

becomes a synonym with different synonym words over the time, the time series correlation calculated by the method as in Patent Document 1 does not become high. As a result, a synonym relation cannot be extracted.

[0014]

Further, with the method depicted in Patent Document 2, it is possible to generate a synonym candidate for a synonym word by using words used for ciphers or omissions (○, “cl” that is a combination of “c” and “l” for “d”). However, it is not possible to grasp the changes in the meaning of the synonym candidate over time.

With the method depicted in Patent Document 3, time information is used for determining synonym words. However, the information from a same information source (broadcast station) is taken as the target, so that it cannot be employed for a text set gathered from an unspecified large number of sources.

[0015]

Further, with the techniques of Non-Patent Document 1, Patent Documents 1 to 3 described above, and combinations thereof, the synonymity between the synonym candidates and the synonym word cannot be determined correctly when the meanings of the synonym candidates change over time.

[0016]

~~It is therefore an object~~Embodiments of the present invention ~~to provide a synonym relation determination device, a synonym relation determination method, and a program thereof for making it possible to effectively extract and specify the synonym relation of the synonym candidate whose meaning changes over time from natural words used in texts from an unspecified large number of sources.~~

DISCLOSURE OF THE INVENTION

[0017]

According to a first aspect of the present invention, there is provided a synonym relation determination device, comprising: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of a synonym relation are recorded in a mutually corresponding manner; and synonym relation determination/specification ~~means for determining and specifying~~ module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, wherein the synonym relation determination/specification ~~means~~ module comprises: text gathering ~~means for gathering unit which gathers~~ the texts inputted from outside, and ~~generating~~ generates a text set whose issuing time can be specified; synonym candidate detection ~~means for specifying and outputting unit which specifies and outputs~~ a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set gathered by the text gathering ~~means~~ unit; and synonym period specification ~~means for determining and specifying unit which determines and specifies~~ a time interval in which the synonym expression candidate and the synonym source expression are

in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof, wherein the synonym candidate detection unit comprises: synonym candidate search unit which detects and counts the synonym expression candidates from the text set whose issuing time can be specified gathered by the text gathering unit, and specifies the time interval with a largest occurrence number per unit time as a time interval where the synonym expression candidates exist; and synonym source search unit which detects and counts the synonym source expression from the text set whose issuing time can be specified gathered by the text gathering unit, and specifies the time interval with a largest occurrence number per unit time as a time interval where the synonym source expressions exist.

A synonym relation determination device according to an embodiment of the present invention includes: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of the synonym relation are recorded in a mutually corresponding manner; and a synonym relation determination/specification module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard.

Further, the synonym relation determination/specification module includes: a text gathering unit which gathers the texts inputted from outside, and generates a text set whose issuing time can be specified; a synonym candidate detection module which specifies and outputs a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set gathered by the text gathering unit-14; and a synonym period specification module which determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof.

[0018]

According to a second aspect of the present invention, there is provided a synonym relation determination method used for a synonym relation determination device comprising: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of a synonym relation are recorded in a mutually corresponding manner; and a synonym relation determination/specification module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, wherein: a text gathering unit of the synonym relation determination/specification module

gathers the texts inputted from outside and generates a text set whose issuing time can be specified based thereupon; the synonym relation determination/specification module determines and specifies a synonym relation between the synonym expression candidate and the synonym source expression contained in the generated text set based on a specific standard; in a step for specifying the synonym relation, a synonym candidate detection unit of the synonym relation determination/specification module searches and specifies a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set gathered, respectively; and then, a synonym period specification unit of the synonym relation determination/specification module determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof, wherein in a step for detecting the synonym candidate, the synonym expression candidates are detected and counted from the text set gathered by the text gathering unit, and the time interval with a largest occurrence number per unit time is extracted and specified; and at about the same time, the synonym source expressions are detected and counted from the text set gathered by the text gathering unit, and the time interval with a largest occurrence number per unit time is extracted and specified.

A synonym relation determination method according to an embodiment of the present invention is used for a synonym relation determination device which includes: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of the synonym relation are recorded in a mutually corresponding manner; and a synonym relation determination/specification module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, wherein:

a text gathering unit of the synonym relation determination/specification module gathers the texts inputted from outside and generates a text set whose issuing time can be specified based thereupon (a text gathering/generating step);

the synonym relation determination/specification unit determines and specifies a synonym relation between the synonym expression candidate and the synonym source expression contained in the generated text set based on a specific standard (a synonym relation determining/specifying step);

in a step for specifying the synonym relation, a synonym candidate detection module of the synonym relation determination/specification unit searches and specifies a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set, respectively (a synonym candidate detecting step); and

subsequently, a synonym period specification module of the synonym relation

determination/specification unit determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof (synonym period specifying step).

[0019]

According to a third aspect of the present invention, there is provided a synonym relation determination program used for a synonym relation determination device comprising: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of a synonym relation are recorded in a mutually corresponding manner; and a synonym relation determination/specification module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, the program including: a text gathering/generation processing function which gathers the texts inputted from outside and generates based thereupon a text set whose issuing time can be specified; and a synonym relation specification processing function which executes processing for determining and specifying a synonym relation between the synonym expression candidate and the synonym source expression contained in the generated text set based on a specific standard; wherein: the synonym relation specification processing function includes a synonym candidate detection processing function which executes processing for searching and specifying a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set gathered by the text gathering unit, respectively, and a synonym period specification processing function which determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof; and the program causes a computer provided to the synonym relation determination/specification module to execute each of the processing functions, the synonym candidate detection processing function includes a synonym candidate correspondence/time interval specification processing function which detects and counts the synonym expression candidates from the text set gathered by the text gathering unit, and extracts and specifies the time interval with a largest occurrence number per unit time, and a synonym source correspondence/time interval specification processing function which detects and counts the synonym source expressions from the text set gathered by the text gathering unit, and extracts and specifies the time interval with a largest occurrence number per unit time; and the program causes a computer provided to the synonym relation determination/specification module to execute each of the processing functions.

A synonym relation determination program according to an embodiment of the present invention is used for a synonym relation determination device which includes: a synonym expression candidate storage unit 10 in which a prescribed synonym source expression and a plurality of synonym expression candidates are recorded in a mutually corresponding manner; and a synonym relation determination/specification module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, and the program includes:

a text gathering/generation processing function which gathers the texts inputted from outside and generates based thereupon a text set whose issuing time can be specified; and a synonym relation specification processing function which executes processing for determining and specifying a synonym relation between the synonym expression candidate and the synonym source expression contained in the generated text set based on a specific standard, wherein:

the synonym relation specification processing function includes a synonym candidate detection processing function which executes processing for searching and specifying a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set gathered by the text gathering unit, respectively, and a synonym period specification processing function which determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof; and

the program causes a computer provided to the synonym relation determination/specification module to execute each of the processing functions.

[0020]

As described above, embodiments of the present invention are structured to determine the synonymity by capturing the point at which the synonym expression candidates appeared many times, so that the start time at which the synonym relation is established can be outputted. Therefore, it becomes possible to provide the excellent synonym relation determination device, synonym relation determination method, and program thereof, which are not in the related techniques described above and capable of determining the time interval where the synonym relation is established when the synonymity changes over time.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram showing a first embodiment of a synonym relation determination device according to the present invention;

FIG. 2 is an explanatory chart showing an example of a synonym candidate list used in FIG. 1;

FIG. 3 is an explanatory chart showing examples of time transition regarding occurrence

CLAIMS

1. A synonym relation determination device, comprising: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of a synonym relation are recorded in a mutually corresponding manner; and synonym relation determination/specification ~~means for determining and specifying~~ module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, wherein

the synonym relation determination/specification ~~means~~ module comprises:

text gathering ~~means for gathering~~ unit which gathers the texts inputted from outside, and ~~generating~~ generates a text set whose issuing time can be specified;

synonym candidate detection ~~means for specifying and outputting~~ unit which specifies and outputs a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set gathered by the text gathering ~~means~~ unit; and

synonym period specification ~~means for determining and specifying~~ unit which determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof, wherein

the synonym candidate detection unit comprises: synonym candidate search unit which detects and counts the synonym expression candidates from the text set whose issuing time can be specified gathered by the text gathering unit, and specifies the time interval with a largest occurrence number per unit time as a time interval where the synonym expression candidates exist; and synonym source search unit which detects and counts the synonym source expression from the text set whose issuing time can be specified gathered by the text gathering unit, and specifies the time interval with a largest occurrence number per unit time as a time interval where the synonym source expressions exist.

2. The synonym relation determination device as claimed in claim 1, wherein

~~the synonym candidate detection means comprises: synonym candidate search means~~

~~for detecting and counting the synonym expression candidates from the text set whose issuing time can be specified gathered by the text gathering means, and specifying the time interval with a largest occurrence number per unit time as a time interval where the synonym expression candidates exist; and synonym source search means for detecting and counting the synonym source expression from the text set whose issuing time can be specified gathered by the text gathering means, and specifying the time interval with a largest occurrence number per unit time as a time interval where the synonym source expressions exist.~~

32. The synonym relation determination device as claimed in claim 1 ~~or 2~~, wherein

the synonym period specification ~~means~~ unit comprises: synonym relation extraction ~~means for extracting unit which extracts~~ at least the synonym source expression whose occurrence number is largest within the text set in the time interval in which the synonym expression candidates are detected by the synonym candidate detection ~~means~~ unit as a synonym source to be in a synonym relation; and synonym period commencement determination ~~means~~ for determining which determines that the extracted synonym expression is in a synonym relation with the synonym expression candidate, ~~taking~~ takes a start point of the time interval in which the both are in the synonym relation as a start point of the synonym period, and ~~registering~~ registers it along with the synonym period to a synonym dictionary provided in advance.

43. The synonym relation determination device as claimed in claim 32, wherein

the synonym period commencement determination ~~means~~ unit includes a function which, within the set in the time intervals where the synonym candidates are detected in the text set, determines also the synonym source expression having a largest occurrence ratio between the occurrence number per unit time in the time interval and the occurrence number per unit time before the time interval as being in a synonym relation with the synonym expression candidate.

54. The synonym relation determination device as claimed in claim 1, 2 ~~or 3, 3, or 4~~ wherein

the synonym period specification ~~means~~ unit comprises synonym period end determination ~~means for determining unit which determines~~ that the synonym relation is cancelled at a point where the occurrence number per unit period of the synonym expression candidate determined and specified as being in the synonym relation by the synonym period commencement determination ~~means~~ unit of the synonym period specification ~~means~~ unit becomes equal to or lower than a threshold value set in advance.

65. A synonym relation determination method used for a synonym relation determination device comprising: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of a synonym relation are recorded in a mutually corresponding manner; and a synonym relation determination/specification module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, wherein:

a text gathering unit of the synonym relation determination/specification module gathers the texts inputted from outside and generates a text set whose issuing time can be specified based thereupon;

the synonym relation determination/specification module determines and specifies a synonym relation between the synonym expression candidate and the synonym source expression contained in the generated text set based on a specific standard;

in a step for specifying the synonym relation, a synonym candidate detection unit of the synonym relation determination/specification module searches and specifies a time interval where the synonym expression candidates are detected many times and a time interval where the synonym source expression is detected many times from the text set gathered, respectively; and

then, a synonym period specification unit of the synonym relation determination/specification module determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof, wherein

in a step for detecting the synonym candidate, the synonym expression candidates are detected and counted from the text set gathered by the text gathering unit, and the time interval with a largest occurrence number per unit time is extracted and specified; and

at about the same time, the synonym source expressions are detected and counted from the text set gathered by the text gathering unit, and the time interval with a largest occurrence number per unit time is extracted and specified.

~~7. The synonym relation determination method as claimed in claim 6, wherein:~~

~~in a step for detecting the synonym candidate, the synonym expression candidates are detected and counted from the text set gathered by the text gathering unit, and the time interval with a largest occurrence number per unit time is extracted and specified;~~

~~at about the same time, the synonym source expressions are detected and counted from the text set gathered by the text gathering unit, and the time interval with a largest occurrence number per unit time is extracted and specified; and~~

~~operation contents in specifying steps of each time interval mentioned above are executed by the synonym candidate detection unit.~~

86. The synonym relation determination method as claimed in claim 65, wherein:

in s a step for specifying the synonym interval, at least the synonym source expression whose occurrence number is largest within the text set in the time interval in which the synonym expression candidates are detected in the step for detecting the synonym candidate is extracted as a synonym source to be in a synonym relation;

at the same time, the extracted synonym expression is determined to be in a synonym relation with the synonym expression candidate, a start point of the time interval in which the both are in the synonym relation is taken as a start point of the synonym period, and it is registered along with the synonym period to a synonym dictionary provided in advance; and

operation contents of each of the steps for extraction/registration are executed by the synonym period specification module.

97. A synonym relation determination program used for a synonym relation determination device comprising: a synonym expression candidate storage unit in which a prescribed synonym source expression and a plurality of synonym expression candidates as targets of a synonym relation are recorded in a mutually corresponding manner; and a synonym relation determination/specification module which determines and specifies a synonym relation between the synonym expression candidates and the synonym source expression in texts inputted from outside based on a specific standard, the program including:

a text gathering/generation processing function which gathers the texts inputted from outside and generates based thereupon a text set whose issuing time can be specified; and

a synonym relation specification processing function which executes processing for determining and specifying a synonym relation between the synonym expression candidate and the synonym source expression contained in the generated text set based on a specific standard; wherein:

the synonym relation specification processing function includes

a synonym candidate detection processing function which executes processing for searching and specifying a time interval where the synonym expression candidates are detected

many times and a time interval where the synonym source expression is detected many times from the text set gathered by the text gathering unit, respectively, and

a synonym period specification processing function which determines and specifies a time interval in which the synonym expression candidate and the synonym source expression are in a synonym relation as a synonym period based on a positional relation between the time interval where the synonym expression candidate is detected in the text set and the time interval where the synonym source expression is detected in the text set and based on detection frequency thereof; and

the program causes a computer provided to the synonym relation determination/specification module to execute each of the processing functions,

the synonym candidate detection processing function includes a synonym candidate correspondence/time interval specification processing function which detects and counts the synonym expression candidates from the text set gathered by the text gathering unit, and extracts and specifies the time interval with a largest occurrence number per unit time, and

a synonym source correspondence/time interval specification processing function which detects and counts the synonym source expressions from the text set gathered by the text gathering unit, and extracts and specifies the time interval with a largest occurrence number per unit time; and

the program causes a computer provided to the synonym relation determination/specification module to execute each of the processing functions.

~~10. The synonym relation determination program as claimed in claim 9, wherein:~~

~~the synonym candidate detection processing function includes a synonym candidate correspondence/time interval specification processing function which detects and counts the synonym expression candidates from the text set gathered by the text gathering unit, and extracts and specifies the time interval with a largest occurrence number per unit time, and~~

~~a synonym source correspondence/time interval specification processing function which detects and counts the synonym source expressions from the text set gathered by the text gathering unit, and extracts and specifies the time interval with a largest occurrence number per unit time; and~~

~~the program causes a computer provided to the synonym relation determination/specification module to execute each of the processing functions.~~