determine if the context length should contain only a few words, or a larger portions of the string. Similarly, it is also a challenge to decide whether all context words or only a selected word, such as words in certain part of speech or a certain grammatical relations to the target word, are to be considered for context content. There is also question whether the selected the selected words should be weighted based on their distance apart from the target word, or be treated as a “bag of words”.

10

Summary

[0004] In accordance with one aspect of the present invention, there is a system for disambiguating word sense from a text containing document having sentences. The system comprises an entity recognition module adapted for extracting possible entities from the sentence using a Linked Data; a text preprocessor adapted for tokenizing sentence into lemmatized words, the text processor includes a word recognizer adapted to identify if a verb and nouns from the sentence, a lemmatizer for lemmatizing the words of the sentence, and a polysemy checker for counting a number of possible sense of the words to determine if the words are ambiguous, an index builder (134) adapted for creating an index of schema graphs for each identified verb and to extract all possible sense description for nouns; a disambiguator adapted for disambiguating word senses, wherein the disambiguator extracts all the schemas for the identified verb and placing all the identified nouns into the schemas to determine the most suitable word sense, and a disambiguation rules is utilized for disambiguating word sense.

[0005] In one embodiment, the disambiguator is operable to determine if the entity is a verb by referring to a linguistic resource and subsequently retrieve all possible schemas related to the verb.

[0006] In another embodiment, the text processor is operable to determine if a word is a verb or nouns through a linguistic resource. The index builder may be adapted for extracting schemas through the use of a linguistic resources and building an index reference for each word entry with all the related sense descriptions.

[0007] In yet another embodiment, the disambiguator may create a context word vector from the nouns extracted from the sense description, wherein the words in the context word vector are checked for semantic constraints with reference to a concept hierarchy from the linguistic resources, and the disambiguation rules is utilised when the ambiguous word cannot be resolved.

[0008] In another aspect, the present invention has a method of disambiguating word sense from a text containing document having sentences. The method comprises extracting possible entities from the sentence using a Linked Data; tokenizing sentence into lemmatized words; lemmatizing the words of the sentence; counting a number of possible sense of the words to determine if the words are ambiguous through a polysemy checker; identifying verb and nouns from the sentence; lemmatizing the words of the sentence; counting a number of possible sense of the words to determine if the words are ambiguous through a polysemy checker; creating an index of schema graphs for each identified verb and to extract all possible sense description for nouns; disambiguating word senses through extracting all the schemas for the identified verb

and placing all the identified nouns into the schemas to determine the most suitable word sense; utilizing disambiguation rules disambiguating word sense.

[0009] In one embodiment, the disambiguating word sense includes determining if the entity is a verb through referring to a linguistic resource and retrieving all possible
5 schemas related to the verb.

[0010] In another embodiment, the identifying verb and nouns includes matching verb and nouns with a linguistic resource.

[0011] Further, the index builder may be adapted for extracting schemas through the use of a linguistic resources and building an index reference for each word entry
10 with all the related sense descriptions.

[0012] Yet, disambiguating the word sense includes creating a context word vector from the nouns extracted from the sense description, wherein the words in the context word vector are checked for semantic constraints with reference to a concept hierarchy from the linguistic resources, and the disambiguation rules is utilised when
15 the ambiguous word cannot be resolved.

Brief Description of the Drawings

[0013] Preferred embodiments according to the present invention will now be described with reference to the figures accompanied herein, in which like reference
20 numerals denote like elements;

[0014] FIG. 1 illustrates a block diagram of a word sense disambiguation system in accordance with one embodiment of the present invention;

[0015] FIG. 2 illustrates a process carried out by the disambiguation module of FIG. 1 in accordance with one embodiment of the present invention; and

[0016] FIGS. 3A-D exemplify an example of a sentence that is being processed to resolve ambiguity.

5

Detailed Description

[0017] Embodiments of the present invention shall now be described in detail, with reference to the attached drawings. It is to be understood that no limitation of the scope of the invention is thereby intended, such alterations and further modifications in the illustrated device, and such further applications of the principles of the invention as
10 illustrated therein being contemplated as would normally occur to one skilled in the art to which the invention relates.

[0018] FIG. 1 illustrates a block diagram of a word sense disambiguation system 100 in accordance with one embodiment of the present invention. The system 100 is adapted for automatically identifying which sense of a word (i.e. meaning) is
15 used in the sentence context. It is particularly useful for words that are polysemous or have multiple meanings. Briefly, the system comprises an Entity Recognition Module 102, a text pre-processor 102, and a disambiguation module 103.

[0019] The entity recognition module 101 provides a preprocessing to the sentence or a target string to be processed to identify the possible entities based on a
20 link data 112. Any entity recognition engine that is known in the market is suitable for identifying the relevant entities.

[0020] The entity recognition module **101** is configured to recognize entities from content of a document. Many systems and methods for recognizing entities are well known in the art and they can be adapted for the present invention. In another embodiment, the entity recognition engine or module disclosed in the Malaysia patent application entitled "SYSTEM AND METHOD FOR AUTOMATED ENTITY RECOGNITION" filed on the same day as the present application can also be adapted wherein.

[0021] The text preprocessor **102** comprises a word recognizer **122**, a Lemmatizer **124**, and a Polysemy Checker **126**. The word recognizer **122** is adapted to works with the entity recognition module **101** to distinguish from the sentence if a word is a noun or a verb. The word recognizer **122** also takes references from the linguistic resources **138** to perform its recognitions. The lemmatizer **124** is adapted to tokenize sentence into lemmatized word form to identify ambiguous words. The polysemy checker **126** is utilized to identify which ambiguous word is to be disambiguated.

[0022] An index builder **134** is used to create an index of schema graphs/maps for each verb. It extracts all possible sense descriptions from the linguistic resources **138** to determine if the word is a noun, while extracting possible related schemas to the targeted verb. The index is referenced by the disambiguator for disambiguating the word sense.

[0023] It can be seen that the use of the polysemy checker **126** to identify ambiguous word, and the word recognizer **122** to distinguish verb and noun from the sentence and the index builder to extract sense description (for noun) and schemas (for

verb) would be able to address the need for substantial amounts of training examples/tagged datasets.

[0024] The disambiguation module 103 receives the word and disambiguates its word sense based on disambiguate rules 132. With the use of semantic information (i.e. schemas), and the context vectors created from the sense descriptions to determine for selectional constraint for the semantic role, a high accuracy for fine-grained sense distinctions can be achieved. The disambiguation rules 132 are well known in the art.

[0025] FIG. 2 illustrates a process carries out by the disambiguation module 103 of FIG. 1 in accordance with one embodiment of the present invention. The process starts with selecting a sentence or target sentence to be processed to extract word sense of the containing words at step 202. From the target sentence, each of the words is being determined if it is a verb at step 204 through the use of the linguistic resources 138. This is done through the text preprocessor 102. When a word is determined to be a verb at step 206, all possible schemas related to the target verb are retrieved at step 208. A word that is not identified as verb, in general, it would be a noun for being salient and meaningful, which otherwise, it will proceed under step 212. At the step 212, it is further being determined if that word is ambiguous by using the polysemy checker. The polysemy checker 126 calculates a total count of different possible senses for the word. When the word has more than one possible sense, the higher count on the polysemy, the higher in likelihood that the word is ambiguous. If the word is being determined to be ambiguous at step 214, all related sense descriptions for the word (potential ambiguous word) are retrieved through the index rendered by the index builder 134 at step 216 and subsequently all nouns are extracted from the sense description to create context vectors of the word at step 218.

[0026] Returning to the step 214, if the word is determined to be not ambiguous, at step 222, the word is being matched with the schemas of the word retrieved in step 208. A best schema, being the maximum number of concept matched, is selected at step 224.

5 [0027] At step 226, each of the context vectors of the word is checked if it satisfies selectional constraints of the semantic role of the best schema that identified earlier. The selectional constraints check is done with the reference to a concept hierarchy from the linguistic resources 138 at step 228.

[0028] When the selectional constraints above are satisfied, at step 232, a best
10 sense to the ambiguous word is selected and assigned to that word. If this can be resolved in step 234, the sense of that word is identified. If the sense of that word cannot be resolved, i.e. selectional constraints check not satisfied, at step 234, the disambiguation module 103 applies disambiguation rules to give the word a word sense.

[0029] FIGs. 3A-3C exemplify an example of a sentence that is being processed
15 to disambiguate the word sense thereof. The exemplified sentence is "*The boy fishes the bass from the river.*". The exemplified sentence can also herewith refer as a target sentence. As shown in FIG. 3A, the target sentence is scanned through by the present system 100 to identify entity/noun phrase through the entity recognition module 101 with reference to the Linked Data 112. In this case, the word "boy", "bass" and "river"
20 shall be identified. Subsequently, verb(s) are identified from the target sentence using the linguistic resources 138. In this case, "fish" may be identified. The word "fish" will be tokenized into lemmatized form.

[0030] Accordingly, as shown in **FIG. 3B**, all the schemas relating to “fish” are extracted as shown as S1, S2, ..., Sn. The ambiguous words are being identified through polysemy checker 126. In this case, “fish” and “bass” are polysemous, and since “fish” is identified as verb, it can be remove from the consideration to
5 disambiguate. Subsequently putting the words from the sentence to all the schemas extracted before. In this case, S2 can be identified as the closest match for the sentence.

[0031] Referring now to **FIG. 3C**, all the possible sense descriptions are retrieved through the index builder 134. Following that, all the nouns from the sense description is extracted to create a context vector.

10 [0032] As shown in **FIG. 3D**, each of the context word is checked if the semantic constraints are satisfied. In the illustrated sentence, you can identify that the context word “fish” is an animate from the concept hierarchy. And therefore, the word “bass” shall be assigned with an appropriate sense from the sense description that has the word “fish” as shown in **FIG. 3C**.

15 [0033] While specific embodiments have been described and illustrated, it is understood that many changes, modifications, variations, and combinations thereof could be made to the present invention without departing from the scope of the invention.

Claims

1. A system for disambiguating word sense from a text containing document having sentences, the system comprising:

5 an entity recognition module (101) adapted for extracting possible entities from the sentence using a Linked Data (112);

a text preprocessor (102) adapted for tokenizing sentence into lemmatized words, the text processor (102) includes a word recognizer (122) adapted to identify if a verb and nouns from the sentence, a lemmatizer (124) for lemmatizing the words of the
10 sentence, and a polysemy checker (126) for counting a number of possible sense of the words to determine if the words are ambiguous,

an index builder (134) adapted for creating an index of schema graphs for each identified verb and to extract all possible sense description for nouns;

a disambiguator (103) adapted for disambiguating word senses, wherein the
15 disambiguator (103) extracts all the schemas for the identified verb and placing all the identified nouns into the schemas to determine the most suitable word sense, and a disambiguation rules (132) is utilized for disambiguating word sense.

2. The system according to claim 1, wherein the disambiguator (103) is operable to
20 determine if the entity is a verb through referring to a linguistic resource (138) and retrieve all possible schemas related to the verb.

3. The system according to claim 1, wherein the text processor (102) is operable to determine if a word is a verb or nouns through a linguistic resource (138).
4. The system according to claim 2, wherein the index builder (134) adapted for
5 extracting schemas through the use of a linguistic resources (138) and building an index reference for each word entry with all the related sense descriptions.
5. The system according to claim 1, wherein the disambiguator (103) creates a context word vector from the nouns extracted from the sense description, wherein the
10 words in the context word vector are based a concept hierarchy from the linguistic resources (138), and the disambiguation rules (132) is utilised when the ambiguous word cannot be resolved.
6. A method of disambiguating word sense from a text containing document
15 having sentences, the method comprising:
extracting possible entities from the sentence using a Linked Data (112);
tokenizing sentence into lemmatized words
lemmatizing the words of the sentence
counting a number of possible sense of the words to determine if the words are
20 ambiguous through a polysemy checker;
identifying verb and nouns from the sentence,
lemmatizing the words of the sentence,

counting a number of possible sense of the words to determine if the words are ambiguous through a polysemy checker (126);

creating an index of schema graphs for each identified verb and to extract all possible sense description for nouns;

- 5 disambiguating word senses through extracting all the schemas for the identified verb and placing all the identified nouns into the schemas to determine the most suitable word sense;

utilizing disambiguation rules (132) disambiguating word sense.

- 10 7. The method according to claim 6, wherein disambiguating word sense includes determining if the entity is a verb through referring to a linguistic resource (138) and retrieving all possible schemas related to the verb.

8. The method according to claim 6, wherein identifying verb and nouns includes
15 matching verb and nouns with a linguistic resource (138).

9. The method according to claim 6, wherein the index builder (138) adapted for extracting schemas through the use of a linguistic resources and building an index reference for each word entry with all the related sense descriptions.

- 20 10. The method according to claim 6, wherein disambiguating the word sense includes creating a context word vector from the nouns extracted from the sense description, wherein the words in the context word vector are based on a concept

hierarchy from the linguistic resources (138), and the disambiguation rules (132) is utilised when the ambiguous word cannot be resolved.

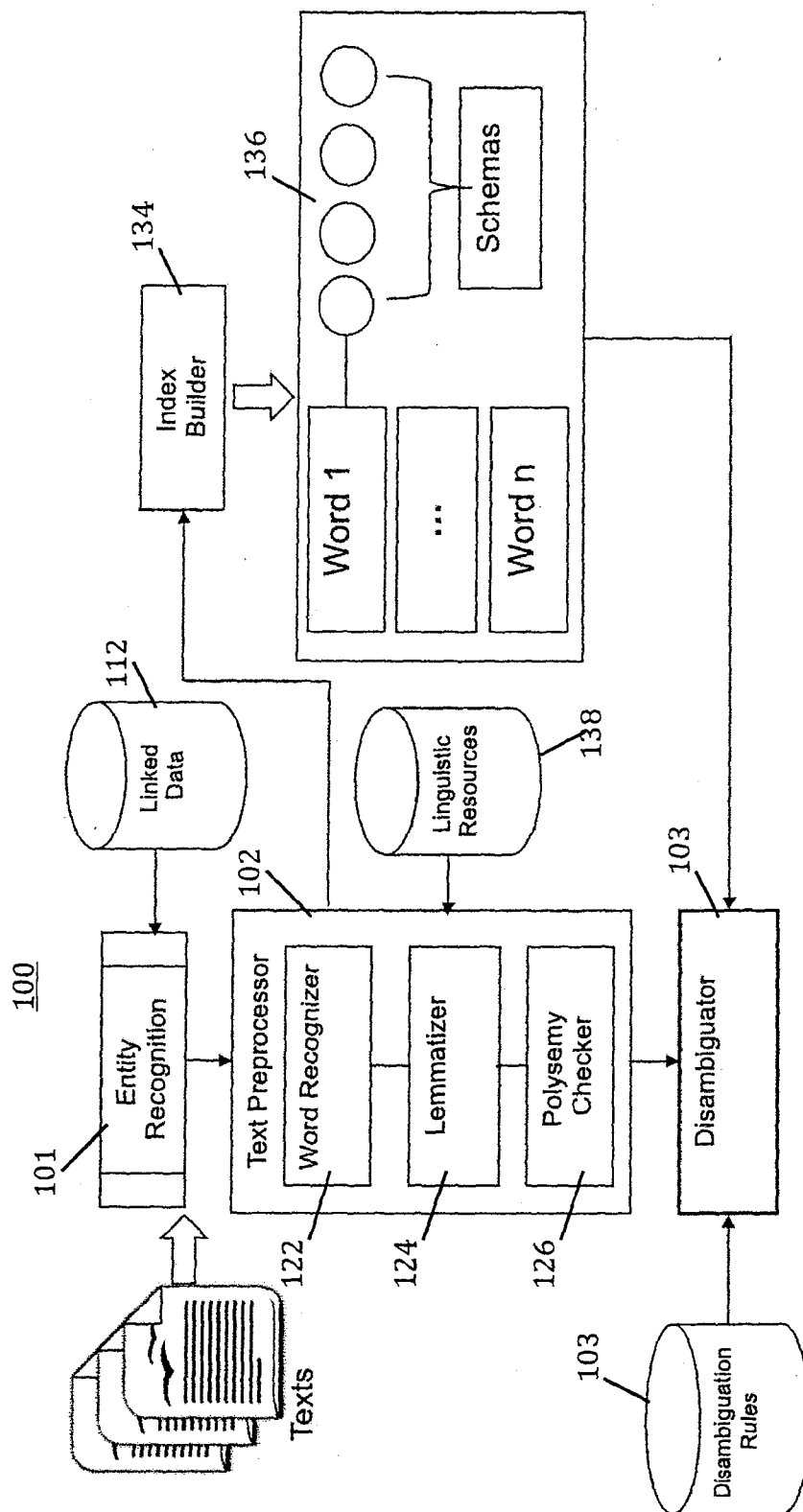


FIG. 1

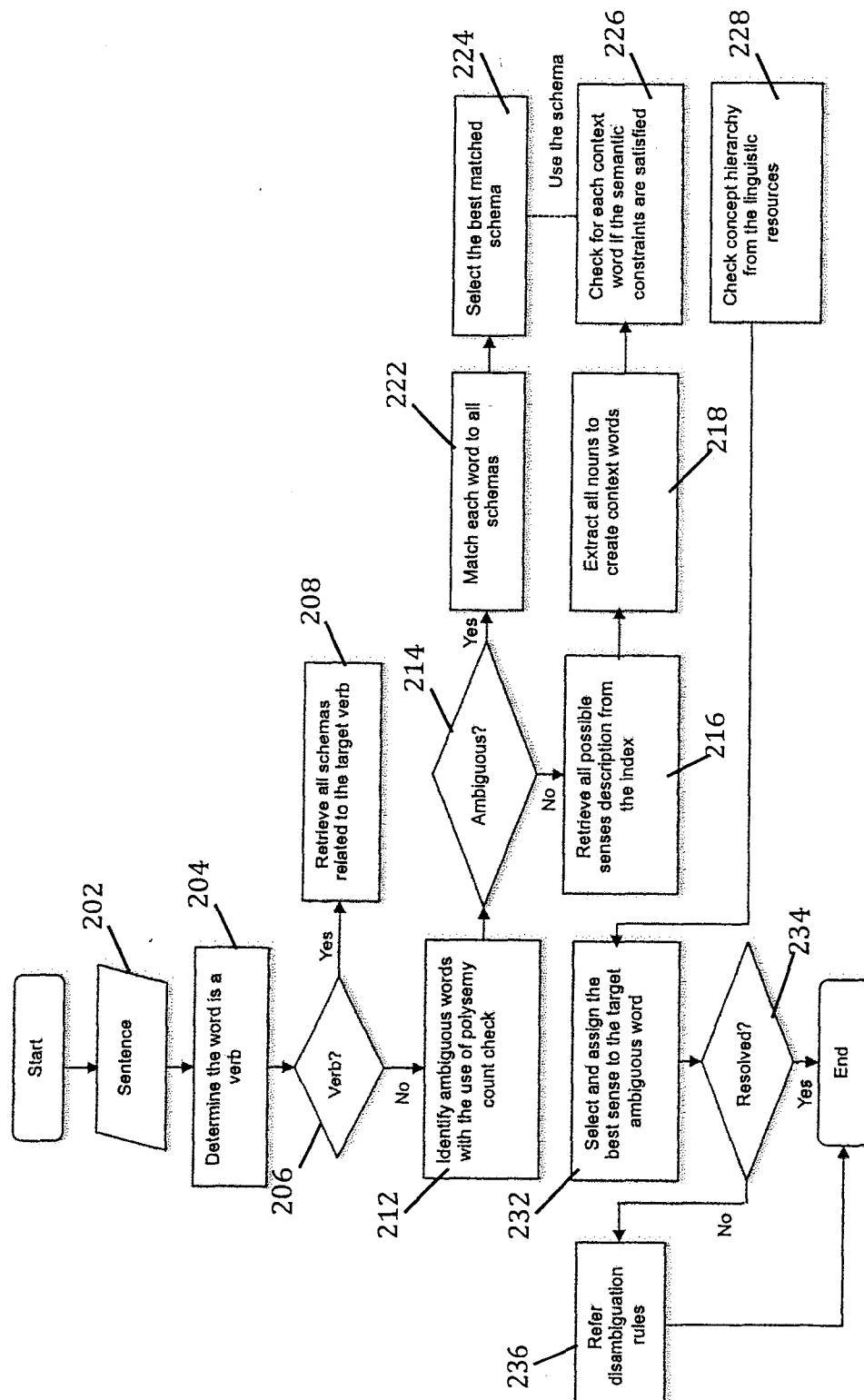


FIG. 2

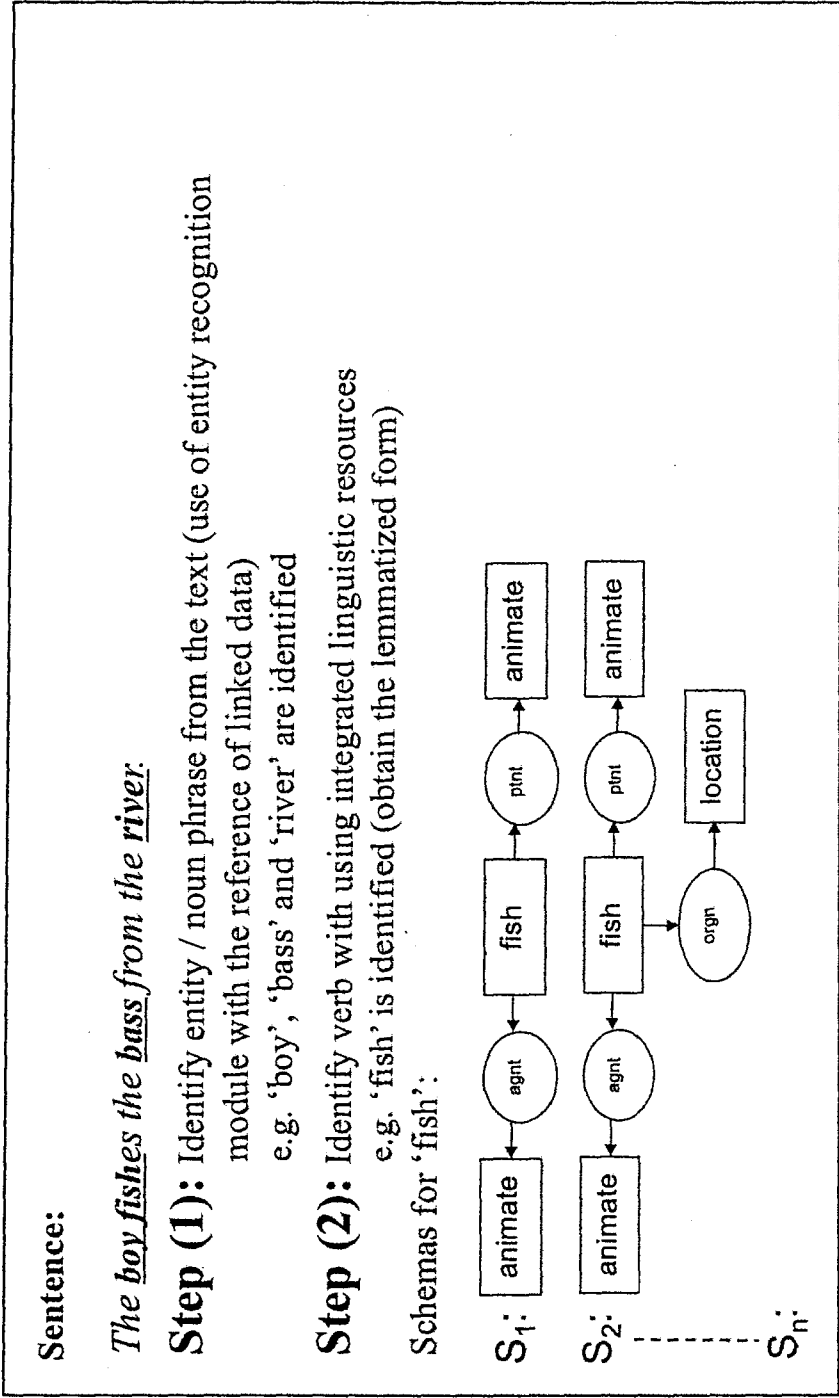


FIG. 3A

Sentence:

The boy fishes the bass from the river.

Step (3): Retrieve all possible schemas related to the target verb

Step (4): Identify ambiguous word with the use of polysemy count check
(For this case, we can identify 'fish' and 'bass' are polysemous; and since 'fish' is identified as verb, remove from consideration to disambiguate)

Step (5): Match each of the words from the sentence to all schemas extracted
(For this case, we can identify Schema 2 is the most closely match for the sentence)

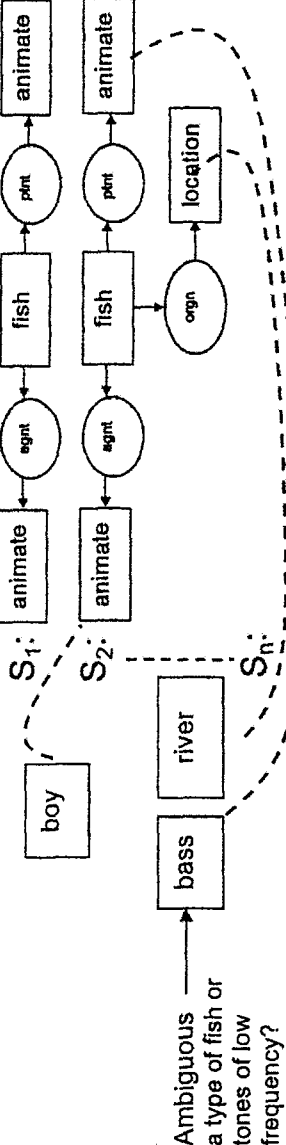


FIG. 3B

Sentence:

The boy fish the bass from the river.

Step (6): Retrieve all possible sense description from the index

Sense	Description
Bass #1	The lowest <u>part</u> of the musical <u>range</u>
Bass #2	The lowest <u>part</u> in polyphonic <u>music</u>
Bass #3	An <u>adult</u> <u>male</u> <u>singer</u> with the lowest <u>voice</u>
Bass #4	The lean flesh of a saltwater <u>fish</u> of the family Serranidae

Step (7): Extract all nouns from the sense description to create a context vector

FIG. 3C

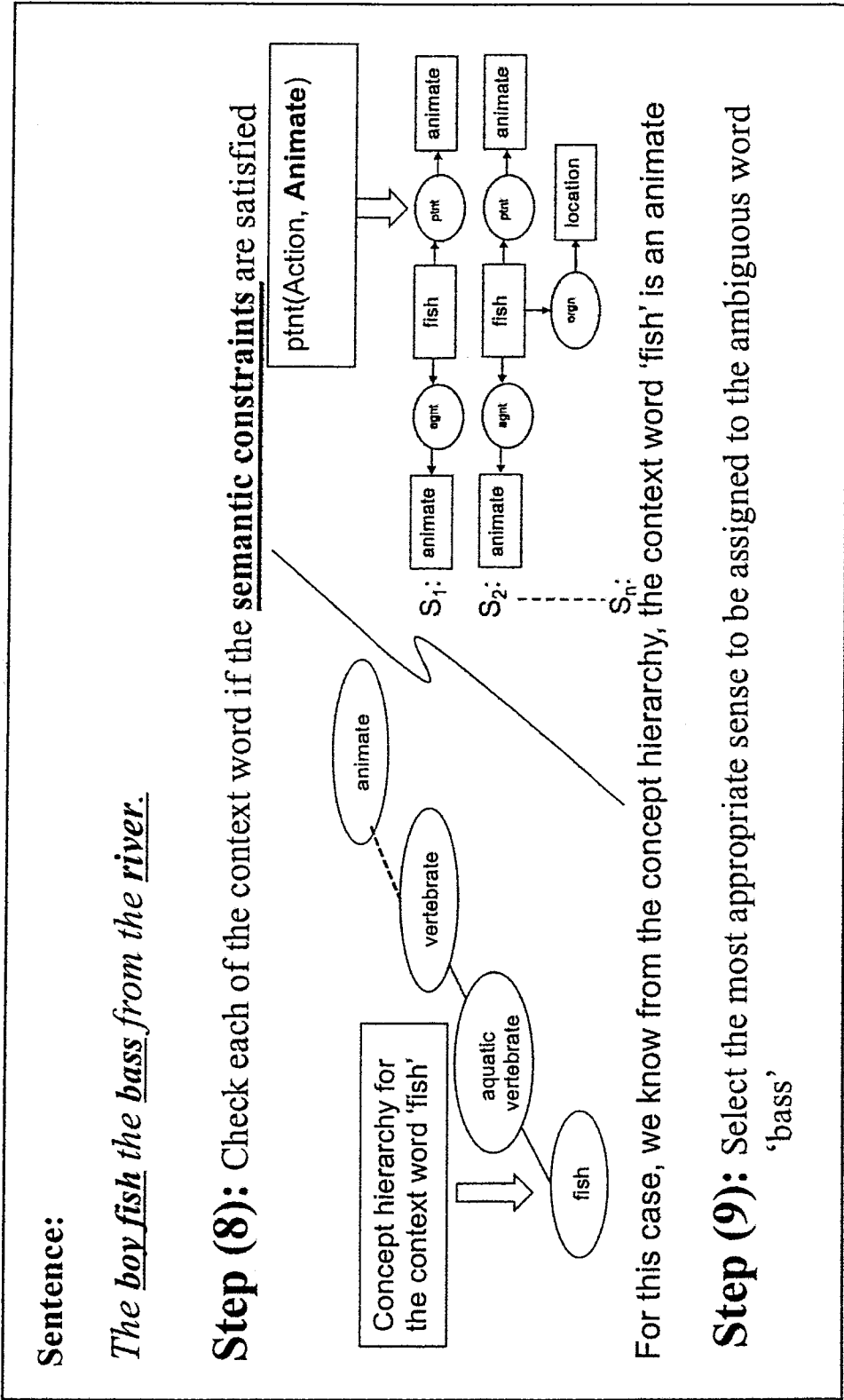


FIG. 3D

PATENT COOPERATION TREATY

PCT

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a), Rules 13ter.1(c) and Rule 39)

Applicant's or agent's file reference 10146.P061PCT	IMPORTANT DECLARATION	Date of mailing(<i>day/month/year</i>) 18 February 2015 (18-02-2015)
International application No. PCT/MY2014/000154	International filing date(<i>day/month/year</i>) 29 May 2014 (29-05-2014)	(Earliest) Priority date(<i>day/month/year</i>) 27 November 2013 (27-11-2013)
International Patent Classification (IPC) or both national classification and IPC G06F17/2785		
Applicant MIMOS BERHAD		

This International Searching Authority hereby declares, according to Article 17(2)(a), that **no international search report will be established** on the international application for the reasons indicated below

1. ☐ The subject matter of the international application relates to:
 - a. ☐ scientific theories
 - b. ☐ mathematical theories
 - c. ☐ plant varieties
 - d. ☐ animal varieties
 - e. ☐ essentially biological processes for the production of plants and animals, other than microbiological processes and the products of such processes
 - f. ☐ schemes, rules or methods of doing business
 - g. ☐ schemes, rules or methods of performing purely mental acts
 - h. ☐ schemes, rules or methods of playing games
 - i. ☐ methods for treatment of the human body by surgery or therapy
 - j. ☐ methods for treatment of the animal body by surgery or therapy
 - k. ☐ diagnostic methods practised on the human or animal body
 - l. ☐ mere presentations of information
 - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art
2. ☒ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:

☒ the description
☒ the claims
☐ the drawings
3. ☐ A meaningful search could not be carried out without the sequence listing; the applicant did not, within the prescribed time limit:

☐ furnish a sequence listing on paper complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.
☐ furnish a sequence listing in electronic form complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it.
☐ pay the required late furnishing fee for the furnishing of a sequence listing in response to an invitation under Rule 13ter.1(a) or (b).
4. Further comments:

Name and mailing address of the International Searching Authority  European Patent Office, P.B. 5818 Patentlaan 2 NL-2280 HV Rijswijk Tel. (+31-70) 340-2040 Fax: (+31-70) 340-3016	Authorized officer GOLZE, Doreen Tel: +49 (0)30 25901-704
---	---

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 203

1. The system for disambiguating word sense from a text of claim 1 (and the corresponding method of claim 6) relies on several component modules, which are not clearly defined (Article 6 PCT), (neither individually nor w.r.t. to their interaction). Moreover, the application as a whole does not disclose the corresponding subject matter in a manner sufficiently clear and complete for the invention to be carried out by a person skilled in the art (Article 5 PCT).

The lack of clarity (Article 6 PCT) and disclosure (Article 5 PCT) leave it impossible to conduct a meaningful search into the prior art, for the reasons set out below.

2.1 According to Article 6 PCT, claims shall be clear, concise and fully supported by the description. When evaluating the entity recognition module of claim 1 (and corresponding claim 6), the person skilled in the art thus looks to the description, according to which (see paragraph 20), the entity recognition module, is assumed to be of a type known to the person skilled in the art; the only further indication as to what sort of module is meant is a reference to a module disclosed in an unpublished patent application by the applicant, which may be used in an embodiment. From such a vague reference to prior art modules, there is no way the person skilled in the art can know what sort of output to expect, beyond the indication that Linked Data is used. Claims 1 and 6 are thus not supported by the description (Article 6 PCT).

Even if the person skilled in the art assumes that the output should be Linked Data, the term is so vague that this does not give the person skilled in the art a sufficient indication as to how the data should subsequently be used. As this information is lacking from the entire application, the application fails to meet the requirements of Article 5 PCT.

As the module is neither properly defined in the claim nor sufficiently supported in the description, claim 1 (and corresponding claim 6) fails to meet the requirements of Article 6 PCT. Moreover, there appears to be no way of overcoming the objection as the module is not sufficiently disclosed in the description either, a mere reference to prior art recognisers not being sufficient disclosure such that the person skilled in the art could construct the specific entity recognition module which would be required for this application. Therefore, the requirements of Article 5 PCT are not met either, making a search of such a feature impossible.

2.2.1 The system of claim 1 (and the corresponding method of claim 6) uses an index builder updated for creating an index of schema for each identified verb. Neither the claim itself, nor the remainder of the application provides the person skilled in the art with the features required to create such an index. There is no indication as to where the schemas should come from, what form they should take, how such an index should be built and how it should look.

2.2.2 The feature continues by stating that the index builder should also extract all possible sense descriptions from nouns, a vague mention of WordNet notwithstanding, there is no indication as to where the sense descriptions should be extracted from and what role the index builder should play in doing so.

2.2.3 The application thus fails to meet the requirements of the PCT because the index builder is neither clearly defined (Article 6 PCT), nor disclosed in a manner sufficiently clear and complete for the invention to be carried out by a person skilled in the art (Article 5 PCT).

2.3 A disambiguator is adapted to extract all the schemas for the

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 203

identified word. This is rendered impossible for the person skilled in the art because he or she has no indication as to what the schemas are. Furthermore, there is no indication as to how the module interacts with the remaining modules of claim 1 (and corresponding claim 6). For these reasons too, the application fails to meet the requirements of Articles 5 and 6 PCT.

3. None of the three main modules of independent claims is disclosed clearly or supported by the description (Article 6 PCT). Moreover, the invention has not been disclosed in the description in such a way that the person skilled in the art could implement it (Article 5 PCT). It is thus impossible to carry out a meaningful search into the prior art w.r.t. claims 1 and 6.

4. As dependent claims 2-5 and 7-10 depend on claims 1 and 6 respectively, but do not themselves disclose additional features which may make their subject matter clear (Article 6 PCT) no search can be carried out w.r.t. claims 2-5 and 7-10 either.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure. If the application proceeds into the regional phase before the EPO, the applicant is reminded that a search may be carried out during examination before the EPO (see EPO Guidelines C-IV, 7.2), should the problems which led to the Article 17(2) declaration be overcome.