uses the created statistical model to score terms in a target language that are considered as translation candidates for the concept, wherein the creation unit creates the statistical model using a covariance matrix and a mean vector of the input terms' context vectors.

Another exemplary aspect of the present invention is a term translation acquisition method which includes: creating a statistical model based on a set of input terms' context vectors, wherein the set of terms, including at least two terms, are in the same source language and describe the same concept; and using the created statistical model to score terms in a target language that are considered as translation candidates for the concept, wherein the statistical model is created using a covariance matrix and a mean vector of the input terms' context vectors.

Yet another exemplary aspect of the present invention is a computer-readable recording medium storing a program that causes a computer to execute: a creation function of creating a statistical model based on a set of input terms' context vectors, wherein the set of terms, including at least two terms, are in the same source language and describe the same concept; and a ranking function of using the created statistical model to score terms in a target language that are considered as translation candidates for the concept, wherein the creation function creates the statistical model using a covariance matrix and a mean vector of the input terms' context vectors.

According to embodiments of the present invention, the problem of sparse context vectors related to low-frequent terms, as well as the problem of noisy context vectors related to ambiguity of input terms can be mitigated. As a consequence, translation accuracy may be improved.

BRIEF DESCRIPTION OF DRAWINGS

Figure 1 is a block diagram showing the functional structure of the term translation system related to Non-Patent Document 1.

Figure 2 is a block diagram showing the functional structure of the term translation syste m related to Non-Patent Document 2.

CLAIMS

1. A term translation acquisition apparatus comprising:

a creation unit which creates a statistical model based on a set of input terms' context vectors, wherein the set of terms, including at least two terms, are in the same source language and describe the same concept; and

a ranking unit which uses the created statistical model to score terms in a target language that are considered as translation candidates for the concept,

wherein the creation unit creates the statistical model using a covariance matrix and a mean vector of the input terms' context vectors.

~~2. The apparatus according to claim 1, wherein the creation unit creates the statistical model using a covariance matrix and a mean vector of the input terms' context vectors.~~

32. The apparatus according to claim 1 ~~or 2~~, wherein the ranking unit scores each translation candidate in the target language according to the created statistical model using similarity between each translation candidate and the statistical model.

43. The apparatus according to claim 32, wherein the ranking unit uses, as the similarity, the probability that each translation candidate is observed given the created statistical model.

54. The apparatus according to claim 32, wherein the ranking unit uses, as the similarity, the posterior probability of a statistical model's parameter assuming a prior distribution over each translation candidate.

65. The apparatus according to any one of claims 1 to 54, wherein a user's input includes a

single term in the source language, and

the apparatus further comprises an extension unit which extends the single input term to the set of input terms, including at least two terms, and supplies the extended set of input terms to the creation unit.

~~7~~6. The apparatus according to claim ~~6~~5, wherein the extension unit comprises a storage unit which stores a monolingual dictionary including synonymous terms in the source language, and

the extension unit looks up synonyms which are a set of terms which are synonymous to the single input term in the monolingual dictionary, selects, among the looked up synonyms, terms which context vector is closer to the single input term's context vector than context vectors of the other terms, and supplies the selected terms and the single input term to the creation unit as the set of input terms.

~~8~~7. A term translation acquisition method comprising:

creating a statistical model based on a set of input terms' context vectors, wherein the set of terms, including at least two terms, are in the same source language and describe the same concept; and

using the created statistical model to score terms in a target language that are considered as translation candidates for the concept,

wherein the statistical model is created using a covariance matrix and a mean vector of the input terms' context vectors.

~~9~~8. A computer-readable recording medium storing a program that causes a computer to execute:

a creation function of creating a statistical model based on a set of input terms' context vectors, wherein the set of terms, including at least two terms, are in the same

source language and describe the same concept; and

a ranking function of using the created statistical model to score terms in a target language that are considered as translation candidates for the concept,

wherein the creation function creates the statistical model using a covariance matrix and a mean vector of the input terms' context vectors.