[DESCRIPTION] SPECIFICATION

[TITLE OF INVENTION]

TEXT MINING SYSTEM, TEXT MINING METHOD, AND PROGRAM

[TECHNICAL FIELD]

[0001]

The present invention relates to a text mining system for obtaining, from text information, analysis of trends in contents thereof and useful knowledge.

[BACKGROUND ART]

[0002]

Text mining is a data analysis technique for, from an input of text data written in a natural language, such as texts in a free comment field in a questionnaire, grasping overall trends of their contents and finding useful knowledge. For example, in a call center, this makes it possible to grasp contents in an inquiry from an answering note, or find problems or improvements for a product from a questionnaire about the product.

[0003]

For example, PTL 1 is for extracting a syntactic dependency relationship between two or more words, summing up a frequency of appearances of syntactic dependencies, arranging words into predefined categories, and thereafter, displaying a network of syntactic dependency relationships of words (FIGs. 8, 10 and 12 in PTL 1). For example, categories may include the part of speech of words or the function in a sentence, such as a subject and an object. Moreover, in displaying the aforementioned network, display conditions regarding the category, word, and syntactic dependency relationship may be adjusted to achieve analysis at one's discretion from a wide-range overview to refined details of syntactic dependency relationships.

[CITATION LIST]

[PATENT LITERATURE]

[0004]

PTL 1: JP P2007-293685A

[SUMMARY OF INVENTION]

[TECHNICAL PROBLEM]

[0005]

However, there is a problem that such a syntactic dependency-based inter-word network does not allow for analysis with contents of texts directly understood. A reason thereof is that contents of a text cannot be accurately read from syntactic dependency relationships. For example, assuming that there are two sentences: "The personal computer B operates slow when a lot of applications are started up" (Sentence 1) and "The personal computer B is started up slow" (Sentence 2), syntactic dependency relationships "personal computer B - slow" and "started up - slow" are included in common. However, "The personal computer B operates slow" in Sentence 1 is on condition that "when a lot of applications are started up," so that what is "started up" is different between Sentences 1 and 2; thus, information is partially read from the syntactic dependency relationships based on which these sentences are related to each other. Accordingly, to understand specific contents of a text, a human being should read a sentence in the text, including its syntactic dependencies, and rearrange it.

[0006]

Embodiments of the present invention has been made in view of such a problem, and aim to provide a text mining system, a text mining method, and a program capable of analysis with contents of texts directly understood. In one aspect of the invention, a text mining system is provided. The text mining system comprises a synonym cluster acquiring section configured to acquire synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous text. The text mining system further comprises an implication relationship acquiring section configured to acquire implication relationships among said synonym clusters and an implication graph generating section configured to generate an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters. In a second aspect of the invention, a text mining method executed by a computer is provided. The text mining method comprises acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts. The text mining method further comprises acquiring implication relationships among said synonym clusters and generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters. In a third aspect of the invention, a program characterized in causing a computer to execute is provided. The program causes the computer to execute synonym cluster

acquiring processing of acquiring synonym clusters from texts in text data to be analyzed. The synonym clusters each is a collection of synonymous texts. The program also causes the computer to execute implication relationship acquiring processing of acquiring implication relationships among said synonym clusters. The program further causes the computer to execute implication graph generating processing of generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

[DISCLOSURE OF THE INVENTION]

[0007]

The present invention is a text mining system, characterized in comprising: a synonym cluster acquiring section for acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts; an implication relationship acquiring section for acquiring implication relationships among said synonym clusters; and an implication graph generating section for generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

[0008]

The present invention is a text mining method, characterized in comprising: for acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts; acquiring implication relationships among said synonym clusters; and generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym

cluster from the implication relationships among said synonym clusters.

[0009]

The present invention is a program characterized in causing a computer to execute: synonym cluster acquiring processing of acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts; implication relationship acquiring processing of acquiring implication relationships among said synonym clusters; and implication graph generating processing of generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

[0010]

According to the present invention, texts are handled in an unchanged form, for which the synonym cluster acquiring section acquires clusters each comprised of texts having different expressions but the same content (which will be referred to as a synonym cluster). The text implication relationship acquiring section acquires implication relationships among all of the generated synonym clusters. As used herein, an implication relationship is defined such that a first text implies a second text in case that contents of the second text can be read from contents of the first text. According to the configuration of the present invention, since synonymous texts are already gathered and rearranged by the synonym cluster acquiring section, an implying side includes contents differential with respect to those of an implied side. The implication graph generating section generates an implication graph having vertices of synonym clusters and directed edges each drawn from the implied side to the implying side. A vertex in the implication graph is an original text remaining unchanged. Moreover, by tracing the implication graph in a direction of the directed edge, that is, from the implied side (start point) to the implying side (end point), analysis may be achieved as to how many texts there are that include contents of a text at the start point, and what specific texts they have.

[ADVANTAGEOUS EFFECTS OF INVENTION]

[0011]

According to the present invention, analysis may be achieved with contents of texts directly understood.

[BRIEF DESCRIPTION OF DRAWINGS]

[0012]

[FIG. 1] FIG. 1 is a block diagram showing a configuration of a first embodiment of the present invention.

[FIG. 2] FIG. 2 is a flow chart showing an operation of the first embodiment of the present invention.

[FIG. 3] FIG. 3 is a diagram showing an example of data stored in a text data storing section.

[FIG. 4] FIG. 4 is a diagram showing an example of a result output from a synonym cluster acquiring section.

[FIG. 5] FIG. 5 is a diagram showing an example of a result output from an implication graph generating section.

[FIG. 6] FIG. 6 is a diagram showing an example of a result output from the implication graph generating section.

[FIG. 7] FIG. 7 is a diagram showing an example of a result from the implication graph generating section displayed on a display device.

[FIG. 8] FIG. 8 is a block diagram showing a configuration of a second embodiment of the present invention.

[FIG. 9] FIG. 9 is a flow chart showing an operation of the second embodiment of the present invention.

[FIG. 10] FIG. 10 is a diagram showing an example of a result from a frequency summing section displayed on the display device.

[DESCRIPTION OF EMBODIMENTS]

[0013]

Now embodiments of the present invention will be described with reference to the drawings.

[0014]

(First Embodiment)

A first embodiment of the present invention will be described in detail with reference to the drawings. FIG. 1 is a block diagram showing a configuration of the first embodiment of the present invention. Referring to FIG. 1, the embodiment of the present invention comprises a text data storing section 1, a text mining apparatus 2, and a display device 3.

[0015]

The text data storing section 1 stores therein data of texts to be analyzed. A text to be analyzed is a unit whose contents are understandable by a human being, and particularly, is

desirably a text comprised of one or more sentences.

[0016]

The text mining apparatus 2 comprises a synonym cluster acquiring section 20, a text implication relationship acquiring section 21, and an implication graph generating section 22.

[0017]

The synonym cluster acquiring section 20 retrieves data of the texts to be analyzed from the text data storing section 1, and acquires synonym clusters each comprised of texts having the same content. The synonym cluster acquiring section 20 may employ any method of deciding synonymity between texts.

[0018]

The text implication relationship acquiring section 21 acquires implication relationships among synonym clusters generated by the synonym cluster acquiring section 20. The text implication relationship acquiring section 21 may employ any existing implication deciding method.

[0019]

The implication graph generating section 22 generates an implication graph comprised of vertices of synonym clusters and directed edges each reflecting a direction of a decided implication between synonym clusters from a result of the implication decision made by the text implication relationship acquiring section 21.

[0020]

The display device 3 displays an implication graph generated by the text mining apparatus 2.

[0021]

Next, an operation of the present embodiment will be particularly described with reference to FIG. 1 and a flow chart in FIG. 2.

[0022]

First, the synonym cluster acquiring section 20 reads text data from the text data storing section 1, and conducts processing of acquiring a cluster comprised of synonymous texts on the read text data (Step S1). FIG. 3 shows an example of text data stored in the text data storing section 1. As shown in FIG. 3, data of a text is comprised of a text and an ID denoting its identifier. "X" represents a specific product name. Step S1 may employ any existing method of clustering synonymous texts. In the present description, the synonym cluster acquiring section 20 conducts clustering such that texts are considered to be synonymous in case that a cosine similarity between vectors composed of content words is equal to or greater than 0.8, and

synonyms are considered to be also synonymous. For example, a text with ID 1 and a text with ID 2 result in vectors "(X, Good)," and they are considered to be synonymous because their similarity is 1.0. In case that another text having a similarity equal to or greater than 0.8 to the text with ID 2 is found, it is regarded as synonymous with the text with ID 1 even though it has a similarity less than 0.8 to the text with ID 1.

[0023]

FIG. 4 shows an example of synonym clusters obtained by the synonym clustering. A synonym cluster is comprised of an ID, which is an identifier of the synonym cluster, a list of IDs of texts that belong to the synonym cluster, and a label for the ID of the synonym cluster. The label may be created by extracting any text from the synonym cluster.

[0024]

Next, the text implication relationship acquiring section 21 conducts processing of acquiring implication relationships among synonym clusters obtained by the synonym cluster acquiring section 20 (Step S2). The text implication relationship acquiring section 21 may employ any existing text implication deciding method. In the present description, the text implication relationship acquiring section 21 decides that a first text implies a second text in case that content words appearing in the first text covers 80 % or more of those of the second text. As the text, the label of a synonym cluster is employed. At Step S2, pairs of synonym clusters are listed from all synonym clusters, and an implication relationship between each pair of synonym clusters is decided with one member as the first text and the other as the second text.

[0025]

For example, when the first text is in a synonym cluster having ID 1, and the second text is in a synonym cluster having ID 2, the content words of the first text are "(X, Good)" and those of the second text are "(X, Design, Good)," so that the content word coverage is 66 % (= 2/3), which does not constitute an implication relationship. On the other hand, when the first text and the second text switch positions, the content words of the first text are "(X, Design, Good)" and those of the second text are "(X, Good)," resulting in coverage of 100 % (= 2/2), which is decided to constitute an implication relationship. This processing is applied to all pairs of synonym clusters. While a result of a decision on an implication between labels of synonym clusters is defined here as a result of a decision on an implication between synonym clusters, results of a decision on an implication between paired texts belonging to two synonym clusters may be combined to make an implication decision between synonym clusters because a synonym cluster has a plurality of texts. This may result in an improved likelihood to make a correct decision on an implication when a result of a decision on an implication between texts of the

labels of synonym clusters would be incorrect.

[0026]

Finally, the implication graph generating section 22 conducts processing of creating, from the results of the decision on an implication between synonym clusters obtained by the text implication relationship acquiring section 21, an implication graph comprised of vertices of the synonym clusters and directed edges from implied synonym clusters to implying synonym clusters (Step S3).

[0027]

FIG. 5 shows an example of a created implication graph. In the example shown in FIG. 5, a directed edge is drawn from an implied side to an implying side. The implication graph in FIG. 5 has two implication graphs comprised of a collection of texts implying "X is good" and a collection of texts implying "X is no good." In other words, this means that there is neither a text implied by "X is good" nor a text implied by "X is no good." For example, it can be seen from this implication graph that there are generally opinions "X is good" and "X is no good," and other opinions containing the meaning of (implying) "X is good" include "Liquid crystal of X is good," "X is light-weight and is good," "Design of X is good," and the like, by tracing arrows.

[0028]

Since an implication relationship holds a transitive principle, an implication graph may sometimes become redundant. For example, while it is possible to trace from "X is good" via "Liquid crystal of X is good" to "Liquid crystal of X is clear," it is also possible to directly trace from "X is good" to a "Liquid crystal of X is clear." In this case, since to take an overview of texts, rearrangement according to a semantic hierarchical structure is desirable, it is more desirable to arrange the implication graph so that the fact that there is also a text "Liquid crystal of X is good" in between can be recognized, rather than directly tracing from "X is good" to "Liquid crystal of X is clear". Accordingly, in case that a plurality of paths are found between an arbitrary pair of vertices, the implication graph generating section 22 may construct an implication graph in which only the longest path are retained. The length of a path may be calculated using, for example, the number of edges included in that path. At that time, the implication graph on the left of FIG. 5 is generated as in FIG. 6. For example, paths tracing from "X is good" to "Liquid crystal of X is clear" include a path of a directly tracing edge and a path tracing along two edges routing via "Liquid crystal of X is good." In this case, the path of a directly tracing edge is eliminated by selecting the longest path. This processing may be applied to all pairs of vertices to construct an implication graph so as to have as deep a hierarchy as possible, as in FIG. 6.

[0029]

Moreover, the implication graph generating section 22 may calculate a score of a likelihood of implication for an implication relationship between each pair of synonym clusters, and give the score to the implication relationship. By using the score, it is possible to implement processing of conducting analysis in which only edges of implication with higher possibility (with a score higher than a predetermined value) are retained. In this case, it may be contemplated that the length of a path is defined as a sum of the score of edges included in that path.

[0030]

The display device 3 displays the implication graph thus created by the implication graph generating section 22. Since the whole implication graph is extensive, a method of displaying desirably involves displaying a synonym cluster that cannot be traced from any vertex as a start point, and by clicking on the synonym cluster, developing proximate synonym clusters. FIG. 7 shows an example of the method of displaying the implication graph shown in FIG. 6. FIG. 6 has two vertices that cannot be traced from any synonym cluster: "X is good" and "X is no good." Once "X is good" is clicked on, display processing is conducted involving developing "Liquid crystal of X is good," "X is light-weight and good," and "Design of X is good" that can be traced from "X is good" along one edge. Other display methods may be applied here, including a method of developing a proximity that can be traced by a certain number of edges, instead of one edge, or display processing involving defining a limit such that a number smaller than a specified value of synonym clusters are displayed, and developing a proximity within that limit. Such display control processing may be achieved by the implication graph generating section 22 in the text mining apparatus 2 or by the display device 3.

[0031]

According to the invention in the first embodiment, analysis may be achieved with contents of texts directly understood. A reason thereof is that texts are handled in an unchanged form, and an implication graph having vertices of synonym clusters and directed edges each representing an implication relationship is generated, so that a vertex represents information in a text as-is, and other texts substantially including contents of the text may be traced according to the implication graph.

[0032]

(Second Embodiment)

Next, a second embodiment of the present invention will be described with reference to FIG. 8. Referring to FIG. 8, the second embodiment of the present invention further comprises a

frequency summing section 40, in addition to the text mining apparatus in the invention of the first embodiment.

[0033]

The frequency summing section 40 conducts processing of summing up a frequency of information represented by each synonym cluster using an implication graph generated by the invention of the first embodiment, and additionally using frequencies of information in synonym clusters that can be traced from that synonym cluster according to the implication graph.

[0034]

An operation of the present embodiment will be particularly described with reference to FIG. 8 and a flow chart in FIG. 9.

[0035]

Steps S1 through S3 in FIG. 8 are similar to those in the flow chart in FIG. 2 showing the first embodiment. The frequency summing section 40 sums up the frequency of information in a synonym cluster of each vertex in an implication graph as the number of text IDs of texts included in the synonym cluster of that vertex plus those included in synonym clusters of vertices that can be moved from that vertex (Step S4). This is achieved by, for example, searching the implication graph from each vertex in a depth-first manner to list a collection of vertices that can be reached, and taking a sum of the numbers of the text IDs in the synonym clusters of the vertices. Besides, one may employ any algorithm capable of listing a collection of vertices that can be reached from a certain vertex.

[0036]

The display device 3 displays a result of the frequency summation by the frequency summing section 40. Basically, a method that may be contemplated involves displaying a result of frequency summation over each synonym cluster node in FIG. 7. Another method that may also be contemplated involves displaying a result in a tabular format. FIG. 10 shows an example of a result of frequency summation according to the implication graph in FIG. 7 displayed in a tabular format. First, there are two synonym clusters displayed: "X is good" and "X is no good," and respective frequencies of the synonym clusters are displayed. Once "X is good" is clicked on, commenced is processing of displaying synonym clusters that can be traced therefrom along one edge, and their frequencies in a table. Such display control processing may be achieved by the implication graph generating section 22 or the frequency summing section 40 in the text mining apparatus 2, or by the display device 3.

[0037]

According to the invention in the second embodiment described above, it is possible to

sum up the frequency, including texts in which the meaning of each text is substantially contained. A reason thereof is that the invention in the first embodiment can generate an implication graph according to the implication relationship, and synonym clusters that substantially contain, i.e., imply, information in a text at each vertex may be listed with reference to the implication graph.

[0038]

The text mining apparatus in accordance with the embodiments of the present invention described above may be implemented by a CPU (Central Processing Unit) loading and executing an operation program, etc., stored in a storage section, or alternatively, they may be configured in hardware. Only part of functions of the embodiments described above may be implemented by a computer program.

[0039]

Part or all of the preceding embodiments may be described as in the following appendices, although not limited thereto.

[0040]

(Supplementary note 1)

A text mining system, characterized in comprising:

a synonym cluster acquiring section for acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts;

an implication relationship acquiring section for acquiring implication relationships among said synonym clusters; and

an implication graph generating section for generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

[0041]

(Supplementary note 2)

The text mining system as recited in Supplementary note 1, characterized in that:

in case that a plurality of paths are found between two said synonym clusters, said implication graph generating section creates an implication graph in which only edges constituting the longest path are retained.

[0042]

(Supplementary note 3)

The text mining system as recited in Supplementary note 1 or 2, characterized in further

comprising:

a frequency summing section for summing up a frequency of information represented by each said synonym cluster additionally using frequencies of information in synonym clusters that can be traced from said synonym cluster in said implication graph.

[0043]

(Supplementary note 4)

The text mining system as recited in any one of supplementary notes 1 to 3, characterized in further comprising:

a display device for, in analyzing each synonym cluster using said implication graph, displaying proximate synonym clusters in said implication graph.

[0044]

(Supplementary note 5)

A text mining method, characterized in comprising:

acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts;

acquiring implication relationships among said synonym clusters; and

generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

[0045]

(Supplementary note 6)

The text mining method as recited in Supplementary note 5, characterized in:

in case that a plurality of paths are found between two said synonym clusters, creating an implication graph in which only edges constituting the longest path are retained.

[0046]

(Supplementary note 7)

The text mining method as recited in Supplementary note 5 or 6, characterized in:

summing up a frequency of information represented by each said synonym cluster additionally using frequencies of information in synonym clusters that can be traced from said synonym cluster in said implication graph.

[0047]

(Supplementary note 8)

The text mining method as recited in any one of supplementary notes 5 to 7, characterized in:

in analyzing each synonym cluster using said implication graph, displaying proximate synonym clusters in said implication graph.

[0048]

(Supplementary note 9)

A program characterized in causing a computer to execute:

synonym cluster acquiring processing of acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts;

implication relationship acquiring processing of acquiring implication relationships among said synonym clusters; and

implication graph generating processing of generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

[0049]

(Supplementary note 10)

The program as recited in Supplementary note 9, characterized in that:

in case that a plurality of paths are found between two said synonym clusters, said implication graph generating processing creates an implication graph in which only edges constituting the longest path are retained.

[0050]

(Supplementary note 11)

The program as recited in Supplementary note 9 or 10, characterized in causing said computer to further execute:

frequency summing processing of summing up a frequency of information represented by each said synonym cluster additionally using frequencies of information in synonym clusters that can be traced from said synonym cluster in said implication graph.

[0051]

(Supplementary note 12)

The program as recited in any one of supplementary notes 9 to 11, characterized in causing said computer to execute:

processing of displaying, in analyzing each synonym cluster using said implication graph, proximate synonym clusters in said implication graph.

[0052]

While the present invention has been described with reference to the preferred

embodiments, it is not necessarily limited to the embodiments described above, and may be practiced with several modifications within a scope of the technical idea thereof.

The present application claims priority based on Japanese Patent Application No. 2012-101017 filed on April 26, 2012, disclosure of which is incorporated herein in its entirety.

[REFERENCE SIGNS LIST]

[0053]

1 Text data storing section

2 Text mining apparatus

20 Synonym cluster acquiring section

21 Text implication relationship acquiring section

22 Implication graph generating section

3 Display device

4 Text mining apparatus

40 Frequency summing section

1. A text mining system comprising:

a synonym cluster acquiring section configured to acquire synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts;

an implication relationship acquiring section configured to acquire implication relationships among said synonym clusters; and

an implication graph generating section configured to generate an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

2. The text mining system according to claim 1, wherein:

in case that a plurality of paths are found between two said synonym clusters, said implication graph generating section creates the implication graph in which only edges constituting the longest path are retained.

3. The text mining system according to claim 1 or 2, further comprising:

a frequency summing section configured to sum up a frequency of information represented by each said synonym cluster additionally using frequencies of information in synonym clusters that can be traced from said synonym cluster in said implication graph.

4. The text mining system according to any one of claims 1 to 3, further comprising:

a display device configured to, in analyzing each synonym cluster using said implication graph, display proximate synonym clusters in said implication graph.

5. A text mining method executed by a computer comprising:

acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts;

acquiring implication relationships among said synonym clusters; and

generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

6. The text mining method according to claim 5, comprising:

in case that a plurality of paths are found between two said synonym clusters, creating the implication graph in which only edges constituting the longest path are retained by said computer.

7. The text mining method according to claim 5 or 6, comprising:

summing up a frequency of information represented by each said synonym cluster additionally using frequencies of information in synonym clusters that can be traced from said synonym cluster in said implication graph by said computer.

8. The text mining method according to any one of claim 5 to 7, comprising:

in analyzing each synonym cluster using said implication graph, displaying proximate synonym clusters in said implication graph by said computer.

9. A program characterized in causing a computer to execute:

synonym cluster acquiring processing of acquiring synonym clusters from texts in text data to be analyzed, said synonym clusters each being a collection of synonymous texts;

implication relationship acquiring processing of acquiring implication relationships among said synonym clusters; and

implication graph generating processing of generating an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among said synonym clusters.

[DOCUMENT NAME] ABSTRACT

TEXT MINING SYSTEM, TEXT MINING METHOD, AND PROGRAM

The present invention is a text mining system comprising a synonym cluster acquiring section configured to acquire synonym clusters from texts in text data to be analyzed, the synonym clusters each being a collection of synonymous texts, an implication relationship acquiring section configured to acquire implication relationships among the synonym clusters, and an implication graph generating section configured to generate an implication graph including vertices of synonym clusters and directed edges each indicating a direction from an implied synonym cluster to an implying synonym cluster from the implication relationships among the synonym clusters.