



- (51) International Patent Classification:
G09B 7/08 (2006.01)
- (21) International Application Number:
PCT/MY2014/000126
- (22) International Filing Date:
28 May 2014 (28.05.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
PI 2013701825 27 September 2013 (27.09.2013) MY
- (71) Applicant: **MIMOS BERHAD** [MY/MY]; Technology Park Malaysia, 57000 Kuala Lumpur (MY).
- (72) Inventors: **JASBEER SINGH ATMA, Singh**; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **CHU IN XIAN, Benjamin**; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **DICKSON, Lukose**; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY). **LIU, Qiang**; c/o Mimos Berhad, Technology Park Malaysia, 57000 Kuala Lumpur (MY).
- (74) Agent: **H. A. RASHID, Ahmad Fadzlee**; Ipvolusi Sdn Bhd, A-3-3A, Centrio Pantai Hillpark, No. 1, Jalan Pantai Murni, 59200 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:

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- of inventorship (Rule 4.17(iv))

[Continued on next page]

(54) Title: A SYSTEM AND METHOD FOR CREATING LEARNING OBJECTS FROM A DOCUMENT

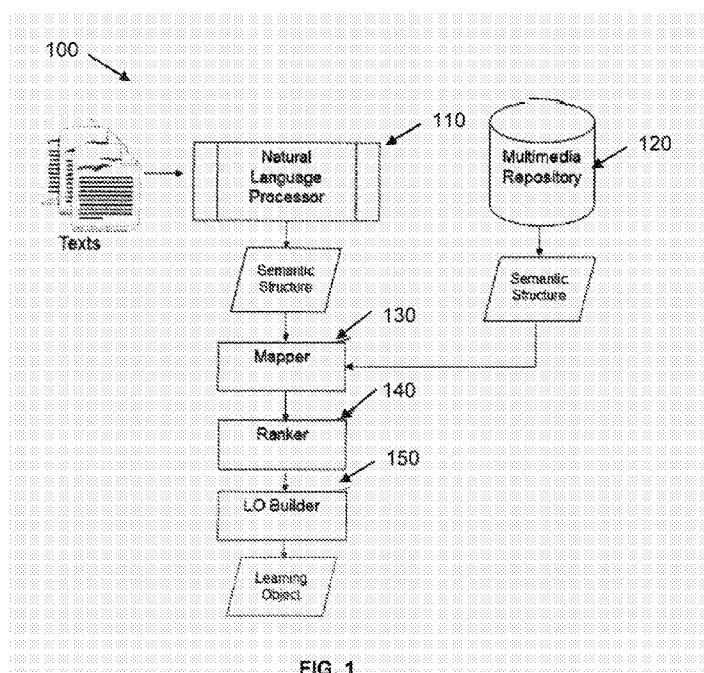


FIG. 1

(57) Abstract: The present invention relates to a system and method for creating a Learning Object (LO) from a document. Particularly, the system (100) automatically generates the LO at sentence level, paragraph level and topic level in different modalities such as text, image, audio and video. The system (100) comprises a mapper (130), a ranker (140) and a Learning Object Builder (150). The method for creating a Learning Object from a document includes the steps of analysing and processing a document received from a user, matching a retrieved semantic structure by sentence index with a semantic structure from a multimedia repository (120) by the mapper (130), selecting and ranking projected graphs for each type of multimedia by the ranker (140) and producing the Sentence Learning Object, Paragraph Learning Object and Topic Learning Object by the Learning Object Builder (150).



Published:

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

**A SYSTEM AND METHOD FOR CREATING LEARNING OBJECTS FROM A
DOCUMENT**

FIELD OF INVENTION

5 The present invention relates to a system and method for creating Learning Object (LO) from a document. Particularly, the system automatically generates the LO at sentence level, paragraph level and topic level in different modality such as text, image, audio and video.

10 **BACKGROUND OF THE INVENTION**

 Learning Object (LO) is an entity that is used in learning, education or training system. The main idea of Learning Object is to break educational content down into small chunks that can be reused in various learning environments, object-oriented programming. In other words, learning object is a representation designed to afford
15 uses in different educational contexts. There are a few examples of system and method that use the concept of Learning Object for education purpose.

 An example of such system and method is disclosed in US Publication No. 2008/0189684 A1 whereby the system and method of an e-learning authorship is
20 based on meta-tagged media specific learning objects. In one embodiment, a method includes searching for learning objects in a repository based on a content query data publishing the learning objects to a content management system through applying a meta-data describing a context to the learning objects to such that the learning objects become readily accessible based on a search of the context and generating a
25 set of executable files based on the learning objects and a number of auxiliary files associated with the set of executable files to create an e-learning course of a specific media type. In addition, the method may include comparing a context query data and the meta-data of the each of the learning objects to perform a context search which is used to select a subset of the learning object matching the context query data.
30 However, the prior art does not generate learning object at a sentence level, a paragraph level and a topic level.

 Another example is disclosed in US Publication No. 2004/0205578 A1 whereby a system and method converts a document to create a reusable learning
35 object. The reusable learning object is used to create a course of instruction for an

online or virtual classroom. The reusable learning object is shared and modified by different users for each course. A document stored on a computing platform coupled to the system is made available for the reusable learning object. The document is created by an application and contains data that may not be in a format useable by the reusable learning object. An autopilot and an export filter are invoked by a prompt or command from the application or another source. The export filter converts the data within the document into metadata to support an extensible mark-up language ("XML") format. The autopilot queries and receives information on the document to create the reusable learning object. The reusable learning object is created by the converted data and received information. Thus, the document is available for online training purposes despite not being in a format supported by the reusable learning object. However, the prior art generates learning object using a metadata and information wherein the metadata is converted from a data using an export filter which is manually defined by a user.

15

Hence, there is a need to improve upon the existing system and method for automated creation of Learning Object (LO) from a document. The improvement shall expand the scope of Learning Object for education purpose.

20 SUMMARY OF INVENTION

A system (100) for creating Learning Object (LO) from a document comprises of a natural language processor (110) to analyse and process a document received from a user, a multimedia repository (120) to store multimedia items in different modality, and a learning object repository (160) to store Learning Object; wherein the system (100) is characterised in that the Learning Object Creation system (100) includes a mapper (130) to perform semantic structure matching between semantic structure of the multimedia items and semantic structure of the text of the document; a ranker (140) to select the matched semantic structures, project graphs and calculate the average score of matched semantic structures; and a Learning Object Builder (150) to create a Paragraph Learning Object, a Sentence Learning Object and a Topic Learning Object.

A method for creating Learning Object (LO) from a document by using the system comprises the steps of analysing and processing a document; matching a retrieved semantic structure by sentence index with a semantic structure from the

35

multimedia repository (120) by a mapper (130); selecting and ranking projected graphs for each type of multimedia by a ranker (140); and producing the Sentence Learning Object, Paragraph Learning Object and Topic Learning Object by a Learning Object Builder (150).

5

Preferably, matching the retrieved semantic structure by sentence index with a semantic structure from the multimedia repository (120) by the mapper (130) includes the steps of retrieving the semantic structure by sentence index by the mapper (130); selecting the multimedia type and the semantic structure from the multimedia repository (120) by the mapper (130); performing a matching of both semantic structures selected from the multimedia repository (120) by the mapper (130) by using a projection technique; storing the results in a Learning Object Repository (160) if the semantic structures match each other; and determining whether there are more sentence indexes to be selected and completed by the mapper (130).

15

Preferably, selecting and ranking projected graphs for each type of multimedia by the ranker (140) includes the steps of selecting the projected graphs for each type of multimedia by the ranker (140); ranking the projected graphs for each type of multimedia by the ranker (140); selecting the top ranking graph by the ranker (140); selecting the matched semantic structures by the ranker (140); and performing an average scoring by the ranker (140) by using different ranking methods.

20

Preferably, producing the Sentence Learning Object, Paragraph Learning Object and Topic Learning Object by the Learning Object Builder (150) includes the steps of aggregating the Sentence Learning Object by the system (100); creating the Paragraph Learning Object from the output of aggregation; and aggregating the Paragraph Learning Object to create the Topic Learning Object.

25

Preferably, aggregating the Sentence Learning Object by the system (100) includes the steps of selecting the sentence index for aggregation; selecting the top ranked projected graph for each of the multimedia type for the respective sentence index; arranging the projected graphs according to the order of the text, video, audio and image; producing the Sentence Learning Object from the arrangement; and

30

storing the produced Sentence Learning Object in the Learning Object repository (160).

5 Preferably, creating the Paragraph Learning Object from the output of aggregation includes the steps of selecting the paragraph index for aggregation; selecting and grouping related projected graph from each Sentence Learning Object into the type of multimedia accordingly; sorting the projected paragraphs by sentence order for each multimedia type; producing the Paragraph Learning Object from the sortation; and storing the produced Paragraph Learning Object in the Learning
10 Object repository (160).

Preferably, aggregating the Paragraph Learning Object to create the Topic Learning Object includes the steps of selecting the topic index for aggregation; selecting and grouping related projected graph from each Paragraph Learning Object
15 into the type of multimedia accordingly; sorting the projected paragraphs by paragraph order for each multimedia; producing the Topic Learning Object from the sortation; and storing the produced Topic Learning Object in the Learning Object repository (160).

20 Preferably, ranking methods include a Conceptual Similarity, a Relational Similarity, a Conceptual and Relational Similarity, a Domain Similarity and/or a combination thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

25 The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments of the invention and, together with the description, serve to explain the principles of the invention.

30 FIG. 1 illustrates a block diagram of a system (100) for creating Learning Object (LO) from a document according to an embodiment of the present invention.

FIG. 2 illustrates a flowchart of a method for creating Learning Object (LO) from a document according to an embodiment of the present invention.

FIG. 3 illustrates a flowchart of sub-steps for matching a retrieved semantic structure by sentence index with a semantic structure of the method of **FIG. 2**.

FIG. 4 illustrates a flowchart of sub-steps for selecting and ranking projected graphs for each type of multimedia of the method of **FIG. 2**.

FIG. 5 illustrates a flowchart of sub-steps for producing a Sentence Learning Object, a Paragraph Learning Object and a Topic Learning Object of the method of **FIG. 2**.

FIG. 6 illustrates a flowchart of sub-steps for aggregating Sentence Learning Object of the Learning Object Builder (150) of the method of **FIG. 5**.

FIG. 7 illustrates a flowchart of sub-steps for aggregating Paragraph Learning Object of the Learning Object Builder (150) of the method of **FIG. 5**.

FIG. 8 illustrates a flowchart of sub-steps for aggregating Topic Learning Object of the Learning Object Builder (150) of the method of **FIG. 5**.

DESCRIPTION OF THE PREFERRED EMBODIMENT

A preferred embodiment of the present invention will be described herein below with reference to the accompanying drawings. In the following description, well known functions or constructions are not described in detail since they would obscure the description with unnecessary detail.

FIG. 1 shows a system (100) for creating Learning Object (LO) from a document according to an embodiment of the present invention. The system (100) comprises of a mapper (130), a ranker (140) and a Learning Object Builder (150). The mapper (130) performs semantic structure matching between semantic structure of the multimedia items and semantic structure of the text of the document. The ranker (140) is used to select and project graphs. Moreover, the ranker (140) is used to calculate the average score of matched semantic structures. The Learning Object Builder (150) is used to create a Paragraph Learning Object, a Sentence Learning Object and a Topic Learning Object. The system (100) further comprises of a natural language processor (110) and a multimedia repository (120). The natural language processor (110) analyses and processes a document received from the user. The

natural language processor (110) processes the document into a set of semantic structures which represent the document. The document can be in the form of text, video, audio and image. Each type of document is represented with a respective semantic structure. A multimedia repository (120) comprise of items in different modality such as text, video, audio and image. All the items in the multimedia repository (120) are indexed and semantically structured.

FIG. 2 shows a method for creating Learning Object (LO) from a document according to an embodiment of the present invention. The method includes the steps of analysing and processing a document received from the user as in step 201, matching the semantic structure by sentence index with a semantic structure from the multimedia repository (120) by the mapper (130) as in step 202, selecting and ranking projected graphs for each type of multimedia by the ranker (140) as in step 203 and producing the Sentence Learning Object, Paragraph Learning Object and Topic Learning Object by the Learning Object Builder (150) as in step 204.

FIG. 3 shows a flowchart of sub-steps for matching a retrieved semantic structure by sentence index with a semantic structure of the method of FIG. 2. In step 131, the mapper (130) retrieves the semantic structure by sentence index. In step 132, the mapper (130) selects the multimedia type and thereon, the mapper (130) selects the semantic structure from the multimedia repository (120) as in step 133. Later on, the mapper (130) performs a matching of both semantic structures selected from the multimedia repository (120) by using a projection technique as in step 134. The projection technique includes the matching of the available sentence semantic structure information with one of the multimedia semantic structure, wherein the selection of the multimedia semantic structures is done by iterating all the multimedia items in the multimedia repository (120) based on the sentence index. In decision 135, if the semantic structures do not match each other, the method returns to step 134. If the semantic structures match each other, the matched results are stored in a Learning Object Repository (160). In decision 137, the mapper (130) determines whether there are more sentence indexes to be selected and completed. If there are more sentences indexes to be selected and completed, the method returns to step 131. If there is no more sentence index to be selected and completed, the method ends.

FIG. 4 shows a flowchart of sub-steps for selecting and ranking projected graphs for each type of multimedia of the method of **FIG. 2**. The ranker (140) selects the projected graphs for each type of multimedia, ranks them and selects the top ranking graph. The ranker (140) selects the matched semantic structures and then performs an average scoring by using different ranking methods such as Conceptual Similarity, Relational Similarity, Conceptual and Relational Similarity, Domain Similarity and/or a combination thereof. The Conceptual Similarity describes the similarity between two concepts by calculating the distance between them. The distance between the two concepts is calculated by their respective positions in the concept hierarchy. The Relational Similarity describes the similarity between two concepts wherein it defines the similarity between the two relations in the graph. The distance between the two relations is calculated by their respective positions in the relation hierarchy. The Domain Similarity expresses the similarity of concepts based on the overlapping of both graphs in terms of common concepts and relations to the domain context of the document.

FIG. 5 shows a flowchart of sub-steps for producing a Sentence Learning Object, a Paragraph Learning Object and a Topic Learning Object of the method of **FIG. 2**. In step 300, the Sentence Learning Object is first aggregated by the system (100). From the output of aggregation, the method continues to create the Paragraph Learning Object as in step 400. The Paragraph Learning Object is further aggregated to create the Topic Learning Object as in step 500.

FIG. 6 shows a flowchart of sub-steps for aggregating Sentence Learning Object of the Learning Object Builder (150) of the method of **FIG. 5**. In step 301, the sentence index is selected for aggregation. In step 302, top ranked projected graph is selected for each of the multimedia type for the respective sentence index. Later on, the projected graphs are arranged according to the order of the text, video, audio and image as in step 303. The Sentence Learning Object is produced from the arrangement and then stored in the Learning Object repository (160). In decision 304, if there are more sentence indexes to be selected, the method returns to step 301. If there is no more sentence index to be selected, the method ends.

FIG. 7 shows a flowchart of sub-steps for aggregating Paragraph Learning Object of the Learning Object Builder (150) of the method of **FIG. 5**. In step 401, the

paragraph index is selected for aggregation. In step **402**, all related projected graph from each Sentence Learning Object are selected and grouped into the type of multimedia accordingly. The projected paragraphs are sorted by sentence order for each multimedia type as in step **403**. Later on, the Paragraph Learning Object is produced from the sortation and then stored in the Learning Object repository (**160**).
5 In decision **404**, if there are more paragraph indexes to be selected, the method returns to step **401**. If there is no more paragraph index to be selected, the method ends.

10 **FIG. 8** illustrates shows of sub-steps for aggregating Topic Learning Object of the Learning Object Builder (**150**) of the method of **FIG. 5**. In step **501**, the topic index is selected for aggregation. In step **502**, all related projected graph from each Paragraph Learning Object are selected and grouped into the type of multimedia accordingly. The projected paragraphs are sorted by paragraph order for each
15 multimedia type as in step **503**. Later on, the Topic Learning Object is produced from the sortation and then stored in the Learning Object repository (**160**). In decision **504**, if there are more topic indexes to be selected, the method returns to step **501**. If there is no more topic index to be selected, the method ends.

20 While embodiments of the invention have been illustrated and described, it is not intended that these embodiments illustrate and describe all possible forms of the invention. Rather, the words used in the specifications are words of description rather than limitation and various changes may be made without departing from the scope of the invention.

25

CLAIMS

1. A system (100) for creating Learning Object (LO) from a document comprises of:
 - 5 a) a natural language processor (110) to analyse and process a document received from a user,
 - b) a multimedia repository (120) to store multimedia items in different modality, and
 - c) a learning object repository (160) to store Learning Object;wherein the system (100) is characterised in that the Learning Object
10 Creation system (100) includes:
 - a) a mapper (130) to perform semantic structure matching between semantic structure of the multimedia items and semantic structure of the text of the document;
 - 15 b) a ranker (140) to select the matched semantic structures, project graphs and calculate the average score of matched semantic structures; and
 - c) a Learning Object Builder (150) to create a Paragraph Learning Object, a Sentence Learning Object and a Topic Learning Object.
- 20 2. A method for creating Learning Object (LO) from a document by using the system (100) as claimed in claim 1 comprises the steps of:
 - a) analysing and processing a document;
 - b) matching a retrieved semantic structure by sentence index with a semantic structure from the multimedia repository (120) by a mapper
25 (130);
 - c) selecting and ranking projected graphs for each type of multimedia by a ranker (140); and
 - d) producing the Sentence Learning Object, Paragraph Learning Object and Topic Learning Object by a Learning Object Builder (150).
- 30 3. The method as claimed in claim 2, wherein matching the retrieved semantic structure by sentence index with a semantic structure from the multimedia repository (120) by the mapper (130) includes the steps of:
 - 35 a) retrieving the semantic structure by sentence index by the mapper (130);

- b) selecting the multimedia type and the semantic structure from the multimedia repository (120) by the mapper (130);
- c) performing a matching of both semantic structures selected from the multimedia repository (120) by the mapper (130) by using a projection technique;
- 5 d) storing the results in a Learning Object Repository (160) if the semantic structures match each other.
4. The method as claimed in claim 2, wherein selecting and ranking projected graphs for each type of multimedia by the ranker (140) includes the steps of:
- 10 a) selecting the projected graphs for each type of multimedia by the ranker (140);
- b) ranking the projected graphs for each type of multimedia by the ranker (140);
- 15 c) selecting the top ranking graph by the ranker (140);
- d) selecting the matched semantic structures by the ranker (140); and
- e) performing an average scoring by the ranker (140) by using different ranking methods.
- 20 5. The method as claimed in claim 2, wherein producing the Sentence Learning Object, Paragraph Learning Object and Topic Learning Object by the Learning Object Builder (150) includes the steps of:
- a) aggregating the Sentence Learning Object;
- b) creating the Paragraph Learning Object from the output of aggregation; and
- 25 c) aggregating the Paragraph Learning Object to create the Topic Learning Object.
6. The method as claimed in claim 5, wherein aggregating the Sentence Learning Object includes the steps of:
- 30 a) selecting the sentence index for aggregation;
- b) selecting the top ranked projected graph for each of the multimedia type for the respective sentence index;
- c) arranging the projected graphs according to the order of the text, video, audio and image;
- 35

- d) producing the Sentence Learning Object from the arrangement; and
- e) storing the produced Sentence Learning Object in a Learning Object repository (160).

5 7. The method as claimed in claim 5, wherein creating the Paragraph Learning Object from the output of aggregation includes the steps of:

- a) selecting the paragraph index for aggregation;
- b) selecting and grouping related projected graph from each Sentence Learning Object into the type of multimedia accordingly;
- 10 c) sorting the projected paragraphs by sentence order for each multimedia type;
- d) producing the Paragraph Learning Object from the sortation; and
- e) storing the produced Paragraph Learning Object in a Learning Object repository (160).

15

8. The method as claimed in claim 5 wherein, aggregating the Paragraph Learning Object to create the Topic Learning Object includes the steps of:

- a) selecting the topic index for aggregation;
- b) selecting and grouping related projected graph from each Paragraph Learning Object into the type of multimedia accordingly;
- 20 c) sorting the projected paragraphs by paragraph order for each multimedia;
- d) producing the Topic Learning Object from the sortation; and
- e) storing the produced Topic Learning Object in a Learning Object repository (160).

25

9. The method as claimed in claim 4, wherein the ranking methods include a Conceptual Similarity, a Relational Similarity, a Conceptual and Relational Similarity, a Domain Similarity and/or a combination thereof.

30

1/5

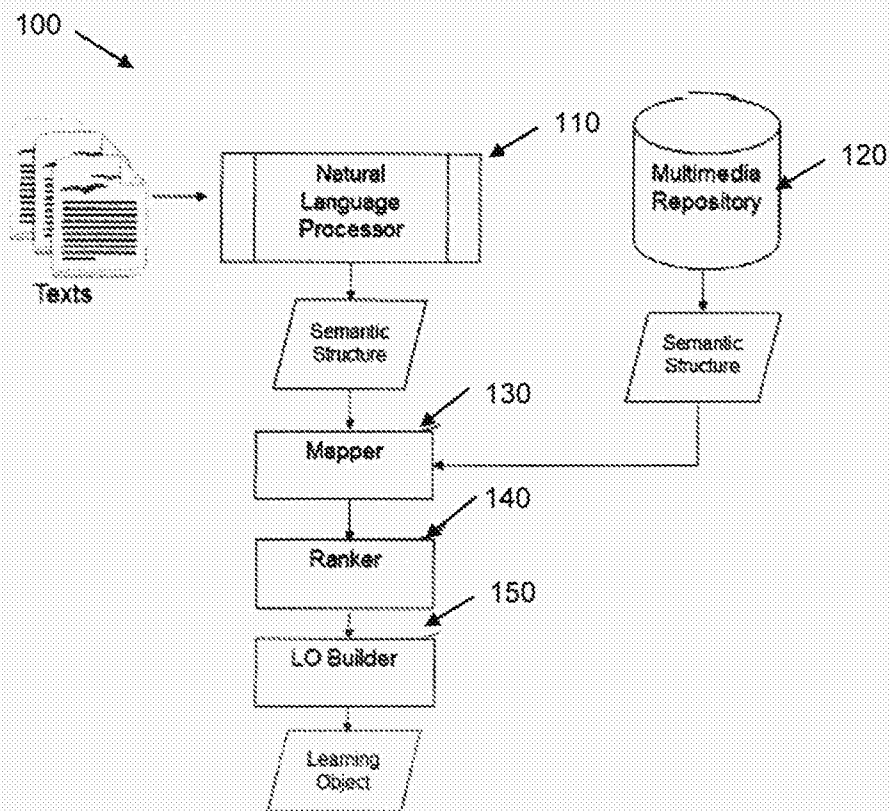


FIG. 1

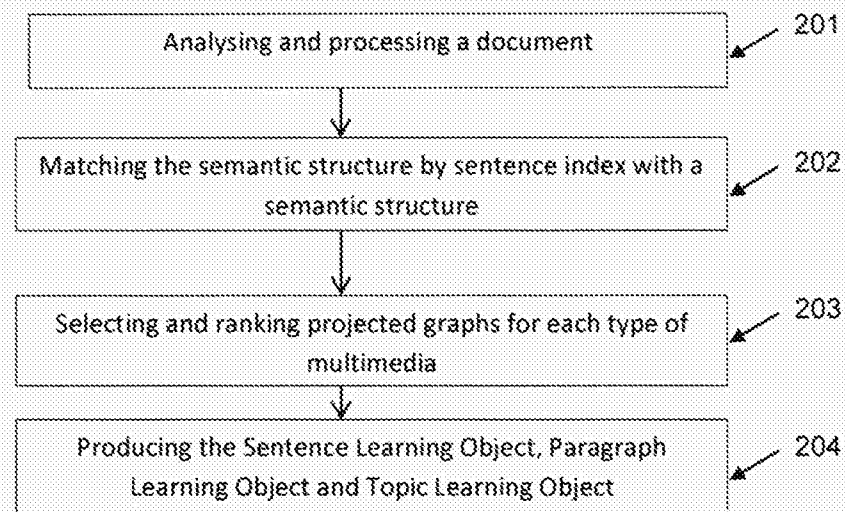


FIG. 2

2/5

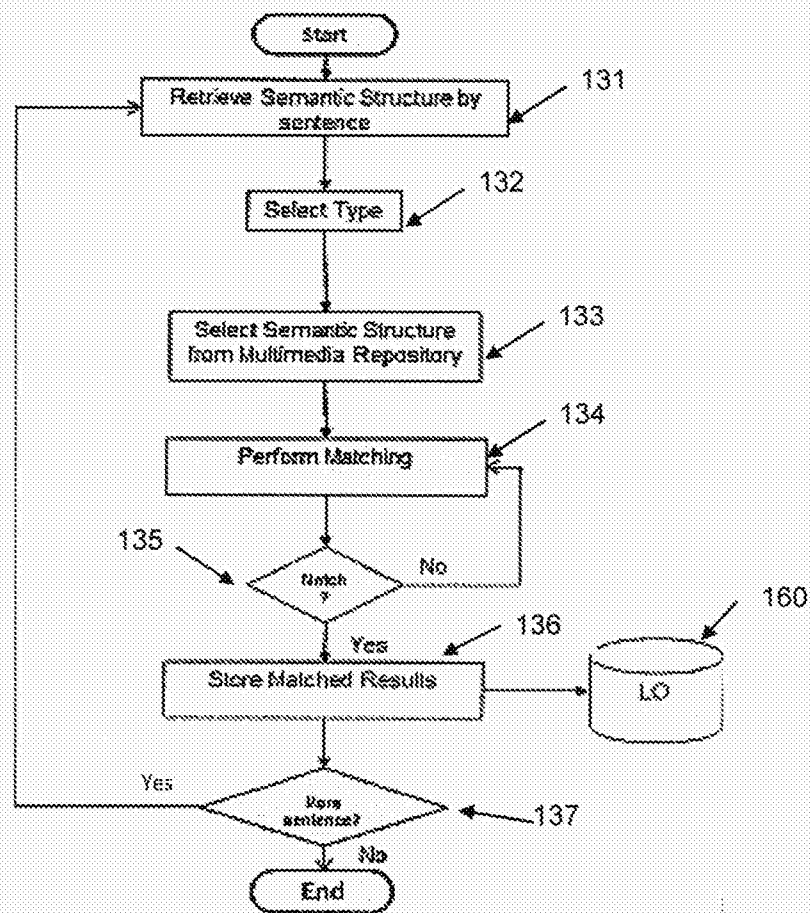


FIG. 3

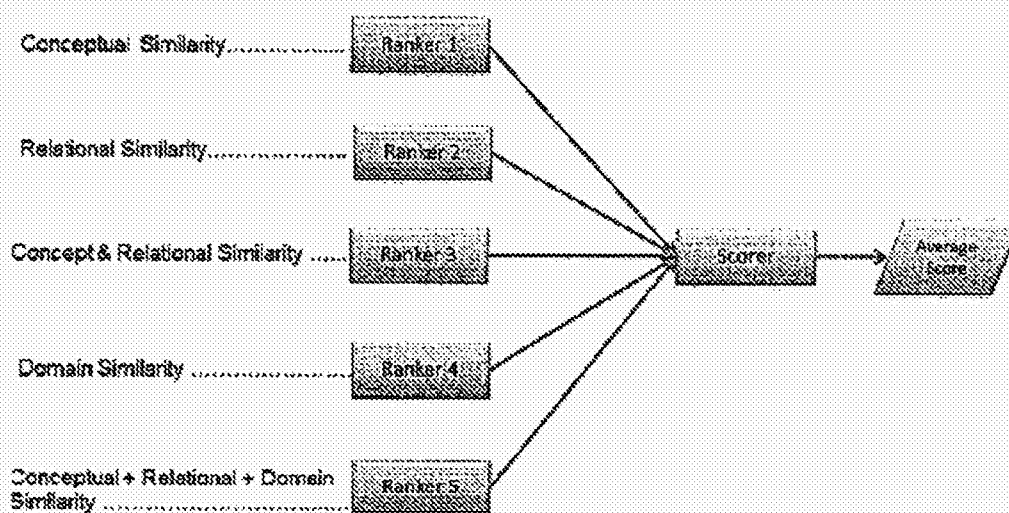


FIG. 4

3/5

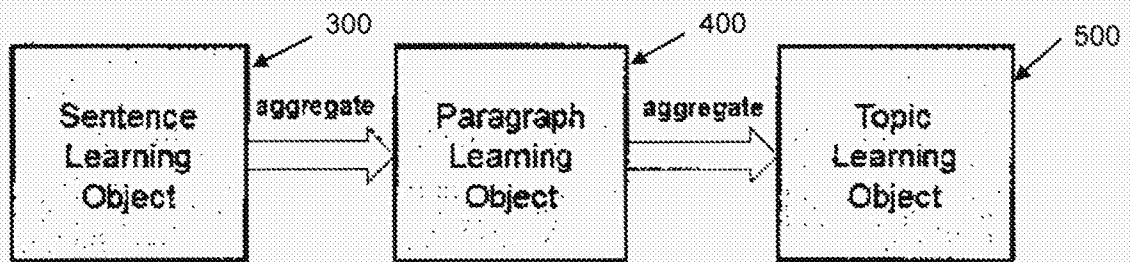


FIG. 5

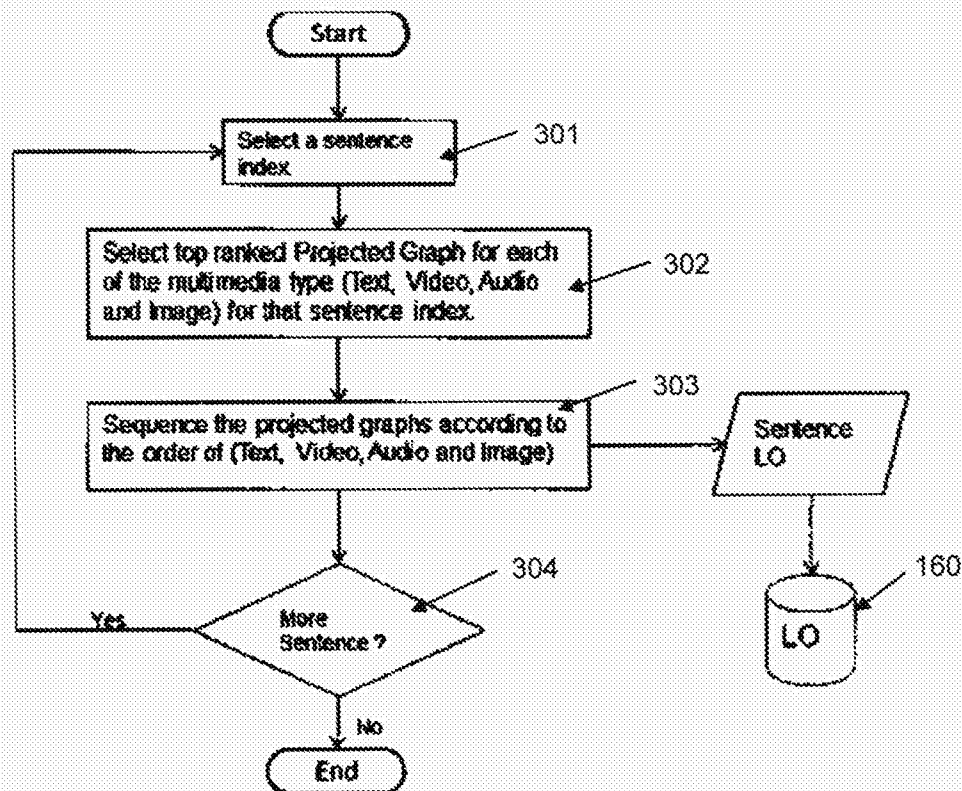


FIG. 6

4/5

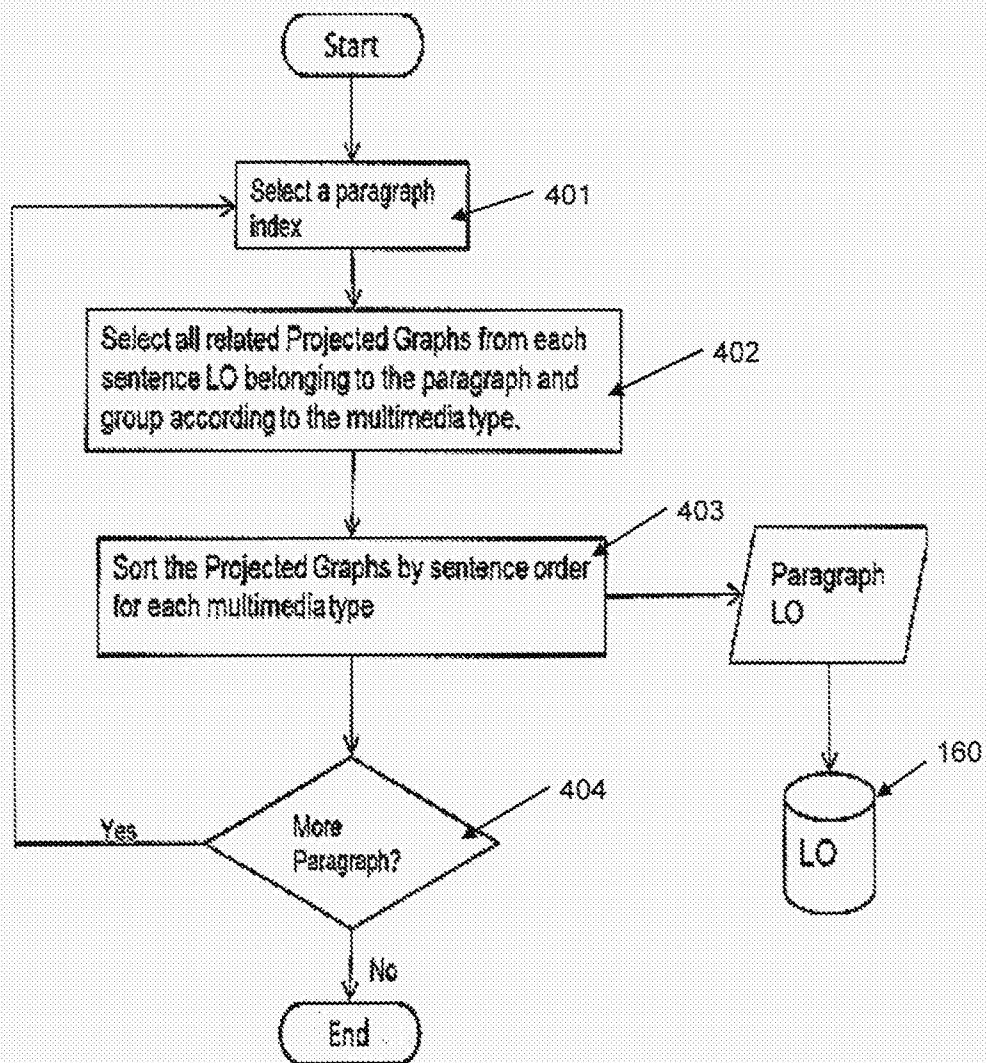


FIG. 7

5/5

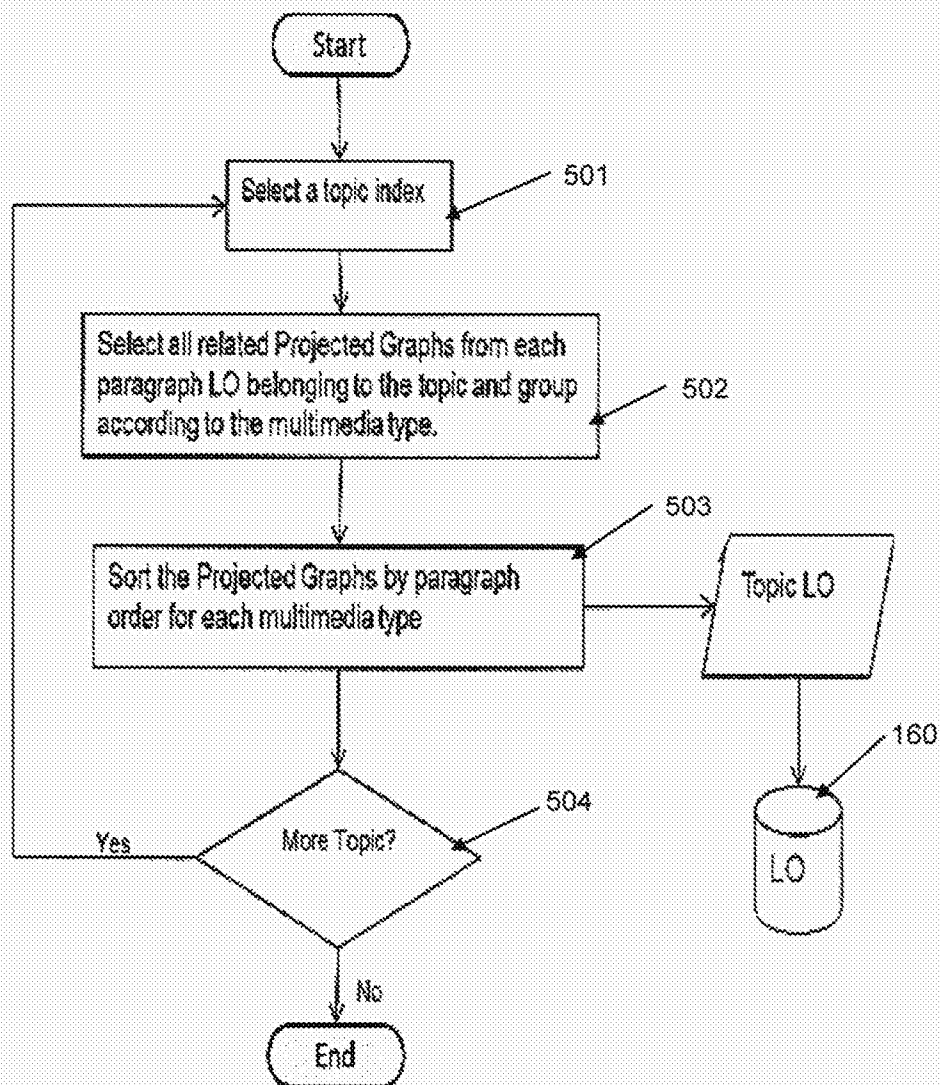


FIG. 8

INTERNATIONAL SEARCH REPORT

International application No

PCT/MY2014/000126

A. CLASSIFICATION OF SUBJECT MATTER
 INV. G09B7/08
 ADD.

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)
 G09B

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)

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2008/189684 A1 (SAREDAY SESHUKUMAR [US] ET AL SAREDAY SESHKUUMAR [US] ET AL) 7 August 2008 (2008-08-07) cited in the application the whole document	1-9
X	US 2004/205578 A1 (WOLFF ALAN S [US] ET AL) 14 October 2004 (2004-10-14) cited in the application the whole document	1-9



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

"E" earlier application or patent but published on or after the international filing date

"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)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search

9 December 2014

Date of mailing of the international search report

16/12/2014

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040,
 Fax: (+31-70) 340-3016

Authorized officer

Giemsa, Falk

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/MY2014/000126

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2008189684	A1	07-08-2008	NONE

US 2004205578	A1	14-10-2004	NONE
