

[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]

~~The Embodiments of the~~ present invention has been made in view of such a problem, and ~~its object is 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

~~[DOCUMENT NAME] SCOPE OF CLAIM FOR PATENT~~

1 A text mining system comprising:

5 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

10 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:

15 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:

20 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:

25 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 comprising:

30 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 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~~

- 5      ~~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.~~

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.