# Evaluation Metrics for Ontology Complexity and Evolution Analysis*

Zhe YANG[1, 2] Dalu Zhang[1] Chuan YE[1]
[1] Department of Computer Science and Technology, Tongji University, Shanghai 200092, China
[2] School of Computer Science & Technology, Soochow University, Suzhou 215006, China
*hatasen@hotmail.com, daluz@ieee.org, shtjjsjyjx@hotmail.com*

## Abstract

*With the tremendous development in size, the complexity of ontology increases. Thus ontology evaluation becomes extremely important for developers to determine the fundamental characteristics of ontologies in order to improve the quality, estimate cost and reduce future maintenance. Our research examines the concepts and their hierarchy in conceptual model, the common feature of the most ontologies, which reflects the fundamental complexity. We suggest a well-defined metrics suite of complexity, which mainly examine the quantity, ratio and correlativity of concepts and relationships, to evaluate ontologies from the viewpoint of complexity and its evolution. In the study, we measure three ontologies in GO to verify our metrics. The results indicate that this metrics suite works well, and the biological process ontology is the most complex one from the view of complexity, and the molecular function ontology is the unsteadiest one from the view of evolution.*

## 1. Introduction

With rising importance of knowledge exchange, ontologies have become a key technology to provide shared knowledge models to semantic-driven applications. In the last decade there has been very active research in ontology engineering. The majority of research is focused on construction issues. Neither the ontology construction itself nor its development is a single-person enterprise. Large standardized ontologies are often developed by several researchers in parallel (e.g. GO [1] and SUO [2]); a number of ontologies grow in the context of peer-to-peer applications (e.g. Edutella [3]); other ontologies are constructed dynamically [4].

Ontology construction and development are necessarily an iterative, dynamic and parallel process [5]. The complexity of ontology itself increases as ontology grows in size. And as ontology evolving, the complexity of management and maintenance increases. It becomes important to determine fundamental characteristics of ontologies [6]. Although it is an essential requirement for useful ontologies, there are still very few commonly agreed methodologies and metrics for analyzing ontology complexity and evaluating its evolution [5, 7].

The ontology complexity should be managed both in engineering and at the display level to make the ontologies more tractable. Thus, ontology metrics are expected to give some insight for developers to help them design ontologies, improve quality and reduce maintenance requirements, as well as help users to choose ontology which meet their needs best.

In this paper, we specifically focus on the problem that ontology complexity has to remain tractable during evolution. We analyze the concepts and their hierarchy in conceptual model, the common feature of the most ontologies, which reflects the fundamental complexity. And we present a well-defined complexity metrics suite, which mainly examine the quantity, ratio and correlativity of concepts and relationships. Then we measure the Gene Ontology with these metrics to analyze the complexity and its evolution trend.

The rest of this paper is structured as follows: section 2 is some related works about ontology evaluation and metrics. Section 3 introduces our research objectives and some considerations. Section 4 is the proposed metrics, includes the primitive and complex ones. In section 5, the complexity analysis results of Gene Ontology are given to show its evolution trend. Section 6 is some conclusions and outlook for future works.

* Supported by National Natural Science Foundation of China under Grant No. 90204010

## 2. Related works

As ontology development becomes a more ubiquitous and collaborative process, ontology evolution and versioning becomes an important area of ontology research. The variety of causes and consequences of the ontology changes makes ontology evolution a very complex operation that should be considered as both an organizational and a technical process [8].

The approach for ontology evolution given in [9] used an ontology to specify the semantics of possible changes of a knowledge base. Their six-phase evolution model checks the ontology after changes are made with the possibility to take back these changes. Noy and Klein are working on an ontology evolution system by extracting the operations from two versions of one ontology [10]. They developed a framework [11] for managing ontology evolution by extracting the operations leads from one ontology version to another. Like [9] they are using an ontology for specifying change operations.

In contrast to the ontology evolution that allows access to all data only through the newest ontology, ontology versioning allows access to data through different versions of the ontology. Thus, ontology evolution can be treated as a part of the ontology versioning mechanism that is analyzed in [12]. Authors provide an overview of causes and consequences of the changes in the ontology. However, the most important flaw is the lack of a detailed analysis of the effect of specific changes.

Oliver et al. [13] discuss the kinds of changes that occur in medical ontologies and propose the CONCORDIA concept model to cope with these changes. The main aspects of CONCORDIA are that all concepts have a permanent unique identifier. Concepts are given a retired status instead of being physically deleted. Moreover special links are maintained to track the retired parents and children of each concept. However, this approach is insufficient for managing a change on the Semantic Web especially while there are no possibilities to control the whole process.

In contrast to the above researches of ontology evolution and versioning, only little empirical work has focused on ontology evaluation [14]. Ontology metrics is a desirable in ontology evaluation. Although some metrics have been suggested [15], more work is needed [14]. The most existing metrics are proposed to evaluate the syntactical, semantic, structural features of ontology conceptual model. There are few metrics investigating the ontology complexity and evolution.

Burton et al. assessed the effectiveness of the DAML ontologies [16]. They suggested an ontology auditor metrics suite, and mainly considered the syntactic, semantic, pragmatic and social quality of ontologies. Each metric has two or three attributes. And they used overall quality as the total quality metric, which is a weighted function of four constituents above.

Literature [17] proposed a set of ontology cohesion metrics to measure the modular relatedness of OWL ontologies. These metrics is focused on the number of classes and depth of inheritance tree of all classes. The metrics are collected by a standard XML DOM parser that parses the XML-based OWL ontology syntactically. And it computes cohesion metrics conceptually based on predefined OWL primitives, which explicitly defined tree-based semantic hierarchies in OWL ontologies. These metrics are theoretically validated using standard metrics validation frameworks. And they are empirically validated by comparing them statistically to assessments performed by a human team of evaluators.

In literature [18], authors use weighted class dependence graphs to represent a given class diagrams, and then present a structure complexity measure for the UML class diagrams based on entropy distance. It considers complexity of both classes and relationships between the classes, and presents rules for transforming complexity value of classes and different kinds of relations into a weighted class dependence graphs. This method can measure the structure complexity of class diagrams objectively.

Idris studied two conceptual integrity metrics based on graph theory is his PhD thesis [19], which are conceptual coherence and conceptual complexity. Conceptual coherence uses average distance between nodes in a graph to measure the interrelatedness of concepts. And conceptual complexity reflects the average number of relationships per node with the average degree across all nodes in a graph. Nevertheless, these metrics take much more focus on characteristics of single concept.

Chris Mungall researched the increased complexity of Gene Ontology [20]. He measured the average number of paths-to-top of a term and used the path-to-term ratio to measure of complexity in an ontology, which is represented in DAG (directed acyclic graph). However, in the calculation of the total number of terms, the obsolete terms does not be eliminated. While calculating the paths-to-top of terms, the paths of these obsolete terms are not counted. Thus the result of path-to-term ratio is not correct. Furthermore, this metric is very simple, and only reflects one feature of ontology complexity.

## 3. Objectives and Approaches

### 3.1. Objectives

All mentioned methodologies and metrics above belong to different domain and have their own emphases. Most of them evaluate the ontology from the viewpoints of semantic, syntactic and structure. And they use different description methods to meet their special needs. Anyway, they put the emphasis mainly on the description ability of ontology and have little help to complexity evaluation. So in this study, the research objectives include:

1. Provide a set of metrics for evaluating ontologies from the viewpoint of complexity.
2. Make developers clearly realize the size and scale of ontology, and ensure them whether the growth rate is at an appreciate speed?
3. Estimate the cost and future workload when evolve and maintain the ontology.
4. Help developers to decide whether the ontology is over complex that it needs some simplifications or building a lite version when used in practice?

### 3.2. Approaches

Although, different researchers have different definitions and implementations of ontologies, there is an important common feature. That is, whatever domains these ontologies are designed for and however they are defined, developers must build the conceptual model for domain knowledge at first, which is a set of concepts and relations that reflects the concepts hierarchy. By analyzing the concepts and relations, we can have a glimpse at the ontology complexity and its evolution.

### 3.3. Common Formal Notation

We use the following formal notation to represent some terms defined in the ontology conceptual model. And small letters are used to identify the notations related to concepts and relations, while capital letters are used to identify the terminology related to ontology and metrics.

$C=\{c_1,c_2,\cdots,c_m\}$ : the set of $m$ concepts defined in an ontology explicitly. In other ontologies, concept may be named as “class” or “term”.

$R=\{r_1,r_2,\cdots,r_m\}$ : the set of number of relations each concept has. In other ontologies, relation may be named as “slot”. It only includes those inherited relations that reflect the hierarchy of concepts, such as “is a”, “part of”, etc.

In ontology conceptual model, concepts hierarchy is typically expressed in DAG (directed acyclic graph) showed in Figure 1. Each node represents a concept and each directed arc represents relation to present the hierarchical structure between concepts in ontologies.

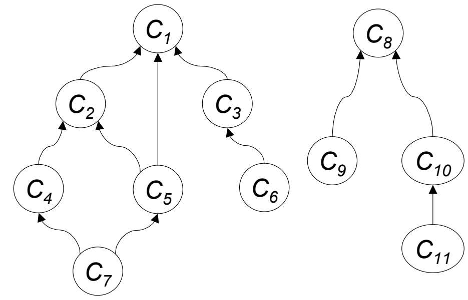


**Figure 1. Ontology in DAG**

In DAG, path is a distinct trace that can be taken from a specific particular concept to the most general concept in the ontology, which is the concept without any parent or superclass (e.g. $c_7$-$c_5$-$c_2$-$c_1$ in figure 1). Let $P=\{p_1,p_2\cdots p_m\}$ be the set of number of paths each concept has. In figure 1, there are 11 concepts ($m$=11), two general concepts ($c_1$ and $c_8$).

Different path has its own length, thus the path length is defined as the sum of relations on a path. Let $pl_i=\{pl_{i,1},pl_{i,2},\cdots,pl_{i,p_i}\}$ be the set of path length of concept $c_i$. Path length of a particular concept indicates the semantic distance between the concept and the general concept. So the set of path length of all concepts in ontology is defined as $PL=\{\{pl_{1,1},\cdots pl_{1,p_1}\},\cdots,\{pl_{m,1},\cdots pl_{m,p_m}\}\}$.

## 4. Proposed Metrics

### 4.1. Primitive Metrics

*TNOC* (Total Number of Concepts): is the sum of concepts in the set $C$. $TNOC=|C|=m$.

*TNOR* (Total Number of Relations): is the sum of relations of each concept. $TNOR=\sum_{i=1}^{m} r_i$ .

*TNOP* (Total Number of Paths): is the sum of paths of each concept. $TNOP=\sum_{i=1}^{m} p_i$ .

$\lambda_i$ : is the longest path length of concept $c_i$. $\lambda_i = max(pl_{i,k}), 1\le k\le p_i$ .

$\overline{\lambda_i}$ : is the average path length of concept $c_i$. $\overline{\lambda_i} = \sum_{k=1}^{p_i} pl_{i,k} \Big/ p_i$ .

The above two metrics measure the semantic distance from general concept to particular ones.

$\Lambda$ : is the max path length of ontology. It is the longest $\lambda_i$ . $\Lambda = max(\lambda_i), 1 \le i \le m$ .

$\overline{\Lambda}$ : is the average path length of ontology. $\overline{\Lambda}$ = $\sum_{i=1,j=1}^{m,p_m} pl_{i,j} \Big/ TNOP$ = $\sum_{i=1,j=1}^{m,p_m} pl_{i,j} \Big/ \sum_{k=1}^{m} p_k$ . (Notice that it is not the same with $\sum_{i=1}^{m} \lambda_i \Big/ m$ .)

These two metrics indicate the radius of the ontology in DAG, and the extension of the general concept. Actually they mainly measure the semantic scope covered by the ontology.

### 4.2. Complexity Metrics

$\mu$ : the average relations per concept. It is the ratio of *TNOR* to *TNOC*. $\mu = TNOR / TNOC = \sum_{i=1}^{m} r_i \Big/ m$ . It indicates the average connectivity degree of a concept.

$\rho$ : the average paths per concept. It is the ratio of *TNOP* to *TNOC*. $\rho = TNOP / TNOC = \sum_{i=1}^{m} p_i \Big/ m$ . For any ontology, $\rho$ must be greater than or equal to 1 (each concept must have a parent except for the general concept). If $\rho$ =1, then the ontology is a tree (each concept has a single parent, and thus a single path to the most general concept). Multi-relation concepts (higher $\mu$ ratio) result in higher $\rho$ ratio for an ontology. This is illustrated in table 1 (see section 5.4) by comparing $\mu$ with $\rho$ for biological process ( $\mu$ =1.67, $\rho$ =24.15 as of Jun. 2005), cellular component (1.42, 5.31 as of the same date) and molecular function (1.18, 1.60 as of the same date).

$\sigma$ : is the ratio of max path length to average path length of the ontology, $\sigma = \Lambda / \overline{\Lambda}$ . This metric examines the concept aggregation and coherence of ontology.

## 5. Experimental Results and Conclusions

We measured the growing complexity of Gene Ontology [1] with the above metrics since we began archiving the ontologies from DEC.2002 to Jun. 2005. GO has three organizing ontologies: BP (biological process), CC (cellular component) and MF (molecular function). The graphs below illustrate the complexity evolution if the three GO ontologies over time.

### 5.1. Biological Process

Figure 2 is the quantity evolution of concepts (*TNOC*), relations (*TNOR*) and paths (*TNOP*) of BP ontology. The left Y-axis shows the increase of concepts and relations. The right Y-axis shows the increase of paths. In this figure, the lines of concepts and relations indicate they increased at a steady but slow rate. The average monthly increase rates are 1.17% and 1.44% respectively. The line of paths indicates that it has a rapid growth in quantity. The average monthly increase rate reaches 8.75%. Moreover, before DEC. 2004, paths increased relative steadily in most time. While after that, it has enormous ladderlike increase, such as at time of DEC. 2004, Jan. 2005 and Apr. 2005.

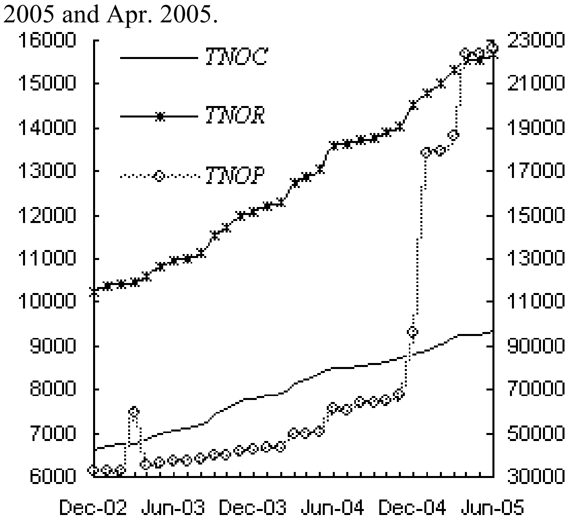


**Figure 2. The evolution of concepts, relations and paths (BP)**

Figure 3 is the evolution of two ratios, $\mu$ (Relation-to-Concept) and $\rho$ (Path-to-Concept). The left Y-axis shows the increase of $\mu$ . The right Y-axis shows the increase of $\rho$ . In this figure, the line of $\mu$ shows that increment of average relations of each concept is very little; the average monthly increase rate is only 0.26%. While the line of $\rho$ indicates the average paths of each concept increased enormously, and the average monthly increase rate is 7.51%. Furthermore, it has the same evolution trend with the line of Paths in figure 2.

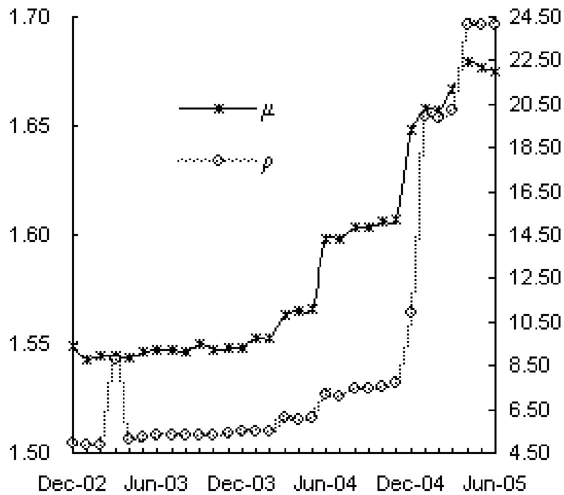

**Figure 3. $\mu$ vs. $\rho$ (BP)**

If compare above figures carefully, we can find that they look like almost the same in some degree if we ignore the line of Concepts in figure 2. Actually, the line of $\mu$ in figure 3 also has the same evolution trend with the line of Relations in figure 2, which is leap increase in quantity over time. This trend is not so obviously in figure 2 only because of the numerical range on left Y-axis. While in figure 3, it is magnified.

And if we examine the leap points on two lines in figure 3 by the time, we can conclude that the $\mu$ and $\rho$ increased synchronously. Because path consists of relations, the increase of $\mu$ means the ontology in DAG becomes more anfractuous and complex, the $\rho$ will grow rapidly and enormously. This feature can be also observed on the lines of Relations and Paths in figure 2, though not markedly. So it is concluded that the complexity of ontology can be indicated by the metrics of $\mu$ and $\rho$ explicitly.

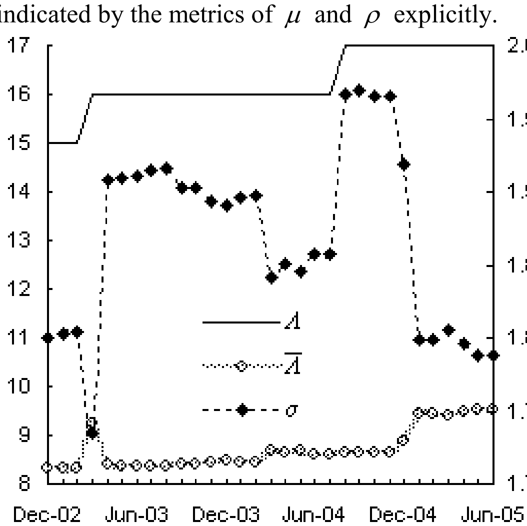

**Figure 4. Path length analysis (BP)**

Figure 4 illustrates the evolution of $\Lambda$ (max path length), $\overline{\Lambda}$ (average path length) and the $\sigma$ ratio between them. The left Y-axis shows the increase of $\Lambda$ and $\overline{\Lambda}$. The right Y-axis indicates the increase of $\sigma$ ratio. From the line of $\Lambda$, it shows an inerratic increase of max path length with two leap points over time. The leap point indicates the ontology extends its knowledge cover further with a “feeler”.

After examining the line of $\sigma$, it is concluded the following results. First, all values of this metric are less than 2, which means most concepts tightly surround the general concept or the core. Professionally, the concept aggregation or coherence is high. Second, when the $\Lambda$ increases, the $\sigma$ metric will have a leaping increment synchronously. After that, when $\Lambda$ remains while $\overline{\Lambda}$ increases, so the $\sigma$ ratio will decrease until the $\Lambda$ increases next time. Moreover, the faster the $\overline{\Lambda}$ increases, the faster the $\sigma$ decreases. So according to the up or down of the line $\sigma$, we can qualitative analysis the variety of $\Lambda$ and $\overline{\Lambda}$ with only one metric.

### 5.2. Cellular Component

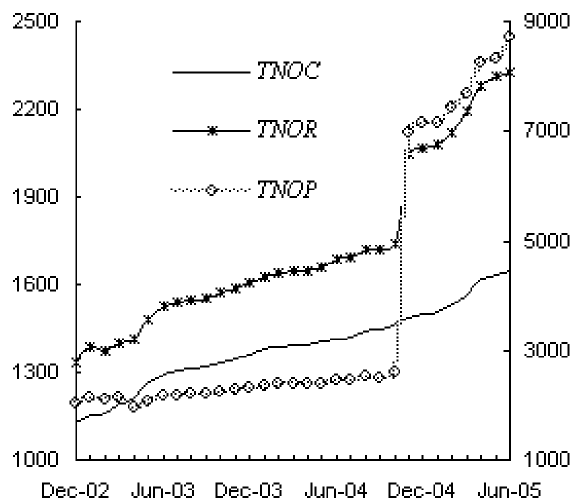

**Figure 5. The evolution of concepts, relations and paths (CC)**

Figure 5 illustrates the evolution of *TNOC*, *TNOR* and *TNOP* of cellular component ontology. All the configurations are the same as figure 2. So does the figure 8 below. In this figure, it is observed that the evolutions of concepts, relations and paths are almost the same as the figure 2. And the monthly average increase rates are 1.28%, 1.92% and 7.24% respectively, which are numerically close to the values of BP. But, the increase of paths is smoother than BP relatively. It has a big rise at Nov. 2004 only.

Figure 6 is the evolution of $\mu$ and $\rho$ . All the configurations are the same as figure 3. So does the figure 9 below. Like the figure 3, it shows the complexity of CC clearly and magnifies the fluctuation trend of lines (Relations and Paths) in figure 5. The monthly average increase rates of $\mu$ and $\rho$ are 0.63% and 5.88%.

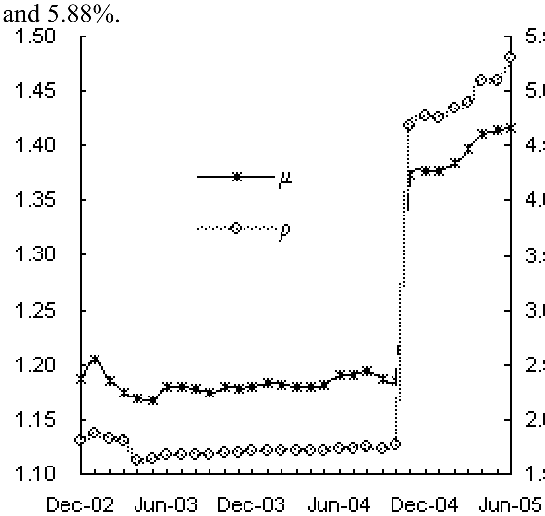


**Figure 6. $\mu$ vs. $\rho$ (CC)**

Figure 7 analyzes the evolution of $\Lambda$ , $\overline{\Lambda}$ and the $\sigma$ ratio. All the configurations are the same as figure 4. So does the figure 10 below. The evolution trends of $\Lambda$ and $\overline{\Lambda}$ are similar to BP ontology in figure 4. The average monthly increase rate of $\overline{\Lambda}$ is 0.47%. However, the speed of $\Lambda$ increase is faster than $\overline{\Lambda}$ . The biggest difference lies in the values of $\sigma$ ratio, most of which are over 2. It means the concept organization and aggregation of CC ontology are loose relatively.

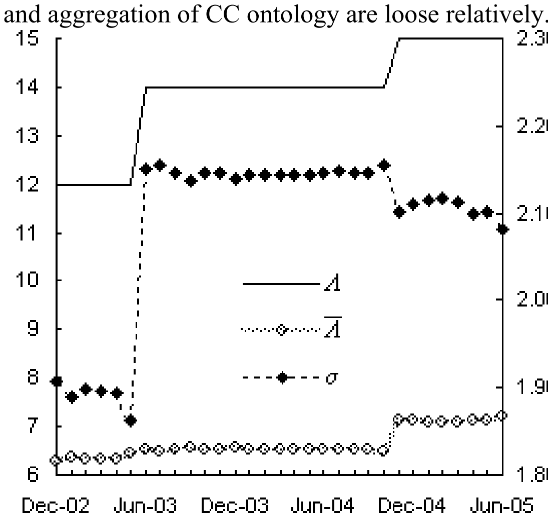


**Figure 7. Path length analysis (CC)**

### 5.3. Molecular Function

Figure 8 illustrates the evolution of *TNOC*, *TNOR* and *TNOP* of molecular function ontology. In this figure, it is observed that concepts and relations increase steadily except for the third quarter of 2003. However the quantity of paths fluctuated a lot over time. The average monthly increase rates of concepts, relations and paths are 1.07%, 0.91% and 0.79%.

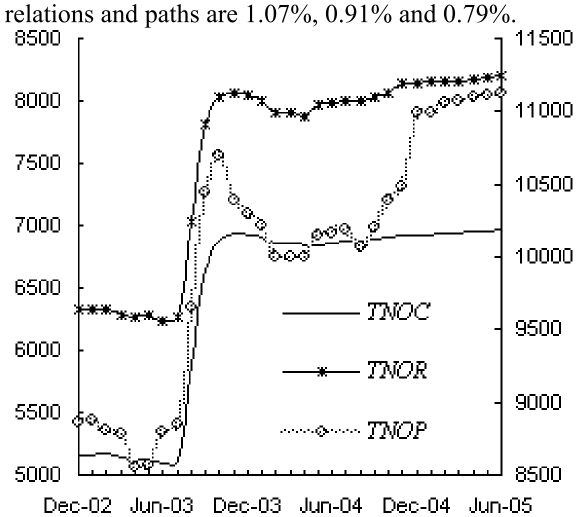


**Figure 8. The evolution of concepts, relations and paths (MF)**

Figure 9 is the contrast of $\mu$ and $\rho$ . These two metrics clearly indicate the fluctuation trend of relations and paths each concept has over time. Overall, the complexity of MF is less than the beginning, and it fluctuated wider and rapidly than BP and CC ontologies in some degree. So the average monthly increase rates of $\mu$ and $\rho$ are -0.14% and -0.23%.

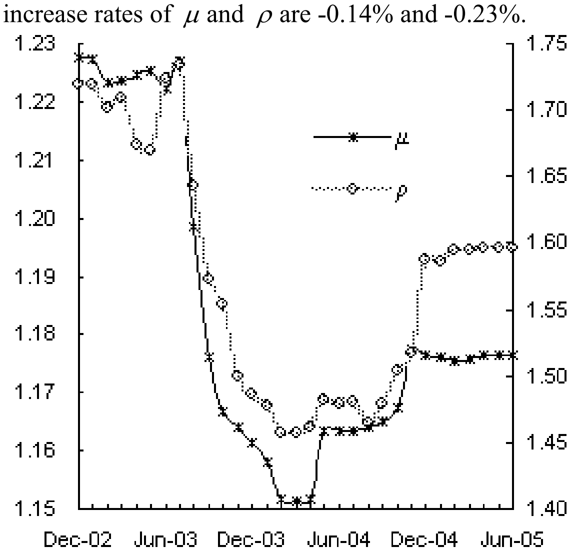


**Figure 9. $\mu$ vs. $\rho$ (MF)**

Figure 10 analyzes the evolution of $\Lambda$ , $\overline{\Lambda}$ and the $\sigma$ ratio. In the figure, it is observed that the $\overline{\Lambda}$ is almost changeless, even a little decrease. The average monthly increase rate of $\overline{\Lambda}$ is -0.02%. The evolution of $\Lambda$ is much more different from the BP and CC ontologies. As we observed in figure 4 and 7, the line of $\Lambda$ increases with time. But in this figure, it fluctuates up and down over time. These lead to the same fluctuation trend of $\sigma$ ratio. And most values of the ratio are over 2, which means the concept organization and aggregation of MF are also loose like CC ontology.

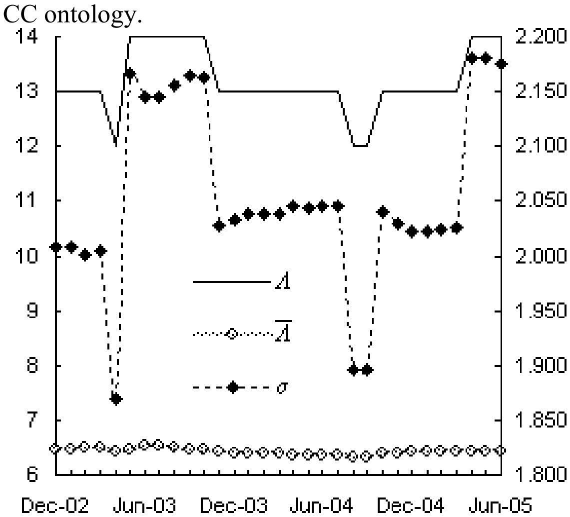


**Figure 10. Path length analysis (MF)**

### 5.4. Compare and Contrast

After analyze the complexity of biological process, cellular component and molecular function ontologies in GO, we get the further conclusions as below.

The first graph (figure 2, 5, 8) of each ontology is an overview of complexity. It outlines the evolution trend of *TNOC*, *TNOR* and *TNOP*, the three primitive metrics of ontology. And with these metrics, we learn that the BP and CC ontologies share some similar evolution characteristics: steadily increase in most time with some big rises (presumably there was some large reorganization at these times). But the MF ontology fluctuated widely during the whole development process. It reflects that the developers do not understand the domain knowledge very well. So iterative development is unavoidable.

The second graph (figure 3, 6, 9) of each ontology is a detailed analysis of two ratios, $\mu$ and $\rho$ , the two complexity metrics. Because of division by the number of concepts, these two ratios eliminate interference factor introduced by the increment of concepts. They reveal the fluctuations of relation and path more precisely than the first graph. In these graphs, we can clearly observe that the path related metrics ( $\mu$ ,*TNOP*) varies with relation related metrics ( $\rho$ ,*TNOR*) synchronously. Because of the increase of relations, every concept has more shortcuts or choices to form paths to the general concept.

In the third graph (figure 4, 7, 9) of each ontology, we examine the metrics related to path length, that are $\Lambda$ , $\overline{\Lambda}$ , and the $\sigma$ ratio between them. The first two metrics reveal the knowledge cover of ontology. All the three ontologies extend their knowledge cover regularly and steadily except for the $\Lambda$ of MF ontology. It fluctuated irregularly over time, which indicates the developers are unfamiliar with the domain knowledge form another side. The ratio of $\sigma$ indicates the concept aggregation and coherence of ontology. From the three graphs, we learn that the concept organization and aggregation of BP is higher than CC and MF.

**Table 1. Complexity contrast**

| Jun.2005 | BP | CC | MF |
|---|---|---|---|
| *TNOC* | 9367 | 1644 | 6972 |
| *TNOR* | 15689 | 2329 | 8202 |
| *TNOP* | 226201 | 8722 | 11126 |
| $\mu$ | 1.67 | 1.42 | 1.18 |
| $\rho$ | 24.15 | 5.31 | 1.60 |
| $\Lambda$ | 17 | 15 | 14 |
| $\overline{\Lambda}$ | 9.507 | 7.205 | 6.436 |
| $\sigma$ | 1.788 | 2.082 | 2.175 |

According to all 9 graphs, we draw the conclusion that the MF is the unsteadier than BP and CC from the view of evolution. And if viewed from the numerical value, the BP is more complex than CC and MF. For example, we list some statistics at a static time slice (Jun. 2005) in table 1 above.

## 6. Summary and Future Works

With the rapid growth in popularity and size, the complexity of ontology increases tremendously. So it becomes very necessary to set up a suite of metrics for developers to understand the fundamental characteristics of ontologies, in order to improve the quality, estimate cost and reduce future maintenance.

In this study, we introduced a well-defined metrics suite for measuring ontology complexity, which mainly examine the quantity, ratio and correlativity of concepts, relationships and paths, to evaluate ontologies from the viewpoint of evolution. And we measured three organizing ontologies, biological process, cellular component and molecular function, in Gene Ontology to verify our metrics. Overall, the results indicate that the biological process is the most complex one from the view of complexity, and the molecular function is the unsteadiest one from the view of evolution.

In the future, we will continue work on the ontology complexity metrics and other ontology metrics. In this research, our metrics come mainly form the hierarchy in conceptual model. We may add more metrics form other sides of ontology. And upon these metrics, we have had some insights of semantic field of ontology. So, future research may include additional works for the metrics on ontology semantic ability.

## 7. References


[1] http://www.geneontology.org/

[2] H. S. Pinto and J. P. Martins. Evolving ontologies in distributed and dynamic settings. In Proceedings of the Eighth International Conference on Principles of Knowledge Representation and Reasoning (KR2002), Toulouse, France, Apr. 22-25, 2002.

[3] W. Nejdl, B. Wolf, C. Qu, S. Decker, M. Sintek, A. Naeve, M. Nilsson, M. Palmr, and T. Risch. Edutella: a p2p networking infrastructure based on rdf. In Proceedings of the eleventh international conference on World Wide Web, pp. 604-615. ACM Press, 2002.

[4] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. Scientific American, 284(5): 34-43, May 2001.

[5] A. Das, W. Wu, D. McGuinness, Industrial Strength Ontology Management', The Emerging Semantic Web, IOS Press, 2002.

[6] Sirin, E., J. Hendler and B. Parsia, 2003. Semiautomatic composition of web services using semantic description. Web Services: Modeling, Architecture and Infrastructure Workshop in Conjunction with ICEIS2003.

[7] Stojanovic L. and Motik B. Ontology Evolution within Ontology Editors. In Proceedings of OntoWeb-SIG3 Workshop at the 13th International Conference on Knowledge Engineering and Knowledge Management 2002, pp 53-62. Siguenza, Spain, 30th September.

[8] S. Staab, H.-P. Schnurr, R. Studer and Y. Sure, Knowledge Processes and Ontologies, IEEE Intelligent Systems. 16(1), Jan./Feb. 2001. Special Issue on Knowledge Management, 2001.

[9] Ljiljana Stojanovic, Alexander Maedche, Boris Motik, Nenad Stojanovic. User-Driven Ontology Evolution Management. Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management. Ontologies and the Semantic Web. pp: 285-300, 2002

[10] N. F. Noy and M. Klein. Ontology evolution: Not the same as schema evolution. Knowledge and Information Systems, Vol.5, 2003.

[11] Michel C. A. Klein, Natalya F. Noy. A Component-Based Framework for Ontology Evolution, Technical Report, IR-504, CS Dept, Vrije Universiteit, Amsterdam, 2003.

[12] M. Klein and D. Fensel. Ontology versioning for the Semantic Web, Proc. International Semantic Web Working Symposium (SWWS), USA, July 30-August 1, 2001.

[13] D. E. Oliver, Y. Shahar, M. A. Musen, and E. H. Shortliffe, Representation of change in controlled medical terminologies, AI in Medicine, 15(1): 53-76, 1999.

[14] Wand, Y. and Weber, R. (2002). Research Commentary: Information Systems and Conceptual Modeling - A Research Agenda. Information Systems Research, 13 (4), pp. 363-376.

[15] Corcho, O. and Gomez-Perez A. A road map to ontology specification languages. In Proceedings of the 12th International Conference on Knowledge Acquisition, Modeling and Management, Juan-les-Pins, France pp. 80-96.

[16] A. Burton-Jones, V.C. Storey, V. Sugumaran, and P. Ahluwalia. Assessing the effectiveness of the daml ontologies for the semantic web. In Eighth International Conference on Applications of Natural Language to Information Systems, June 2003.

[17] Haining Yao, Anthony Mark Orme and Letha Etzkorn. Cohesion Metrics for Ontology Design and Application, Journal of Computer Science 1(1): 107-113, 2005

[18] Dazhou Kang, Baowen Xu, Jianjiang Lu, William C. Chu. A Complexity Measure for Ontology Based on UML, 10th IEEE International Workshop on Future Trends of Distributed Computing Systems (FTDCS'04), pp. 222-228

[19] Idris His. Analyzing the Conceptual Coherence of Computing Applications Through Ontological Excavation, PhD Thesis Proposal, May 13, 2004.

[20] Chris Mungall, Increased complexity in the GO, http://www.fruitfly.org/~cjm/obol/doc/go-complexity.html