Journal of Information & Computational Science 7: 1 (2010) 95–101
Available at http://www.joics.com

Event Ontology and Its Evaluation ⋆

Zhaoman Zhong*, Zongtian Liu, Wei Liu, Yan Guan, Jianfang Shan

School of Computer Engineering and Science, Shanghai University, Shanghai 200072, China

Abstract

An event-oriented ontology model is proposed in this study on the analysis of event class relations and event instance relations based on event elements. Furthermore, corpus-based and user-based evaluation methods for event ontologies are discussed from the perspective of event classes and their relations. In comparison with the concept ontology, the event ontology represents knowledge with a higher granularity, and it will contribute new methods and technologies to semantics-based knowledge processing.

Keywords: Event; Event Class Relation; Event Instance Relation; Event Ontology; Ontology Evaluation

1 Introduction

The concept ontology mainly reflects the existing law of objects, especially classified and non-classified relations between objects, but it has obvious defects on describing events: (1) An event is associated with many concepts, and it is a knowledge unit with a higher granularity compared with a concept; (2) An event is dynamic, and its state is changing; and (3) Many of non-classified relations between concepts in the concept ontology are described by verbs. The relations between concepts are very complex because of the diversification of actions, and it is hard to measure the importance of relations. For the relations between events, they are relatively stationary and definite, and obviously their importance is different. It can not reveal the special relations between events, and furthermore makes the relations more complex to use the method of describing the relations between concepts to represent the relations between events.

In this paper, we take an event as the basic cognitive unit, study the relations between event classes and the relations between event instances from the perspective of event elements. Ontologies are a fundamental data structure for conceptualizing knowledge, and we should be able to say which of them best suits some predefined criterion. We discuss the evaluation methods of corpus-based event class coverage and user-based event class importance.

Related work. Pustejovsky [1] proposed the definition of an event centering on verbs, and he believes that the basic semantic unit of an event is a verb. WordNet gives the extensive definition of 'event'-'In a particular place and time what is happening' [2]. TDT evaluating conference

⋆Project supported by the National Nature Science Foundation of China (No.60975033),and the Shanghai Leading Academic Discipline Project (J50103).

*Corresponding author.
Email address: zhongzhaoman@shu.edu.cn (Zhaoman Zhong).

January 2010

defines an event as a 'narrowly defined topic for search' [3]. The evaluation conference of ACE defines an event as a special process including participants, which is usually described as the change of a status. Filatova [4] defined the atomic event as combination of named entities and the verbs or action-denoting nouns. Liu et al. [5] introduces a six-tuple of event representation from the objective point of view.

The research on event relations focuses on classified relations. Vargas-Vera et al. [6] took events as ontology classes, defined 41 event classes and introduced the hierarchical relations between event classes. ACE divided events into 8 parent types and 36 sub types. The verbs act as event triggers in many cases. VerbNet is the largest on-line verb lexicon currently available for English, which contains 5200 verbs and 237 classes. Kaneiwa et al. [7] proposed a method of distinguishing event class relations and event instance relations. Zhong et al. [8] divided event relations into some relations such as classified, component, cause-effect, follow, concurrence, etc. Various methods of ontology evaluation have been considered in the literature. Broadly speaking, most evaluation approaches fall into one of the following categories [9]: (1) those based on comparing the ontology to a 'golden standard', (2) those based on using the ontology in an application and evaluating the results, (3) those involving comparisons with a source of data about the domain to be covered by the ontology, and (4) those where evaluation is done by humans who try to assess how well the ontology meets a set of predefined criteria, standards, requirements, etc. The remainder of this paper is organized as follows. Section 2 presents an event-oriented ontology model on the basis of analyzing event class relations and event instance relations. Section 3 elaborates the methods of evaluation event ontology. Finally, we give conclusions and discuss future work in section 4.

2 Event Ontology Model

2.1 Events and Event Classes

Definition 1 An event refers to one thing happening at a specific time and location, involving a number of actors, and showing some action characteristics. We use $e =< A, S, T, O, L >$ to formally denote an event, which can be defined as a five-tuple . The elements of an event five-tuple are called event elements, which represent action (A), subject (S), object (O), time (T) and location (L) respectively.

Definition 2 An event class (EC) [5] refers to the set of events which have the same characteristics, denoted by $EC = \{E, A, S, T, O, L\}$, where E is the set of events, called the extension of the event class. A, S, T, O, L are the intension of the event class, which are the set of the same characteristics of certain element for each event.

2.2 Event Class Relations and Event Instance Relations

Event class relations. The event class relations associated by the action element are mainly as follows: (1) Classified relations. Let $EC_1 = (E_1, A_1, S_1, O_1, T_1, L_1)$ and $EC_2 = (E_2, A_2, S_2, O_2, T_2, L_2)$ be two event classes. The relation of EC_1 and EC_2 is a classified relation, if and only if $E_2 \subset E_1$ or $A_1 \subset A_2$, where EC_2 is the sub event class of EC_1 , and EC_1 is the parent event class of EC_2 , denoted by $R_H(EC_1, EC_2)$. (2) Component relations. EC_1 is composed of EC_2 , if each event of EC_1 is composed of the certain event ofEC_2 , denoted by $R_I(EC_1, EC_2)$. (3) Cause-effect relations. The relation of EC_1 and $EC_2$is a cause-effect relation, if the occurrence of events of EC_1 causes the occurrence of events of EC_2 with certain probability, and the probability is bigger than the given threshold α , denoted by $R_{CE}(EC_1, EC_2)$. (4) Follow relations. The relation of EC_1 and EC_2 is a follow relation, if the occurrence of events of EC_1 is followed by the occurrence of events of EC_2 with certain probability, and the probability is bigger than the given threshold β, denoted by $R_F(EC_1, EC_2)$. (5) Concurrence relations. The relation of EC_1 and EC_2 is a concurrence relation, if the event of EC_1 and the event of EC_2 occur at the same time or successively with certain probability, and the probability is bigger than the given threshold γ , denoted by $R_C(EC_1, EC_2)$.

Event class relations. The event instance relations associated by subjects, objects, times and locations are mainly as follows: (1) Identity relations. Let $e_1 = (a_1, s_1, o_1, t_1, l_1)$ and $e_2 = (a_2, s_2, o_2, t_2, l_2)$ be two event instances. The relation of e_1 ande_2 is an identity relation, if and only if $s_1 = s_2$ or $o_1 = o_2$or $t_1 = t_2$ or $l_1 = l_2$, denoted by $r_s(e_1, e_2)$. (2) Inclusion relations. The relation of e_1 ande_2 is an inclusion relation, if and only if $s_1 \subset s_2$ or $o_1 \subset o_2$ or $t_1 \subset t_2$ or $l_1 \subset l_2$, denoted by $r_i(e_1, e_2)$. (3) Subject-object relations. The relation of e_1 ande_2 is a subject-object relation, if and only if $s_1 = o_2$ or $o_1 = s_2$, denoted by $r_{so}(e_1, e_2)$. In some cases, the subject or object of an event is another event, and then, the relations of event instances are as follows: (4) Subject-restriction relations. The relation of e_1 ande_2 is a subject-restriction relation, if and only if $e_1 = s_2$, denoted by $r_{es}(e_1, e_2)$. (5) Object-affiliation relations. The relation of e_1 ande_2 is an object-affiliation relation, if and only if $e_1 = o_2$, denoted by $r_{eo}(e_1, e_2)$.

2.3 Event Ontology Model

Definition 3 An event ontology is a shared, objective, and formal specification of an event class system model. Formally, using EO to denote an event ontology, which can be defined as a quadruple: $EO =< ECS, E^*, W, R >$, where $ECS = \{EC_1, EC_2, ..., EC_n\}$ is the set of event classes, $E^* = \{< EC_i, ECj > |EC_i, EC_j \in ECS\}$ is the set of directed edges, $W = \{w_{ij}|w_{ij}\ is\ the\ correlation\ strength$ $of\ < EC_i, EC_j >\}$, and $R = \{r|r\ is\ the\ relation\ of\ < EC_i, EC_j >, r \in (R_H, R_I, R_{CE}, R_F, R_C)\}$.

3 Event Ontology Evaluation

3.1 Domain Event Ontologies and Evaluation Indexes

We constructed five domain event ontologies for evaluation: 'earthquake', 'fire', 'traffic accident', 'terror attack', and 'food poisoning'. The number of event classes and their relations for each event ontology is shown in Table 1.

From Table 1, the percentage of non-classified relations is 89.3% in all relations.

Fig.1 is the fragment of the domain event ontology of 'earthquake', which is manually constructed by experts and includes 13 event classes. The relation associated by broken line is the relation between an event class and an event instance.

As shown in Fig.1, 'fire' and 'earthquake' are all sub event classes of 'emergency'. 'emergency' results in 'death and injury' and 'rescue', and 'fire' and 'earthquake' all inherit these attributes.

Table 1: Number of event classes and relations

	Number of event classes	Classified relations	Non-classified relations
earthquake	107	35	211
fire	116	22	176
food poisoning	124	27	254
traffic accident	98	19	192
terror attack	105	29	266
Total	550	132	1099

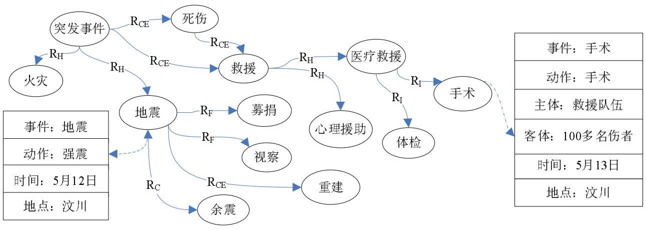

Fig. 1: an example of EO (fragment)

Event ontology contain event classes and relations, and the coverage and importance can be used to evaluate event ontologies. The coverage is given by Eq. (1):

$$C = \frac{NumEOD}{NumD} \tag{1}$$

where $NumEOD$ is the number of event classes included by the event ontology and documents, and $NumD$ is the number of event classes in documents. Because the domain event ontology often contains important event classes in a domain, we should select some important events in documents to evaluate event ontology.

The importance is to measure the importance of event classes in an event ontology. Because the event class importance is determined by event class relations, evaluating event class importance refers to evaluate event class relations. The precision and deviation are used to evaluate event class importance, which are given by Eq. (2) and (3) respectively:

$$Precision = \frac{|CS \bigcap ES|}{|ES|} \tag{2}$$

$$Deviation = \frac{\sum\limits_{|ES|} |NumCS - NumES|}{|ES|} \tag{3}$$

where CS denotes the event class set obtained by using an algorithm, which is ranked by descending order according to event class importance, and ES is the event class set given by users, which is sorted by descending order. $NumCS$ is the serial number of an event class in CS, and $NumES$ is the serial number of the event class in ES. Note that if an event class of CS is not in ES, the deviation of it is $|ES|$. The indexes and applicability are shown in Table 2.

Table 2: The indexes of evaluating event ontology

	Coverage	Importance
Evaluation methods	$Corpus - based$	$User - based$
Applicability	$Alleventclasses$	$Eventclassrelations$

3.2 Corpus-based Event Class Coverage Evaluation

For each category, we randomly download 30 documents from Internet. After identifying events from documents according to identifying event rules, introduced in literature [10], we get event weights through computing events $TF * IDF$.

We select some events to evaluate event ontologies, and the number of which in documents is equal to the number of event classes in event ontologies. The detailed steps of evaluating an domain event ontology are as follows: (1) Identify all events from documents to get set $E1$, and compute event weights according to $TF * IDF$; (2) Select the same number of events from set $E1$ to get set $E2$ according to the number of event classes in an event ontology, and suppose that the number of event classes in an event ontology is $N1$; (3) Compute the number of events included by the event ontology and $E2$; and (4) Use Eq. (1) to get $C = \frac{NumEOD}{|E2|}$. Applying the above steps, the evaluation results for five event ontologies are shown in Table 3.

Table 3: The evaluation results for five domain event ontologies

	earthquake	fire	food poisoning	traffic accident	terror attack
Coverage	81.3%	77.6%	82.3%	80.6%	90.5%

From Table 3, 'terror attack' is an best domain event ontology, and 'fire' is worst according to the coverage index.

3.3 User-based Event Class Importance Evaluation

The users are responsible for giving five important event classes for each category, and order them by descending order according to event class importance. We use link analysis method to get five important event classes from each event ontology, and compare them with event classes given by users.

Quantifying relations for event ontologies. Different relations between event classes have different association strengths. The event class relations can be quantified in practical applications, and be converted into decimal fraction from 0.0 to 1.0. One method of determining the association strengths of event classes is dependent on ontology experts who directly provide the relations, then using rules to quantify the relations. The other is to collect related documents for a domain, then applying computers to calculate associative strengths between events. In this paper, we give the rules of quantification: (1) if $R_H(EC_i, EC_j)$, then $w_{ij} = 1.0/m$, m is the number of sub event classes contained by EC_i ; (2) if $R_I(EC_i, EC_j)$, then $w_{ij} = 1.0$; (3) if $R_{CE}(EC_i, EC_j)$, then $w_{ij} = 0.8$; (4) if $R_F(EC_i, EC_j)$, then $w_{ij} = 0.6$; and (5) $R_C(EC_i, EC_j)$if , then $w_{ij} = 0.4$.

Computing event class importance for event ontologies. Methods of ranking objects have been widely discussed in a variety of papers such as PopRank [11], Swoogle [12], Reverse PageRank [13], and so on. PopRank and Swoogle emphasize on object's authorities, which are the improved algorithms of PageRank. Reverse PageRank employ object's hubs. HITS [14] is a page-oriented link analysis technique, which can rank pages by hubs or authorities.

Wu et al. [15] computed the importance of concepts with Hub rating instead of Authorities rating. Zhi et al. [16] applied the algorithms in literature [15] to compute event importance for an event directed graph. But the relations between events are different from concept relations, and different relations between events imply different associative strengths.

We draw the idea in literature [15] and use Eq. (4) to compute event class importance according to the characteristics of event ontologies. The algorithm of computing event class importance is as follows:

$$R(u)_k = d \sum_{j \in Out(u)} R(j)_{k-1} w_{uj} + \frac{1-d}{n} \tag{4}$$

where $R(u)$ is the importance of u , and $R(u)_k$ denotes the kth iteration of $R(u)$. d is a damping coefficient, ranging from 0.0-1.0, usually $d = 0.85$). According to the surfing model, d refers to how long Web browsers will feel tired. $Out(u)$ refers to the node set in which each node is pointed by u , n is the number of nodes in the network, and w_{uj} is the linking strength from node j to node u.

Ranking effectiveness. Eq. (4) is used to compute event class importance. The precision of iteration convergence is , and . Five important event classes are obtained to compare with the results given by users. The experimental results are shown in Table 4.

Table 4: The evaluation results for five domain event ontologies

	earthquake	fire	food poisoning	traffic accident	terror attack
Precision	80%	60%	80%	80%	60%
Deviation	2.31	3.55	2.16	2.53	3.29

From Table 4, 'food poisoning' is best, and 'fire' is worst according to the importance index.

From Table 4 and 5, the coverage of 'terror attack' is best, but its importance is average, namely the relations between event classes are not reasonable enough. The importance index of 'food poisoning' is best, and its coverage is 82.3%, which means that the importance and coverage are all reasonable.

4 Conclusions and Future Work

The concept ontology model can effectively represent hierarchical structures and semantic relations for static concepts, and it can solve the semantic problems in the field of computer applications to a certain extent. But in the process of modeling the real world, the dynamic events that contain actors and their relations should also be described except representing static concepts and their relations. Therefore, we take an event as the basic unit for knowledge representation, and propose an event ontology model. Evaluating event ontologies mainly relates to event classes and their relations. And the methods of corpus-based coverage and user-based importance are applied to evaluate event ontology according to its characteristics. In future work, we plan to study the semi-automatic constructing technology for the event ontology and its applications in practice.

Acknowledgement

We want to thank the members of our research group, who provided a lot of helpful advice for this paper.

References

[1] Pustejovsky J. Events and the Semantics of Opposition, in Events as Grammatical Objects, Pustejovsky J,Tenny C, Eds.: Stanford: Center for the Study of Language and Information (CSLI Publications), 2000, pp. 445-482

[2] WordNet, a lexical database for the English language [EB/OL]. [2006]http://wordnet.princeton.edu/

[3] Allan J, Carbonnell J, Doddington G, et al. Topic Detection and Tracking Pilot Study: Final report. Proceedings of the DARPA Broadcast News Transcription and Understanding Workshop, 1998

[4] Filatova E, Hatzivassiloglou V. Domain-independent Detection, Extraction, and Labeling of Atomic Events. Proc. of RANLP. Borovetz, Bulgaria, 2003, pp. 145-152

[5] Liu Zongtian, Huang Meili, Zhou Wen, et al. Research on Event-Oriented Model. Computer science, 2009, 36(11):189-192 (in Chinese)

[6] Vargas-Vera M, Celjuska D. Event Recognition on News Stories and Semi-Automatic Population of an Ontology. Web Intelligence, WI 2004. Proc. of IEEE/WIC/ACM International Conference, 2004, pp. 615-618

[7] Kaneiwa D, Iwaztume M, Fukuda D. An Upper Ontology for Event Classifications and Relations. M.A. Orgun and J.Thornton (Eds.): AI 2007, LNAI 4830,2007, pp. 394-403

[8] Zhong Zhaoman, Liu Zongtian, Zhou Wen, et al. The Model of Event Relation Representation. Journal of Chinese Information Processing, 2009, 23(6): 56-60 (in Chinese)

[9] Brank J, Grobelnik M, Mladeni? D. A Survey of Ontology Evaluation Techniques. Proc. of conference on Data Mining and Data Warehouses (SiKDD 2005), Ljubljana, Slovenia, 2005

[10] Zhong Zhaoman, Liu Zongtian. Ranking Events Based on Event Relation Graph for a Single Document. Information Technology Journal, 2010, 9(1): 174-178

[11] Nie Z, Zhang Y, Wen J R, et al. Object-level Ranking: Bringing order to Web objects. Proc. of the WWW.Chiba, Japan, 2005, pp. 567-574

[12] Ding L, Pan R, Finin T, et al. Finding and Ranking Knowledge on the Semantic Web. Proc. of the IS- WC.Galway, Ireland, 2005, pp. 156-170

[13] Fogaras D. Where to start browsing the Web? Proc. of the IICS. Leipzig, Germany, 2003, pp. 65-79

[14] Kleinberg J M. Authoritative Sources in a Hyperlinked Environment. Journal of ACM, 1999,46(5):604-632

[15] Wu Gang, Zhang Kuo, LiI Juanzi, et al. Ranking by Mutually Reinforcing Concepts and Relations in Ontology.Chinese Journal of Computers, 2007, 30(9):1490-1499 (in Chinese)

[16] Zhi H L, Liu Z T. Event Importance Analysis Based on Directed Graph. International Symposium on Intelligent Information Technology Application Workshops. Shanghai, China, 2008, pp. 451-453