

Features, Improvements and Applications of Ontology in the Field of Sports Events during the Era of the Semantic Web

Juan Xiao¹ and Jing Chen²

¹ Department of Sports, Wuhan University, Wuhan 430072, China
973266289@qq.com

² Department of Foreign Language, Hubei University of Technology, Wuhan 430064, China
370958331@qq.com

Abstract. Domain Ontology is a cutting-edge hot topic during the era of the Semantic Web. This paper studies Ontology in the field of sports events. Firstly, it explains the overview of Ontology in the field of sports and Fundamental Ontology, and then makes a comparison between them. It proceeds to provide recommendations for the improvement. The characteristics of Ontology in the field of sports events are summarized and its logical triples interpreted. The intelligent application of Ontology in the field of sports events is consequently discussed.

Keywords: Ontology, Domain Ontology, Semantic Web, intelligent application.

1 Introduction

Ontology is a conceptual model describing the concept and the relationship between concepts, and it explains the concept by the relationship between concepts. It is proper to formalize human knowledge and information, by which it constructs a common information understanding mechanism between the user and the machine, and realizes the sharing of domain knowledge.

2 An Overview of the Ontology in the Field of Sports Events and Domain Ontology

2.1 Fundamental Ontology and Domain Ontology

Knowledge existing outside the human knowledge system belongs to the domain of fundamental knowledge. The Ontology is based on it is Fundamental Ontology. Fundamental Ontologies such as WordNet, FrameNet and CCD are widely used. The Ontology based on a special professional or Domain knowledge is Domain Ontology.

2.2 The Ontology in the Field of Sports Events

The Ontology in the field of sports events is a typical Domain Ontology. It collects some concepts and terminology in the field of sports events, and constructs their relationships.

3 Partial Comparison and Integration of the Ontology in the Field of Sports and Fundamental Ontology

3.1 The Base of Comparison and Integration

In this paper, Fundamental Ontology is based on *Synonymy Thesaurus* by Mei Jiaju and the Ontology in the field of sports events is based on *The Development of Sports Corpus and its Sports Lexical Study* by Chen Wei and *Chinese-English Sports Classification Dictionary* by Chen Naixin. On this basis, it initially develops an Ontology in the field of sports events.

3.2 Partial Comparison and Integration

For lack of space, we only choose some key words to compare and integrate partially the Ontology in the field of sports events and Fundamental Ontology.

3.2.1 From the Perspective of the First-Layer Node

The first-layer node of fundamental Ontology which corresponds to the Ontology in the field of sports events mainly focus on the categories of [man], [thing], [time and space], [abstract thing] and [sport], and nearly have no correspondence to the categories of [characteristic], [psychological], [activity], [phenomenon and condition],[relation][auxiliary vocabulary], and [complimentary vocabulary]. This shows that the semantical categories of the Ontology in the field of sports events have their focuses. The phenomenon that the categories of [organization] and [referee] correspond to the categories in Fundamental Ontology of [abstract thing] [activity] and [thing][action] shows their complexity. By comparing them, we know that there are some differences between each first-layer node of the two Ontology.

3.2.2 From the Perspective of the Second-Layer Node

We here only compare [personnel] of the Ontology in the field of sports events and [human] of Fundamental Ontology. Fundamental Ontology has no categories of [referee], and [assistant clerk] shows its defect, because these two categories are fundamental in fact, having a higher degree of generalization.

3.2.3 From the Perspective of the Third-Layer Node

We only contrast the words in [athletes] of the Ontology in the field of sports events with the words in [athletes] of fundamental Ontology. As is shown in the following, the first layer derives from the Ontology in the sports events [1], and the below layers are from fundamental Ontology [2]:

- 运动员 yundongyuan athlete
- 球员 qiuyuan player 发球员 faqiuyuan server 击球员 jiqiuyuan hitter 接球员 jieqiuyuan receiver 接发球员 jiefaqiuyuan receiver 守门员 shoumenyuan goalkeeper 跑垒员 paoleiyuan base runner 击球手 jiqiushou batter 投手 touthou pitcher 垒手 leishou corner man
- 拳击手 quanjishou typewriter 棋手 qishou chess player 选手 xuanshou contestant 种子选手 zhongzixuanshou seeded player 非种子选手 feizhongzixuanshou unseeded player
- 前锋 qianfeng forward 小前锋 xiao qianfeng small forward 后卫 houwei guard 中锋 zhongfeng center 替补1 tibu alternate 自由人 ziyouren libero 二传手 erchuanshou second pass
- 主队 zhudui host team 客场 kechang away games 击球方 jiqiufang Batting side 发球方 faqiufang serving side 接发球方 jiefaqiufang receiving side
- 运动员 yundongyuan athlete 选手 xuanshou contestant 健儿 jian'er strong man
- 种子 zhongzi seed 健将 jianjiang master
- 前锋 qianfeng forward 中锋 zhongfeng center 左锋 zuofeng left forward 右锋 youfeng right forward 中卫 zhongwei halfback 前卫 qianwei vanguard 左卫 zuowei left back 右卫 youwei right back 后卫 houwei guard 守门员 shoumenyuan goalkeeper
- 一传手 yichuanshou first pass 二传手 erchuanshou second pass 主攻手 zhugongshou main attacker
- 投手 touthou pitcher 捕手 bushou backstop 一垒手 yileishou first baseman 二垒手 erleishou second baseman 三垒手 sanleishou third baseman 左翼手 zuoyishou left wing man 右翼手 youyishou right wing man 中坚手 zhongjianshou backbone man 游击手 youji shou shortstop

We will not translate them again when referring to some of the above words and expressions therein.

There are 28 words in the node of [athlete] about the Ontology in the field of sports, and 27 words in the node of [athlete] about fundamental Ontology. The total number of the words in each Ontology is equivalent. They both have the 8 words as shown above.

Through Fundamental Ontology, we find that the Ontology in the field of sports events has its disadvantages, and the improvement are furnished as follows:

Firstly, the nodes are few and their capacity is too large to highlight the advantages of the concrete of Domain Ontology. For instance, there seems a lack of the terms of “左锋 zuofeng left forward”, “右锋 youfeng right forward”, “中卫 zhongwei halfback” and “前卫 qianwei vanguard”, etc. As is usually known, sport is a national event and an issue of public concern, so the sports terms have a high degree of popularization. We should enlarge the size of the nodes properly during the construction of Fundamental Ontology.

Secondly, the structure level of few nodes is too general. For example, the terms of “前锋 qianfeng forward”, “小前锋 xiao qianfeng small forward” and “后卫 houwei guard” can be disintegrated into different subsidiary nodes. [3]

By contrast, we find the disadvantages of the Ontology in the field of sports events. Likewise, through the Ontology in the field of sports events, we can also find the disadvantages of the Ontology as follows:

On the one hand, the nodes are few and their capacity is too large to include the terms with high a popularizing degree, such as “球员 qiuyuan player”, “棋手 qishou chess player” and “替补 tibu alternate”, etc.

On the other, the subsidiary nodes are so many that some terms are suitable to appear only in the constructing process of the field Ontology, such as “一垒手 yileishou first baseman”, “二垒手 erleishou second baseman”, “三垒手 sanleishou third baseman”, “左翼手 zuoyishou left wing man”, and “右翼手 youyishou right wing man”, etc.

4 Characteristic of the Ontology in the Field of Sports Events

4.1 Dissimilarity of the Structure Level

Within the Ontology in the field of sports events, [man] is in the first floor and [sportsman] is in the third floor. It shows that the same nodes are in the different structure levels when they appear separately in Fundamental Ontology and the Ontology in the field of sports events.

4.2 Scientific of the Structure Level

The Framework for the design and distribution of nodes of the Ontology in the field of sports events is relatively professional and scientific. A very important reason is that the Ontology in the field of sports events can adopt the triple logic form of <concept, attribute, instance>, introducing the attribute, and yet the Fundamental Ontology is short of the attribute.

4.3 Diversification of the Nodes

The category “sportsman” has a high commonality. A small number of lower nodes are separated from the node of [sportsman] for Fundamental Ontology. On the contrary, a large number of lower nodes are separated from the node of [sportsman] for the Ontology in the field of sports events. This shows that the lower nodes of [sportsman] in the field of sports events are getting more, so as to provide the professional and disciplinary terms.

5 Logic Triples of the Node of Ontology in the Field of Sports Events

In theory, the nodes of Ontology in the field of sports events include concepts, attributes, instance and so on. So it could be expressed by the logic triples of <concept, attribute, instance>. Take the node of [athletes] from the Ontology in the field of sports events above, for example. The relationship between it and its lower nodes as shown below in Fig. 1.

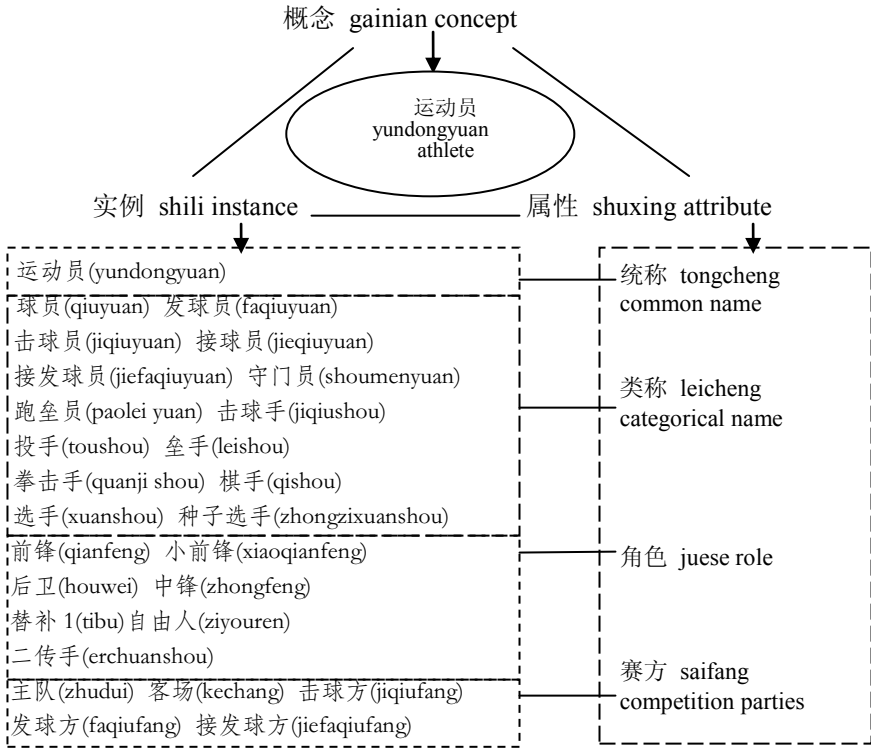


Fig. 1. The relationship between the node of [athlete] and its lower nodes from the perspective of the logic triples

Attribute is reusable. That is to say, attributes that appear in the upper node can also appear in the lower nodes. For example:

<athlete, role, baseman >< baseman, role, first baseman >< baseman, role, second baseman >< baseman, role, third baseman >

The construction of the Ontology in the field of sports events can be achieved by using it. For example:

[athlete] — {role} — [baseman] — {role} — [first baseman]
[second baseman][third baseman]

6 Intelligent Applications on Ontology in the Field of Sports Events

6.1 Improving the Level of Intelligent Retrieval in the Field of Sports Events

At present, there are three types of information retrieval: text retrieval, data retrieval and knowledge retrieval. In dealing with the sport category, the Ontology in the field of sports events has a much better concept hierarchy and supports performance for

logical reasoning, so it can be widely used in the intelligent retrieval for sport information and knowledge.

The following is a true text selected randomly from the webpage¹:

6日晚，C罗头顶脚踢完成了本赛季西甲联赛的第4次“戴帽”，皇马7比1横扫奥萨苏纳。而算上曼联时代，C罗职业生涯的帽子戏法数已提高到了13次。C罗本赛季的联赛进球数提升到了13个，他用进球为自己刚刚拿到的欧洲金靴奖庆祝。

6 ri wan, C luo toudingjiaoti wancheng le ben saiji xijia liansai de di 4 ci daimao, huangma 7 bi 1 hengsao aosasu'na. er suanshang manlian shidai, C luo zhiye shengya de maozi xifa shu yi tigao dao le 13 ci. C luo ben saiji de liansai jinqiushu tisheng dao le 13 ge, ta yong jinqiu wei ziji ganggang nadao de ouzhou jinxue jiang qingzhu.

In this paper, the translation of this text is as follows, and we will not translate it again when referring to some words and expressions therein.

On the evening of (November) 6th, C Luo (Cristiano Ronaldo) won his forth hat-trick of the La Liga this season with his head and foot, and Real Madrid beat Osasuna by 7 to 1. Counting the Manchester United era, the number of C Luo's hat-tricks in his career has been increased to 13. C Luo's has gained 13 marks this season, and he just celebrated his European Golden Shoe with the last shoot.

The Internet user could usually enter the keywords or sentences for search:

6日晚本赛季西甲联赛谁进球？

6 ri wan ben saiji xijia liansai shui jinqiu?

Who scored the La Liga this season on the evening of (November) 6th?

6日晚西甲联赛哪个球队获胜？

6 ri wan ben saiji xijia liansai na ge qiudui huosheng?

Which team won the La Liga this season on the evening of (November) 6th?

Presently, the major search engines are still equipped with the keywords searching method. So the user often only enters the keywords of “进球 jinqiu goal”或“西甲联赛 xijialiansai La Liga”. Obviously, from the text above, the user cannot retrieve the correct information needed. But if the user makes full use of the Ontology in the field of sports events, he will obtain the information needed.

Through the structure tree above, we find that the node of [戴帽(daimao)] is the lower node of [进球(jinqiu)], and the machine/program will judge that“戴帽(daimao)” means“进球(jinqiu)”. We also find that [横扫(hengsao)] is the synonymy node of [获胜(huosheng)], and the machine will judge that“横扫(hengsao)” means“获胜(huosheng)”. Then it will give the answer as follows to the user through a series of computations.

6日晚本赛季西甲联赛谁进球？→6日晚本赛季西甲联赛C罗进球。

→ 6 ri wan ben saiji xijia liansai C luo jinqiu.

→ On the evening of (November) 6, C Luo goaled in the La Liga season.

6日晚西甲联赛哪个球队获胜？→6日晚西甲联赛皇马获胜。

→ 6 ri wan ben saiji xijia liansai huangma huosheng.

→ On the evening of (November) 6, Real Madrid beat the La Liga.

¹ See <http://news.163.com/11/1107/14/7I919TO200014AED.html>

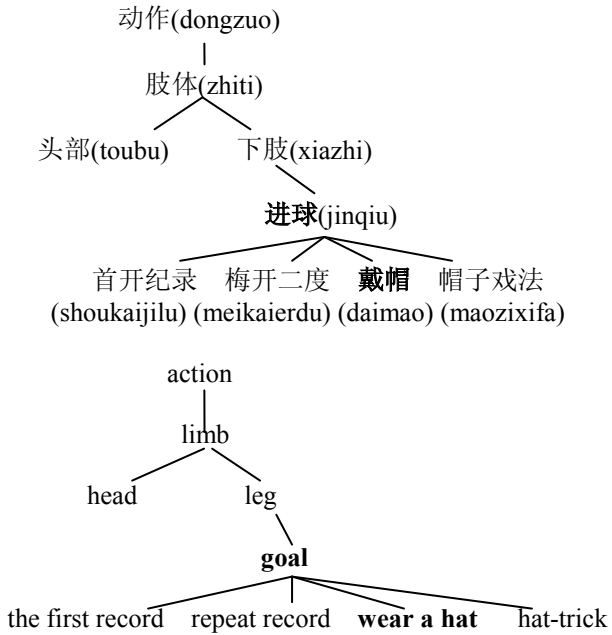


Fig. 2. The structural tree of the nodes of [动作 dongzuo action]

6.2 Promoting the Depth of Text Understanding in the Field of Sports Events

The event predicate can be regarded as a concept and has its attribute which can be granted a certain value. So each sentence can be expressed by the logic triples of <concept, attribute, instance>. For the event predicate, they all have the characteristic of carrying a particular argument, and the value of the characteristic is the predicate’s semantic role of “agent”, “patient, and so on. When we further analyze the text above, and find its structure can be expressed as follows:

- <头顶脚踢(toudingjiaoti), AGENT, C 罗(C luo)>
- <进球(jinqiu), AGENT, C 罗(C luo)>
- <横扫(hengsao), AGENT, 皇马(huangma)>
- <横扫(hengsao), PATENT, 奥萨苏纳(aosasu'na)>
- <提高(tigao), EXPERIENCE, 帽子戏法数(maozixifashu)>
- <获胜(huosheng), AGENT, 皇马(huangma)>
- <提升(tisheng), EXPERIENCE, 进球数(jinqiushu)>
- <庆祝(qingzhu), AGENT, 他(ta)>

Suppose the user continues to enter the following search request:

C罗为什么庆祝？

C luo wei shenme qingzhu?

Why did C Luo celebrate?

Because “他(ta)”和“C罗(C luo)” are of the co-referential relationship, “C罗(C luo)”和“皇马(huangma)” are of the controlled and controller relationship, a chain of the theme events has formed in Fig. 3 on the basis of the structure above.

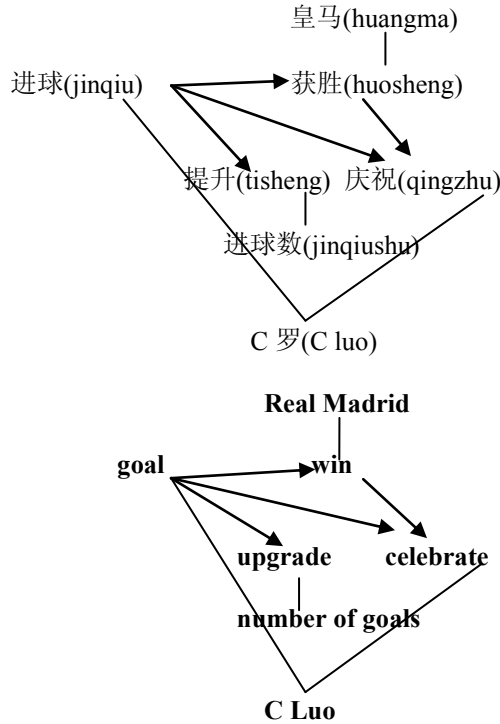


Fig. 3. The chain of the theme events

The chain of the theme events is effective for the computer to understand and to induce the web text information. The computer can judge that the central event of the text is “进球(jin qiu)”, and that the other three events of “获胜(huosheng)”, “提升(tisheng)” and “庆祝(qingzhu)” are all the chain reaction of the central events.

Computer will return the result below:

- C罗为什么庆祝？→因为C罗进球， 皇马获胜， C罗进球数提升。
- Yinwei C luo jin qiu, huangma huosheng, C luo jin qiu shu tisheng
- Because C Luo goaled, Madrid won the match, and C Luo’s goals increased.

6.3 Improving the Quality of Machine Translation in the Field of Sports Events

When the machine translates, we can mark each node of the Ontology in the field of sports events with a semantic code. Because Chinese and English use the same

semantic code, the corresponding relations between them will be clear. See another example below:

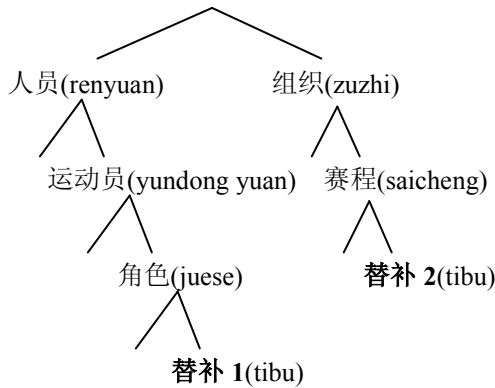
米利安每场比赛都是以替补出场。

Mili'an mei chang bisai jun shi yi tibu chuchang.

Miljan is a bench every game.

If we make the correct semantic code label in advance, then the machine would not translated “替补(tibu)”into “替补2(tibu)”, which is used specially in “race” in sports events. So the only correct result is to translate “替补(tibu)”into “替补1(tibu)”, which is used as “role” of sports events, and we get the result as shown in Fig. 4.:

体育赛事领域Ontology(tiyu saishi lingyu Ontology)



The Ontology in the field of sports events

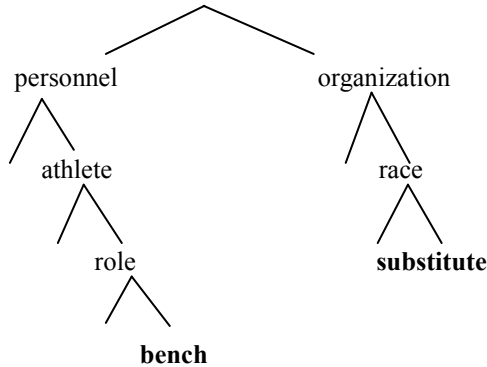


Fig. 4. Part of the Ontology in the field of sports events

6.4 Providing the Base of Web Information Sharing and Exchange in the Field of Sports Events

Ontology in the field of sports events is the base of sharing, and it exchanges the web information on the level of semantics. As we all know, XML and RDF not only provide grammar frameworks but also some semantic description. For example, the

XML fragment of `<Author > Jack < / Author >` shows that Jack is author, the RDF fragment of `< rdf : Description about = "http : //sports.sohu.com/ Home / Steven " >< s : Creator > Steven < / s : Creator > < / rdf : Description >` shows that Steven is the founder of the webpage of "http://sports.sohu.com/Home/Steven". From the point of <concept, attribute, instance>, the attribute scope of XML and RDF is lacking in specifications and restrictions, and is subject to change. For example, "Author" and "Creator" could be replaced by "Writer" in the text, and the result becomes `<Writer > Jack < / Writer >< rdf : Description about = "http : //sports.sohu.com/ Home / Steven" >< s : Writer > Steven < / s : Writer > < / rdf : Description >`. In fact, for Ontology, "Author", "Creator" and "Writer" are of the same concept, and can share a same upper node.

For another example, the node of [instructor] is both appearing on the webpage of a basketball club website and a fitness equipment website, XML and RDF cannot well define it as "coach" or "equipment manual". At this time, the Ontology in the field of sports events can achieve disambiguation by judging that they should belong to different nodes, and the distance of nodes is rather long.

7 Outlook

The construction of Ontology in the field of sports is not a task that can be completed once for all. We should view it from the point of development because it has a bearing on the evolution of the Ontology. Therefore, it is a trend that the evolution of one Ontology in the field of sports could lead to another Ontology of the same domain.

References

1. Chen, W.: A study on the building of the sports register corpus's building and its event vocabulary. *Journal of Nanjing Normal University*, 38–41 (2007)
2. Mei, J.J.: *Synonymy Thesaurus*. Shanghai Dictionary Publishing House, Shanghai (1983)
3. Chen, A.H.: *Sports Dictionary*. Shanghai Lexicographical Publishing House, Shanghai (2000)