Saudi Journal of Biological Sciences (2017) 24, 132–138

King Saud University

Saudi Journal of Biological Sciences

www.ksu.edu.sa
www.sciencedirect.com

ORIGINAL ARTICLE

Margin based ontology sparse vector learning algorithm and applied in biology science

Wei Gao [a,*], Abdul Qudair Baig [b], Haidar Ali [b], Wasim Sajjad [b], Mohammad Reza Farahani [c]

[a] School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China
[b] Department of Mathematics, COMSATS Institute of Information Technology, Attock, Pakistan
[c] Department of Applied Mathematics, Iran University of Science and Technology, Narmak, 16844 Tehran, Iran

Received 29 June 2016; revised 1 September 2016; accepted 1 September 2016
Available online 13 September 2016

KEYWORDS

Ontology;
Similarity measure;
Sparse vector;
Margin

Abstract In biology field, the ontology application relates to a large amount of genetic information and chemical information of molecular structure, which makes knowledge of ontology concepts convey much information. Therefore, in mathematical notation, the dimension of vector which corresponds to the ontology concept is often very large, and thus improves the higher requirements of ontology algorithm. Under this background, we consider the designing of ontology sparse vector algorithm and application in biology. In this paper, using knowledge of marginal likelihood and marginal distribution, the optimized strategy of marginal based ontology sparse vector learning algorithm is presented. Finally, the new algorithm is applied to gene ontology and plant ontology to verify its efficiency.

1. Introduction

The term "Ontology" refers to a knowledge representation and conceptual shared model. It is widely used in gene computing, knowledge management and information retrieval, which also witnesses its effectiveness in the various applications. Besides, the concept semantic model was welcomed and borrowed by scholars in social science, medical science, biology science, pharmacology science and geography science (for instance, see Gregor et al. (2016), Kaminski et al. (2016), Forsati and Shamsfard (2016), Pesaranghader et al. (2016), Huntley et al. (2016), Brown et al. (2016), Palmer et al. (2016), Terblanche and Wongthongtham (2016), Farid et al. (2016) and Carmen Suarez-Figuero et al. (2016)).

Traditionally, we take ontology model as a graph $G = (V, E)$, where each vertex v in the ontology graph G represents a concept and each edge $e = v_i v_j$ of it represents a relationship between concepts v_i and v_j. A few years ago, ontology similarity-based technologies were quite popular among researchers due to its wide range of applications. For instance, GO-WAR algorithm was raised by Agapito et al. (2016) to explore cross-ontology association rules in which GO terms

* Corresponding author.
E-mail address: gaowei@ynnu.edu.cn (W. Gao).
Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.sjbs.2016.09.001
1319-562X

present in its sub-ontologies of GO. What's more, after mining publicly available GO annotated datasets which show how GO-WAR outperforms current state of the art approaches, a deep performance evaluation of GO-WAR was discovered. Chicco and Masseroli (2016) put forward a computational pipeline which can predict novel ontology-based gene functional annotations by means of various semantic and machine learning methods. Then, in order to categorize the predicted annotations by their likelihood of being correct, a new semantic prioritization rule was achieved in their papers. The definition of GO ontological terms, molecular function, biological process and cellular components were given in detail by Umadevi et al. (2016), and he also found the relations to that of the disease genes with p-value < 0.05. Bajenaru et al. (2016) raised the constituent parts and architecture of the proposed ontology-based e-learning system and a framework for its application in reality. However, based on the OWL2 rules and the reasoning process of the OntoDiabetic system, Sherimon and Krishnan (2016) shifted his attention to the modeling and implementation of clinical guidelines. In terms of fuzzy logic, Bobillo and Straccia (2016) extended it to classical ontologies. With the help of pre-existing information about ontologies, such as terminology and ontology structure, Trokanas and Cecelja (2016) worked out a framework for evaluation of ontology for reuse to calculate a compatibility metric of ontology suitability for reuse and hence integration. To illustrate, the framework was explained in a Chemical and Process Engineering perspective. With the aim to allow users to quickly compute, manipulate and explore Gene Ontology (GO) semantic similarity measures, Mazandu et al. (2016) proposed A-DaGO-Fun. Auffeves and Grangier (2016) raised a new quantum ontology to make usual quantum mechanics fully compatible with physical realism. Hence, the physical properties in the ontology are attributed jointly to the system. In addition, Hoyle and Brass (2016) defined a statistical mechanical theory which expresses the process of annotating an object with terms selected from an ontology.

With the consideration of ontology similarity measure and ontology mapping, some effective learning tricks turn out to work well. With the harmonic analysis and diffusion regularization on hypergraph, Gao et al. (2013) proposed a new ontology mapping algorithm. Gao and Shi (2013) raised a novel ontology similarity computation technology considering operational cost in the real applications. Using ADAL trick, an ontology sparse vector learning algorithm was worked out by Gao et al. (2015) to make contributions to the ontology similarity measuring and ontology mapping. Then, Gao et al. (2016) proposed an ontology optimization tactics using distance calculating and learning. Several theoretical analysis of ontology algorithm mentioned but not defined in detail in this paper can refer to Gao et al. (2012), Gao and Xu (2013), and Gao and Zhu (2014).

The marginal based ontology algorithm for ontology similarity computation and ontology mapping are given in the paper. By means of the sparse vector, the ontology graph is mapped into a real line and vertices into real numbers. Then, based on the difference between their corresponding real numbers, the similarity between vertices is measured. The rest of the paper is structured like this: the notations and setting are presented in Section 2; the ontology sparse vector optimization algorithm is raised in Section 3, in addition, the technologies to tackle the details in algorithm are also included here; and, the experiments on gene science and plant science are taken to show the efficiency of the algorithm in the last section.

2. Setting

Let V represent an instance space. Concerning each vertex in ontology graph, a p dimension vector represents the information: its name, instance, attribute and structure, and semantic information of the concept. All the information is related to the vertex and that is contained in name and attributes components of its vector. Let $v = \{v_1, \ldots, v_p\}$ be a vector that corresponds to a vertex v. To make the representation, clearer and further, we take a try to confuse the notations. So we consider using v to denote both the ontology vertex and its corresponding vector. The ontology learning algorithms are set to obtain an optimal ontology (score) function f: $V \to \mathbb{R}$, and the similarity between two vertices is determined by the difference between two corresponding real numbers. The core of this algorithm is dimensionality reduction, i.e., choosing one dimension vector to express p dimension vector. Specifically, an ontology function f is a dimensionality reduction function f: $\mathbb{R}^p \to \mathbb{R}$.

In the real application, one sparse ontology function is expressed by

$$f_\beta(v) = \sum_{i=1}^{p} v_i \beta_i + \delta. \tag{1}$$

Here $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ is a sparse vector and δ is a noise term. The sparse vector $\boldsymbol{\beta}$ is used to decrease the components that is not necessary to zero. Then, we learn the sparse vector $\boldsymbol{\beta}$, so that we can determine the ontology function f.

The general versions for learning $\boldsymbol{\beta}$ is learned in the paper. Let $\{v_i, y_i\}_{i=1}^n$ be a sample set with n vertex, $V \in \mathbb{R}^{n \times p}$ be the matrix of n samples such that each sample vertex lies in a p dimension space, and $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n) \in \mathbb{R}^n$ be the vector of outputs of the these n sample vertex. Hence, the regression function Eq. (1) can be expressed as the linear model:

$$\boldsymbol{y} = \boldsymbol{V}\boldsymbol{\beta} + \delta, \tag{2}$$

where δ is the n dimension vector for noise which are normally distributed, $N_p(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_p)$ with variance σ^2.

An estimate of the sparse vector is obtained from the general regression obtains after solving the optimization problem below:

$$\min_{\boldsymbol{\beta} \in \mathbb{R}^p} l(\boldsymbol{\beta}) + \lambda \|\boldsymbol{\beta}\|_1, \tag{3}$$

where $l(\boldsymbol{\beta}) = \frac{1}{2}\|\boldsymbol{y} - \boldsymbol{V}\boldsymbol{\beta}\|_2^2$ is the loss term, $\|\boldsymbol{\beta}\|_1 = \sum_{i=1}^p |\beta_i|$ is the l_1-norm balance term that measures the sparseness of vector $\boldsymbol{\beta}$, and $\lambda > 0$ is the balance parameter which controls the sparsity level. On the selection of the balance parameter λ, readers can refer to Mancinelli et al. (2013), Mukhopadhyay and Bhattacharya (2013), Ishibuchi and Nojima (2013), and Varmuza et al. (2014) for more details about the method of cross-validation.

3. Ontology algorithm describing

In our paper, we consider the special case of Eq. (3), and it can be stated as

$$\min_{\boldsymbol{\beta}\in\mathbb{R}^p}(\boldsymbol{y}-\boldsymbol{V\beta})^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{V\beta})+\lambda\sum_{j=1}^{p}|\beta_j|^{\gamma}, \tag{4}$$

where $\gamma\in\{0,1,2\}$.

We suppose the normal likelihood $\boldsymbol{y}|\boldsymbol{\beta},\sigma^2\sim N(\boldsymbol{V\beta},\sigma^2\boldsymbol{I})$ have independent priors about the ontology coefficients of the expression $p(\beta_j|\lambda)\propto\exp\{-\lambda|\beta_j|^{\gamma}\}$, $\lambda>0$, $0<\gamma<1$ and a representative conjugate prior have the error precision $\sigma^{-2}\sim\Xi(c_0,d_0)$. It's clear to those who fully grasp the criterion Bayesian patterns in which the exponential power class prior can't be a conjugate prior with the normal likelihood for $\gamma\neq 2$.

Let $N(\alpha,\varsigma)$ be a normal probability distribution function with mean α and variance ς, $g(\tau_j)\propto\tau^{-\frac{1}{2}}q(\tau_j)$ where $q(\tau_j)$ is denoted as the density of stable distribution of index $\frac{\gamma}{2}$. Then, the above mentioned class of distributions can be formulated as

$$p(\beta_j|\lambda,\gamma)=\int_0^{\infty}N(0,\tau_j^{-1}\lambda^{-\frac{2}{\gamma}})g(\tau_j)d\tau_j \tag{5}$$

Let $\Gamma(\cdot)$ be the standard Gamma function. In terms of placing the independent normal priors on the ontology coefficients $p(\beta_j|\tau_j,\lambda,\gamma)=\sqrt{\frac{\tau_j\pi}{2}}\lambda^{\frac{1}{\gamma}}e^{\frac{-\tau_j\lambda^{\frac{2}{\gamma}}\beta_j^2}{2}}$ and considering $g(\tau_j)$ as the hyper-prior on τ_j, we infer

$$\frac{\lambda^{\frac{1}{\gamma}}}{2\Gamma\left(1+\frac{1}{\gamma}\right)}e^{-\lambda|\beta_j|^{\gamma}}=\int_0^{\infty}p(\beta_j|\tau_j,\lambda,\gamma)g(\tau_j)d\tau_j. \tag{6}$$

Our ontology framework will be presented as follows and then approximate marginal distributions for ontology parameters can be obtained, too. The marginal likelihood of the given ontology data in Eq. (4) or in many other non-trivial patterns can't be yielded analytically. But the integral can easily be approximated in which the marginal likelihood conditional on λ and γ can be decomposed. Set $\theta=(\beta,\sigma^2,\tau)$. For fixed λ and γ, we deduce

$$\log p(\boldsymbol{y}|\lambda,\gamma)=L_{\lambda,\gamma}-KL(q||p), \tag{7}$$

where $L_{\lambda,\gamma}=\int_{\Theta}q(\boldsymbol{\theta}|\lambda,\gamma)\log\frac{p(\boldsymbol{\theta},\boldsymbol{y}|\lambda,\gamma)}{q(\boldsymbol{\theta}|\lambda,\gamma)}d\boldsymbol{\theta}$ denote the lower bound on the marginal likelihood and $KL(q||p)=\int_{\Theta}q(\boldsymbol{\theta}|\lambda,\gamma)\log\frac{p(\boldsymbol{\theta}|\boldsymbol{y},\lambda,\gamma)}{q(\boldsymbol{\theta}|\lambda,\gamma)}d\boldsymbol{\theta}$ is the Kullback–Leibler divergence between two distributions. Since $KL(q||p)$ is a strict non-negative function which equals to zero if and only if $p(\boldsymbol{\theta}|\boldsymbol{y},\lambda,\gamma)=q(\boldsymbol{\theta}|\lambda,\gamma)$, the first term $L_{\lambda,\gamma}$ in Eq. (7) forms the lower bond of $\log p(\boldsymbol{y}|\lambda,\gamma)$. Assume $q(\boldsymbol{\theta}|\lambda,\gamma)$ be the approximation of the posterior density $p(\boldsymbol{\theta}|\boldsymbol{y},\lambda,\gamma)$. We have $\max L_{\lambda,\gamma}=\min KL(q||p)$.

Let $\boldsymbol{\theta}_i$ be the subvector of $\boldsymbol{\theta}$ and $E_{i\neq j}(\cdot)$ be the expectation with respect to distributions $q_j(\boldsymbol{\theta}_j)$ with $i\neq j$. We consider the factorized expression

$$q(\boldsymbol{\theta}|\lambda,\gamma)=\prod_i q_i(\boldsymbol{\theta}_i|\lambda,\gamma). \tag{8}$$

In terms of maximizing the lower bound with respect to $q_i(\boldsymbol{\theta}_i|\lambda,\gamma)$, we infer

$$q_i(\boldsymbol{\theta}_i|\lambda,\gamma)=\frac{e^{E_{i\neq j}(\log p(\boldsymbol{y},\boldsymbol{\theta}|\lambda,\gamma))}}{\int_{\Theta_i}e^{E_{i\neq j}(\log p(\boldsymbol{y},\boldsymbol{\theta}|\lambda,\gamma))}d\boldsymbol{\theta}_i} \tag{9}$$

By virtue of the above mentioned normal mixture expression of the exponential power distribution and the solution obtained from Eq. (9), the error variance and the approximate marginal posterior distributions of ontology coefficients can be determined. However, since the mixing distribution $g(\tau_j)$ is unknown, the explicit expression for $q(\tau_j)$ can't be calculated. Fortunately, using the value of $q(\boldsymbol{\beta}|\lambda,\gamma)$, the expression of $E(\tau_j)$ can be deduced.

Let

$$\hat{\boldsymbol{\beta}}=E(\sigma^{-2})\Sigma_{\boldsymbol{\beta}}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{y},$$

$$\Sigma_{\boldsymbol{\beta}}=(\boldsymbol{V}^{\mathrm{T}}\boldsymbol{V}E(\sigma^{-2})+\boldsymbol{T})^{-1},$$

$$\boldsymbol{T}=\lambda^{\frac{2}{\gamma}}\mathrm{Diag}(E(\tau_j)),$$

$$\hat{c}=\frac{n}{2}+c_0,$$

and

$$\hat{d}=\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}-\boldsymbol{y}^{\mathrm{T}}\boldsymbol{V}E(\boldsymbol{\beta})+\frac{1}{2}\sum_{i=1}^{n}v_iE(\boldsymbol{\beta\beta}^{\mathrm{T}})v_i^{\mathrm{T}}+d.$$

The approximate marginal posterior distributions of the ontology coefficients and the error precision can be stated as follows:

$$q(\boldsymbol{\beta}|\lambda,\gamma)\overset{d}{=}N(\hat{\boldsymbol{\beta}},\Sigma_{\boldsymbol{\beta}}), \tag{10}$$

$$q(\sigma^{-2}|\lambda,\gamma)\overset{d}{=}\Xi(\hat{c},\hat{d}). \tag{11}$$

According to the fact that we don't get an available explicit form for its approximate distribution, these above presented moments are clear except for $E(\tau_j)$. We infer $E(\boldsymbol{\beta})=\hat{\boldsymbol{\beta}}$, $E(\boldsymbol{\beta\beta}^{\mathrm{T}})=\Sigma_{\beta}+E(\boldsymbol{\beta})E(\boldsymbol{\beta}^{\mathrm{T}})$, and $E(\sigma^{-2})=\frac{\hat{c}}{\hat{d}}$. In view of (9), we can express the approximate marginal distribution of τ_j as

$$q(\tau_i|\lambda,\gamma)\frac{e^{E(\log p(\beta_j|\tau_j,\lambda,\gamma))+E(\log g(\tau_j))}}{\int_0^{\infty}e^{E(\log p(\beta_j|\tau_j,\lambda,\gamma))+E(\log g(\tau_j))}d\tau_j} \tag{12}$$

In terms of Eq. (6), $q(\tau_i|\lambda,\gamma)$ evaluated at $\beta_j^2=E(\beta_j^2)$ is the normalizing constant for this term. Moreover, we deduce the following fact by derivation of both sides of Eq. (6) with respect to β_j^2 and evaluate it again at $\beta_j^2=E(\beta_j^2)$,

$$\frac{\lambda^{1-\frac{1}{\gamma}}\gamma E(\beta_j^2)^{\frac{\gamma}{2}}}{2\Gamma\left(1+\frac{1}{\gamma}\right)}e^{-\lambda E(\beta_j^2)^{\frac{\gamma}{2}}}=\int_0^{\infty}\tau_j\frac{\lambda^{\frac{1}{\gamma}}\tau_j}{\sqrt{2\pi}}e^{-\frac{1}{2}\tau_j\lambda^{\frac{2}{\gamma}}E(\beta_j^2)}g(\tau_j)d\tau_j. \tag{13}$$

Then, by normalization operation, we have

$$E(\tau_j)=\lambda^{1-\frac{2}{\gamma}}\gamma E(\beta_j^2)^{\frac{\gamma}{2}-1}. \tag{14}$$

From what we discussed above, we conclude the following facts:

$$\begin{aligned}L_{\lambda,\gamma}=&E(\log p(\boldsymbol{y}|\boldsymbol{\beta},\sigma^2))+E(\log p(\boldsymbol{\beta}|\tau,\lambda,\gamma))+E(\log p(\sigma^{-2}))\\&+E(\log g(\tau))-E(\log q(\boldsymbol{\beta}|\lambda,\gamma))-E(\log q(\sigma^{-2}|\lambda,\gamma))\\&-E(\log q(\tau|\lambda,\gamma)),\end{aligned}$$

$$E(\log p(\boldsymbol{y}|\boldsymbol{\beta},\sigma^2)) = -\frac{1}{2}\sum_{i=1}^{n} n_i(\log 2\pi - E(\log \sigma^{-2})) - E(\sigma^{-2})(\hat{d} - d_0),$$

$$E(\log p(\boldsymbol{\beta}|\tau,\lambda,\gamma)) = -\frac{p}{2}\log 2\pi + \frac{1}{2}\sum_{j=1}^{p} E(\log \tau_j) + \gamma^{-1}\log\lambda - \lambda^{\frac{2}{\gamma}}\sum_{j=1}^{p} E(\tau_j)E\left(\beta_j^2\right),$$

$$E(\log p(\sigma^{-2})) = c_0 \log d_0 + (c_0 - 1)E(\log \sigma^{-2}) - d_0 E(\sigma^{-2}) - \log \Gamma(c_0),$$

$$E(\log g(\tau)) = \sum_{j=1}^{p} E(\log g(\tau_j)),$$

$$E(\log q(\boldsymbol{\beta}|\lambda,\gamma)) = -\frac{p}{2}(\log 2\pi + 1) - \frac{1}{2}\log|\Sigma_\beta|,$$

$$E(\log q(\sigma^{-2}|\lambda,\gamma)) = \hat{c}\log\hat{d} + (\hat{c} - 1)E(\log \sigma^{-2}) - \hat{d}E(\sigma^{-2}) - \log\Gamma(\hat{c}),$$

$$E(\log q(\tau|\lambda,\gamma)) = \sum_{j=1}^{p} E(\log p(\beta_j|\tau_j,\lambda,\gamma)) + p\log 2 + \sum_{j=1}^{p} E(\log g(\tau_j)) - p\gamma^{-1}\log\lambda + p\log\Gamma\left(1 + \frac{1}{\gamma}\right) + \lambda\sum_{j=1}^{p} E\left(\beta_j^2\right)^{\frac{\gamma}{2}}.$$

Based on simplifications, we have

$$L_{\lambda,\gamma} = -\frac{n}{2}\log(2\pi) + \frac{p}{2} + \frac{1}{2}\log|\Sigma_\beta| + c_0\log d_0 - \hat{c}\log\hat{d} - \log\Gamma(c_0) + \log\Gamma(\hat{c}) + p\gamma^{-1}\log\lambda - p\log 2 - p\log\Gamma\left(1 + \frac{1}{\gamma}\right) - \lambda\sum_{j=1}^{p} E\left(\beta_j^2\right)^{\frac{\gamma}{2}}. \quad (15)$$

Finally, the iterative produce for ontology coefficients can be stated as

$$E_{k+1}(\boldsymbol{\beta}) = \arg\min_{\boldsymbol{\beta}}(\boldsymbol{y} - \boldsymbol{V}\boldsymbol{\beta})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{V}\boldsymbol{\beta}) + \lambda\gamma E_k(\sigma^{-2})^{-1}\sum_{j=1}^{p}\beta_j^2 E_k\left(\beta_j^2\right)^{\frac{\gamma}{2}-1}, \quad (16)$$

where $E_k(\cdot)$ denotes the expectation determined at the k-th iteration.

4. Simulation studies

In this section, two simulation experiments related to ontology similarity measure are presented. In order to be close to the setting of ontology algorithm, we choose a vector with p dimension to express each vertex's information. All the information of name, instance, attribute and structure of vertex is contained in the vector. Here the instance of vertex means the set of its reachable vertex in the directed (or, undirected) ontology graph.

In order to make comparisons more accurate, the main algorithm runs in C++, in view of available LAPACK and BLAS libraries for linear algebra computations. All experiments are taken on a double-core CPU with memory of 8 GB.

4.1. Experiment on biology data

In the first experiment, we choose "Go" ontology O_1 whose basic structure can be shown in http://www.geneontology.org

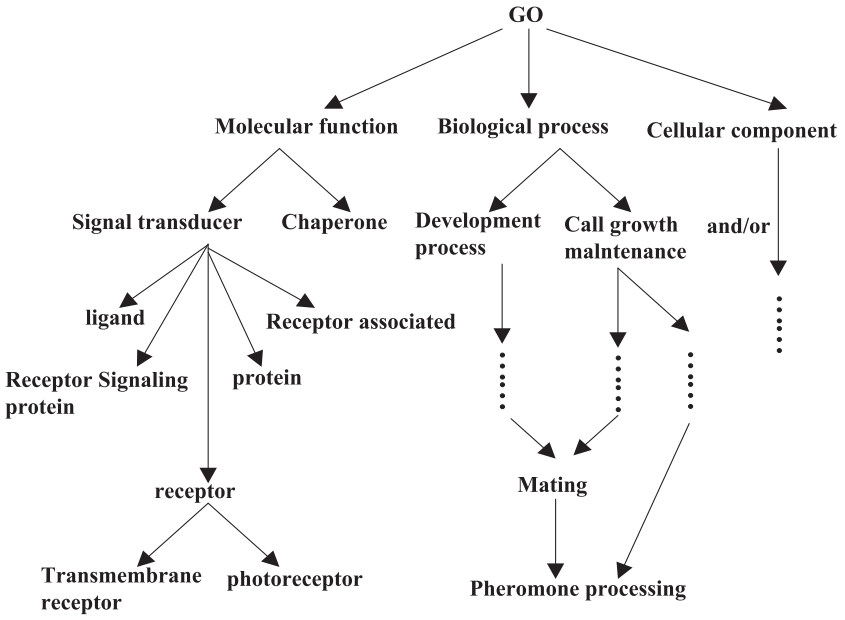

Figure 1 "Go" ontology.

Table 1 Experiment data for ontology similarity measure.

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio	$P@20$ average precision ratio
Our Algorithm	56.49%	68.27%	81.24%	93.71%
Algorithm in Gao et al. (2013)	56.46%	67.72%	78.38%	79.39%
Algorithm in Gao and Shi (2013)	56.44%	65.73%	78.39%	89.72%
Algorithm in Gao et al. (2016)	49.87%	63.64%	76.02%	85.46%

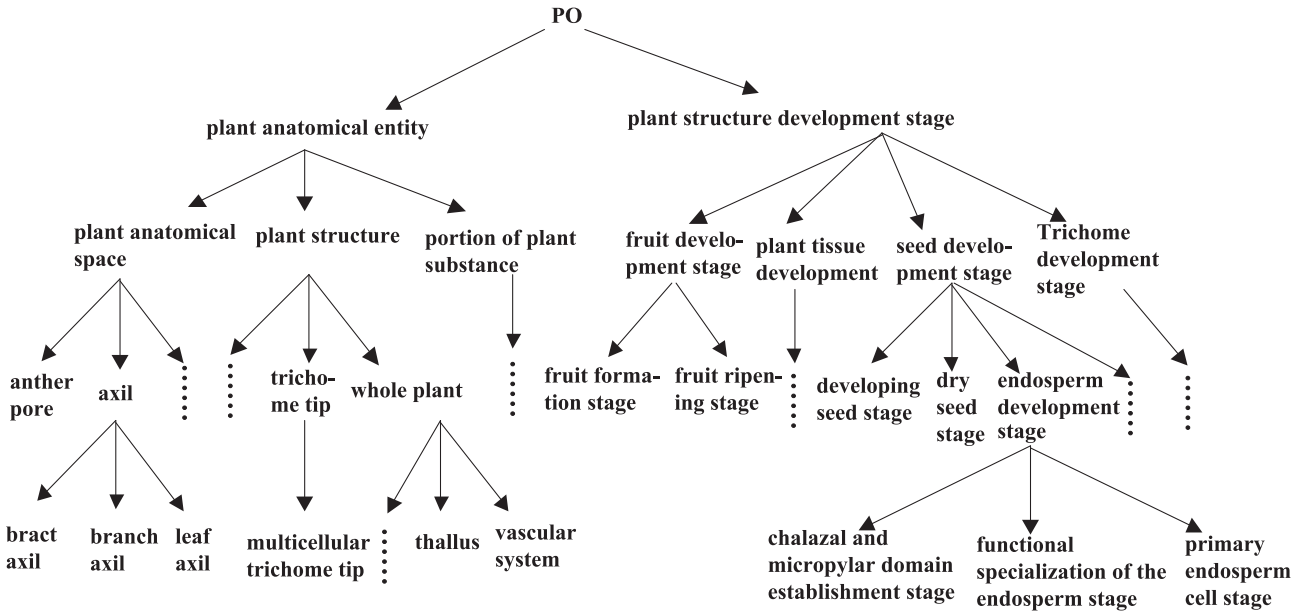

Figure 2 "PO" ontology O_2.

Table 2 Experiment data for ontology similarity measure.

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio	$P@20$ average precision ratio
Our Algorithm	53.60%	66.64%	90.04%	96.73%
Algorithm in Gao et al. (2013)	36.63%	44.60%	58.45%	70.06%
Algorithm in Gao and Shi (2013)	36.96%	45.08%	60.17%	73.99%
Algorithm in Gao et al. (2016)	53.58%	65.17%	88.21%	93.85%

(Fig. 1 shows the basic structure of O_1). $P@N$ (Precision Ratio, see Craswell and Hawking (2003) for more detail) is traditionally used to judge the equality of the experiment, and we also choose it for its efficiency. At first, the experts give the closest N concepts for every vertex on the ontology graph. Then using the algorithm, we compute the precision ratio, so that we can get the first N concepts for every vertex on ontology graph. Ontology algorithms in Gao et al. (2013, 2016) and Gao and Shi (2013) are also applied into "Go" ontology. At last, the precision ratio obtained from the four methods is gotten and given in Table 1.

From the data in Table 1, we can find that when $N = 3$, 5, 10 or 20, the precision ratio obtained from our algorithm is higher than that obtained by algorithms proposed in

Gao et al. (2013, 2016) and Gao and Shi (2013). Particularly, such precision ratios are increasing apparently with N increasing. Thus, our algorithm is better than the method presented by Gao et al. (2013, 2016) and Gao and Shi (2013).

4.2. Experiment on plant data

In this subsection, we use "PO" ontology O_2, whose structure is presented in http://www.plantontology.org. (Fig. 2 shows the basic structure of O_2), to check the efficiency of our new algorithm in ontology similarity measuring. Similarly, we use the $P@N$ again for this experiment. Moreover, the ontology methods in Gao et al. (2013, 2016) and Gao and Shi (2013) are applied to the "PO" ontology. We calculate the data using the three algorithms, and then we compare the results with that gotten from the new algorithm. Part of the data can be referred to Table 2.

From the data in Table 2, we can find that when $N = 3$, 5, 10 or 20, the precision ratio gotten from our algorithm is higher than that from algorithms proposed in Gao et al. (2013, 2016) and Gao and Shi (2013). Particularly, such precision ratios are increasing apparently with N increasing. Thus, our algorithm is better and more effective than the method presented by Gao et al. (2013, 2016) and Gao and Shi (2013).

5. Conclusions

Borrowed the marginal technology for ontology sparse vector computation in this paper, we proposed a new computation algorithm on the basis of the marginal distribution and the analysis of the convergence criterion problem. The simulation data obtained from the experiments shows the high efficiency of our newly proposed algorithm in biology and plant science. Hence, the ontology sparse algorithm sees the promising application prospects for biology science.

Acknowledgments

We thank the reviewers for their constructive comments in improving the quality of this paper. This work was supported in part by NSFC (Nos. 11401519 and 61262070).

References

Auffeves, A., Grangier, P., 2016. Contexts, systems and modalities: a new ontology for quantum mechanics. Found. Phys. 46, 121–137.

Agapito, G., Milano, M., Guzzi, P.H., Cannataro, M., 2016. Extracting cross-ontology weighted association rules from gene ontology annotations. IEEE-ACM Trans. Comput. Biol. 13, 197–208.

Bajenaru, L., Smeureanu, I., Balog, A., 2016. An ontology-based e-learning framework for healthcare human resource management. Stud. Inform. Control 25, 99–108.

Bobillo, F., Straccia, U., 2016. The fuzzy ontology reasoner fuzzyDL. Knowl.-Based Syst. 95, 12–34.

Brown, R.B.K., Beydoun, G., Low, G., Tibben, W., Zamani, R., Garcia-Sanchez, F., Martinez-Bejar, R., 2016. Computationally efficient ontology selection in software requirement planning. Inform. Syst. Front. 18, 349–358.

Carmen Suarez-Figuero, M., Gomez-Perez, A., Fernandez-Lopez, M., 2016. Scheduling ontology development projects. Data Knowl. Eng. 102, 1–21.

Chicco, D., Masseroli, M., 2016. Ontology-based prediction and prioritization of gene functional annotations. IEEE-ACM Trans. Comput. Biol. 13, 248–260.

Craswell, N., Hawking, D., 2003. Overview of the TREC 2003 web track. In: Proceeding of the Twelfth Text Retrieval Conference, Gaithersburg, Maryland, NIST Special Publication 78–92.

Farid, H., Khan, S., Javed, M.Y., 2016. DSont: DSpace to ontology transformation. J. Inf. Sci. 42, 179–199.

Forsati, R., Shamsfard, M., 2016. Symbiosis of evolutionary and combinatorial ontology mapping approaches. Inform. Sci. 342, 53–80.

Gao, W., Gao, Y., Liang, L., 2013. Diffusion and harmonic analysis on hypergraph and application in ontology similarity measure and ontology mapping. J. Chem. Pharm. Res. 5, 592–598.

Gao, W., Gao, Y., Zhang, Y., 2012. Strong and weak stability of k-partite ranking algorithm. Information 15, 4585–4590.

Gao, W., Shi, L., 2013. Ontology similarity measure algorithm with operational cost and application in biology science. BioTechnol. An Indian J. 8, 1572–1577.

Gao, W., Xu, T. W., 2013. Stability analysis of learning algorithms for ontology similarity computation. Abstr. Appl. Anal. http://dx.doi.org/10.1155/2013/174802.

Gao, W., Zhu, L.L., Wang, K.Y., 2015. Ontology sparse vector learning algorithm for ontology similarity measuring and ontology mapping via ADAL technology. Int. J. Bifurcat. Chaos 25. http://dx.doi.org/10.1142/S0218127415400349.

Gao, W., Zhu, L. L., 2014. Gradient learning algorithms for ontology computing. Comput. Intell. Neurosci. http://dx.doi.org/10.1155/2014/438291.

Gao, Y., Farahani, M.R., Gao, W., 2016. Ontology optimization tactics via distance calculating. Appl. Math. Nonlin. Sci. 1, 159–174.

Gregor, D., Toral, S., Ariza, T., Barrero, F., Gregor, R., Rodas, J., Arzamendia, M., 2016. A methodology for structured ontology construction applied to intelligent transportation systems. Comput. Stand. Interf. 47, 108–119.

Hoyle, D.C., Brass, A., 2016. Statistical mechanics of ontology based annotations. Physica A 442, 284–299.

Huntley, R.P., Sitnikov, D., Orlic-Milacic, M., et al, 2016. Guidelines for the functional annotation of microRNAs using the Gene Ontology. RNA 22, 667–676.

Ishibuchi, H., Nojima, Y., 2013. Repeated double cross-validation for choosing a single solution in evolutionary multi-objective fuzzy classifier design. Knowl.-based Syst. 54, 22–31.

Kaminski, M., Nenov, Y., Grau, B.C., 2016. Datalog rewritability of disjunctive datalog programs and non-horn ontologies. Artif. Intell. 236, 90–118.

Mancinelli, G., Vizzini, S., Mazzola, A., Maci, S., Basset, A., 2013. Cross-validation of delta N-15 and FishBase estimates of fish trophic position in a Mediterranean lagoon: the importance of the isotopic baseline. Estuar. Coast. Shelf Sci. 135, 77–85.

Mazandu, G.K., Chimusa, E.R., Mbiyavanga, M., Mulder, N.J., 2016. A-DaGO-Fun: an adaptable gene ontology semantic similarity-based functional analysis tool. Bioinformatics 32, 477–479.

Mukhopadhyay, S., Bhattacharya, S., 2013. Cross-validation based assessment of a new Bayesian palaeoclimate model. Environmetrics 24, 550–568.

Palmer, C., Urwin, E.N., Pinazo-Sanchez, J.M., Cid, F.S., Rodriguez, E.P., Pajkovska-Goceva, S., Young, R.I.M., 2016. Reference ontologies to support the development of global production network systems. Comput. Ind. 77, 48–60.

Pesaranghader, A., Matwin, S., Sokolova, M., Beiko, R.G., 2016. SimDEF: definition-based semantic similarity measure of gene ontology terms for functional similarity analysis of genes. Bioinformatics 32, 1380–1387.

Sherimon, P.C., Krishnan, R., 2016. OntoDiabetic: an ontology-based clinical decision support system for diabetic patients. Arab. J. Sci. Eng. 41, 1145–1160.

Terblanche, C., Wongthongtham, P., 2016. Ontology-based employer demand management. Software Pract. Exper. 46, 469–492.

Trokanas, N., Cecelja, F., 2016. Ontology evaluation for reuse in the domain of process systems engineering. Comput. Chem. Eng. 85, 177–187.

Umadevi, S., Premkumar, K., Valarmathi, S., Ayyasamy, P.M., Rajakumar, S., 2016. Identification of novel genes related to diabetic retinopathy using protein-protein interaction network and gene ontologies. J. Biol. Syst. 24, 117–127.

Varmuza, K., Filzmoser, P., Hilchenbach, M., Krüger, H., Silén, J., 2014. KNN classification-evaluated by repeated double cross validation: Recognition of minerals relevant for comet dust. Chemometr. Intell. Lab. Syst. 138, 64–71.