

[0012] For example, assume that "text T: Company A withdrew from the personal computer business" and "text H: Personal computers from Company A will disappear " exist. The above-described entailment determination system disclosed in Non-Patent Document 1 extracts, as VASs, "withdraw (subject: Company A, object: personal computer, business)" from the text T and "disappear (subject: Company A, from, personal computers)" from the text H.

[0013] In the case where such VASs are extracted, "withdraw" and "disappear" are different verbs. Consequently, in the above-described entailment determination system disclosed in Non-Patent Document 1, the text T is not determined to entail the text H, even though it does.

[0014] Furthermore, in the entailment determination system, even if all verbs are viewed as being equal in the aforementioned case, the subject of "disappear" is different from the subject and object of "withdraw", and therefore, the text T is not determined to entail the text H in this case as well.

[0015] Object of the Invention

Embodiments~~An example of the object of the present invention seeks~~ to solve the above-mentioned problems and to provide a textual entailment recognition apparatus, a textual entailment recognition method, and a computer-readable recording medium that can allow for entailment determination of a plurality of texts that are to undergo determination, even if the argument structures of the predicates are different.

#### Means for Solving the Problem

[0016] According to an aspect of the present invention ~~for achieving the foregoing object,~~ there is provided an textual entailment determination apparatus for determining whether a first text entails a second text, comprising~~including~~:

a vector generation unit that acquires a predicate-argument structure of each of the first text and the second text, and generates, for each of the first text and the second text, a vector having a character string as a component, for each predicate-argument structure by using a word other than a word indicating a type of argument of a predicate in the predicate-argument structure;

a combination identification unit that generates vectors from each ~~of~~compares the vector generated for each predicate-argument structure for the first text and the vector generated for each predicate-argument structure for the second text by converting the character strings that are components of the vectors into numbers, calculates similarity of the two generated vectors, and identifies combinations of the predicate-argument structures of the first text and the predicate-argument structures of the second text based on the calculated similarity~~a result of the comparison~~; and

an entailment determination unit that obtains ~~a feature amount~~ for each of the identified combinations, as a feature amount, either the degree of word coverage or the degree of word matching for only words serving as the arguments between the predicate-argument structures of the first text and the predicate-argument structures of the second text based on a word other than a word indicating a type of argument of a predicate in the predicate-argument structure and determines whether the first text entails the second text based on the obtained feature amounts.

[0017] Also, according to another aspect of the present invention ~~for achieving the foregoing object~~, there is provided an textual entailment determination method for determining whether a first text entails a second text, comprising~~including~~ the steps of:

(a) by a computer, acquiring a predicate-argument structure of each of the first text and the second text and generating, for each of the first text and the second text, a vector having a character string as a component, for each predicate-argument structure by using a word other than a word indicating the type of argument of a  
5 predicate in the predicate-argument structure~~acquiring a predicate-argument structure of each of the first text and the second text, and generating, for each of the first text and the second text, a vector for each predicate-argument structure by using a word other than a word indicating the type of argument of a predicate in the predicate-argument structure;~~

10 (b) by the computer, generating vectors from each of the vector generated for each predicate-argument structure for the first text and the vector generated for each predicate-argument structure for the second text by converting the character string that are components of the vectors into numbers, calculating similarity of the two generated vectors, and identifying combinations of the predicate-argument  
15 structures of the first text and the predicate-argument structures of the second text based on the calculated similarity;~~and comparing the vector generated for each predicate-argument structure for the first text and the vector generated for each predicate-argument structure for the second text, and identifying combinations of the predicate-argument structures of the first text and the predicate-argument~~  
20 ~~structure of the second text based on a result of the comparison; and~~

(c) by the computer, obtaining, for each of the identified combinations, as a feature amount, either the degree of word coverage or the degree of word matching for only words serving as the arguments between the predicate-argument structures of the first text and the predicate-argument structures of the second text based on a word other than a word indicating a type of argument of a predicate in the predicate-argument structure and determining whether the first text entails the second text based on the obtained feature amount~~obtaining a feature amount for each of the identified combinations, and determining whether the first text entails the second text based on the obtained feature amounts.~~

[0018] ~~Furthermore, according to another aspect of the present invention for achieving the foregoing object, there is provided a computer-readable recording medium having recorded thereon a program for determining whether a first text entails a second text by a computer, the program including an instruction for causing the computer to execute the steps of:~~

(a) acquiring a predicate-argument structure of each of the first text and the second text, and generating, for each of the first text and the second text, a vector having a character string as a component, for each predicate-argument structure by using a word other than a word indicating the type of argument of a predicate in the predicate-argument structure;

(b) generating vectors from each of the vector generated for each predicate-argument structure for the first text and the vector generated for each predicate-argument structure for the second text by converting the character string that are components of the vectors into numbers, calculating similarity of the two generated vectors, and identifying combinations of the predicate-argument structures of the first text and the predicate-argument structures of the second text based on the calculated similarity; and comparing the vector generated for each predicate-argument structure for the first text and the vector generated for each predicate-argument structure for the second text, and identifying combinations of the predicate-argument structures of the first text and the predicate-argument structure of the second text based on a result of the comparison; and

(c) obtaining, for each of the identified combinations, as a feature amount, either the degree of word coverage or the degree of word matching for only words serving as the arguments between the predicate-argument structures of the first text and the predicate-argument structures of the second text based on a word other than a word indicating a type of argument of a predicate in the predicate-argument structure and determining whether the first text entails the second text based on the obtained feature amounts; obtaining a feature amount for each of the identified combinations, and determining whether the first text entails the second text based on the obtained feature amounts.

### Effects of the Invention

[0019] As described above, according to embodiments of the present invention, it is possible to allow for entailment determination of a plurality of texts that are to undergo determination, even if the argument structures of the predicates are different.