The Study on the Organization Approach of Agricultural Model Component Library Based on Topic Map

Haiyan Jiang [1,2], Bing Fu [2], Mei Zhang [1,2], Yan Zhu [2], Weixin Cao [2*]

[1] College of Information Science and Technology, Nanjing Agricultural University, Jiangsu Province Nanjing 210095, China

[2] Hi-Tech Key Laboratory of Information Agriculture, Nanjing Agricultural University, Jiangsu Province Nanjing 210095, China

Abstract. The definition and representation on the model base's descriptions are the basis of reuse, integration and management of heterogeneous model resources. The study proposed the organization method of the agriculture model component based on topic maps. On the basis of the meta-modeling techniques and the topic map techniques, the study comprehensively analyzed the agricultural model characteristics, extracted the model component facets and its associated descriptions, and established the agricultural model component library's description model (DM-AMCL). Then, it hierarchically mapped and annotated the model component semantic and syntactic information by using the topic map. We designed the organization framework of the agriculture model component library on topic maps (TM-AMCL). For example of wheat development model, the XTM document instance is build. The result shows that, converting the agriculture models base to TM-AMCL can provide a unified descriptions and access to data for the intelligent decision support system developers, the managers of agriculture model and the computer system.

Keywords：Topic map, model component, topic maps merging, development model, agriculture model

1 Introductions

The agriculture model is a mathematical model or computer model to describe the characteristics, the status and the variety law of agro-ecosystems [1]. The Agricultural Decision Support System based on the agriculture models plays an important role in the field of precision agriculture, crop production management, plant pest forecasting, agriculture economic analysis and forecasting[1],[2],[3]. With the maturation and development of component-based software development approach (CBD), more and more simulation models are encapsulated as model components or web services, the decision support system development efficiency and quality are improved with the component reuse and assembly [4]. The APSIM research group [5] use the component assembly method to integrate the grass seed models built with the StallaTM tool into the APSIM system. van Ittersum[6] etc., established the component-based seamless integration framework (SEAMLESS-IF), the Agricultural model in different spatial and temporal scales such as land use planning, crop productivity forecasting, economic and social benefit analysis, can be integrated in this framework, and the reusable level between model components is promoted. ZHAO Chun-jiang[7] etc., applied the software component technology to agricultural intelligent system platform, raised the platform development efficiency. CAO

Haiyan Jiang and Bing Fu contributed equally to this work. e-mail: jianghy@njau.edu.cn.
Weixin Cao is corresponding author. e-mail: caow@njau.edu.cn

Wei-xing[8-9] etc., developed crop growth simulation system by using the component-based method to encapsulate the model components of wheat, rice, cotton and rape and other major field crops for the objects, efficiently improved the development efficiency of the agricultural decision support system.

However, in the agriculture information research field, although many different research institutions have large amounts of different types of model component resources, but they are present in their development of independent system. These research institutions adopted different modeling approaches and component model, used different development platforms and technologies, and it's difficult to integrate, share and reuse the models. At the same time, agricultural models involve many aspects, and are high complexity, highly professional, the dependency and relationship between the sub-models are close. All of these made intelligent decision support system developers, agricultural science researchers and computer system to have a different semantics and syntax understanding for the model's input / output parameters. Therefore, to research the unified description and representation method of the agricultural model base distributed environment, can helps all kinds of users to unify the various modes of information access to heterogeneous model base. It's the basis for realizing the agriculture model integration and reuse.

Topic Maps technology [10] is a metadata model for expressing and exchanging the structured information, and has now become the ISO / IEC 13250 standard published. Topic Maps technology played an important role in the fields of heterogeneous databases sharing and integration of subject knowledge [11]. Ellouze N [12] raised a global sharing method of topic maps in different languages through continuous collaboration and fusion between topic maps. WU Xiao-fan [13] raised a method to construct the topic map-based knowledge repository warehouse, to promote information integration to knowledge integration. Although, there have many applied research of topic maps technology in fields of information resource integration and information retrieval, there have less study on the representation of agriculture model base resources, especially on the content representation and build process. This study intended to apply the topic maps technology to the information organization of agricultural model base. By analyzing the features of agriculture model and combining the business component model, the study constructed the layered description model of the agriculture model component library, confirmed the construction process and content on the topic maps of the agriculture model component library. The goal is to convert agricultural models to the topic maps, and provide a unified reuse access to agriculture models in the network environment.

2 The representation method of agriculture model component Library based on the topic maps

2.1 The composition and structure of agriculture model base

Software component is a completely defined self-contained software unit of one or a group interface, has syntax formula, time and deployment of content, can be independently delivered and installed at a point of time of component development life cycle, and represents its effectiveness by assembling with other components [14].The effective use of the component-based development (CBD) methods depends on the large number of component library with an effective organization. Currently, the most agriculture models are encapsulated to model components. Agricultural model library resources are commonly composed of agricultural model library, model component library and model database.

Agricultural model base is a collection of agriculture model concepts and its relationship, model component library is specifically represented as the model algorithm, and is composed of the agriculture model's information of the concept terms, component properties, interface specifications and component entities. Model metadata, model parameters and model-driven run-time data is stored in the model database. Therefore, to unified organize and represent the agriculture model base, needs a totally characterization and description of the model component's semantic and syntactic information.

2.2 Topic map elements and the mapping of agricultural model base resources

Topic maps (TM) are usually used to represent knowledge and to locate data structure of information resources. The TM is composed of three elements, which are topic, association and occurrence, and represents the complex relationship between knowledge concept and information resource through the meta-data XML (XML Topic Maps) [10]. The topic describes the abstract concept of things, and the topic type represents type of classified topic. The association describes the relationship between the topics. The association type depicts the interrelated nature of topic concept. Since there is no direction or order about the association, the association role is used to define the participated role in the association. The occurrence, also called event, describes the topic's link to physical resource, the occurrence type indicates the resource type linked by the topic. One topic can correspond to multiple occurrences. From the overall structure view, TM domain is divided into topic domain and resource domain. The topic domain is a collection of topics, and it contains the association. The resource domain is a collection of all physical resources. Therefore, to organize agriculture model base resource using the topic map technology, it is needed to extract the concept term, interface specification and the relationship from the model component library, and further maps to the topic and association, maps the database document and component entity etc., to the resource domain. Figure 1 mapping between agriculture model-base resources and topic map elements

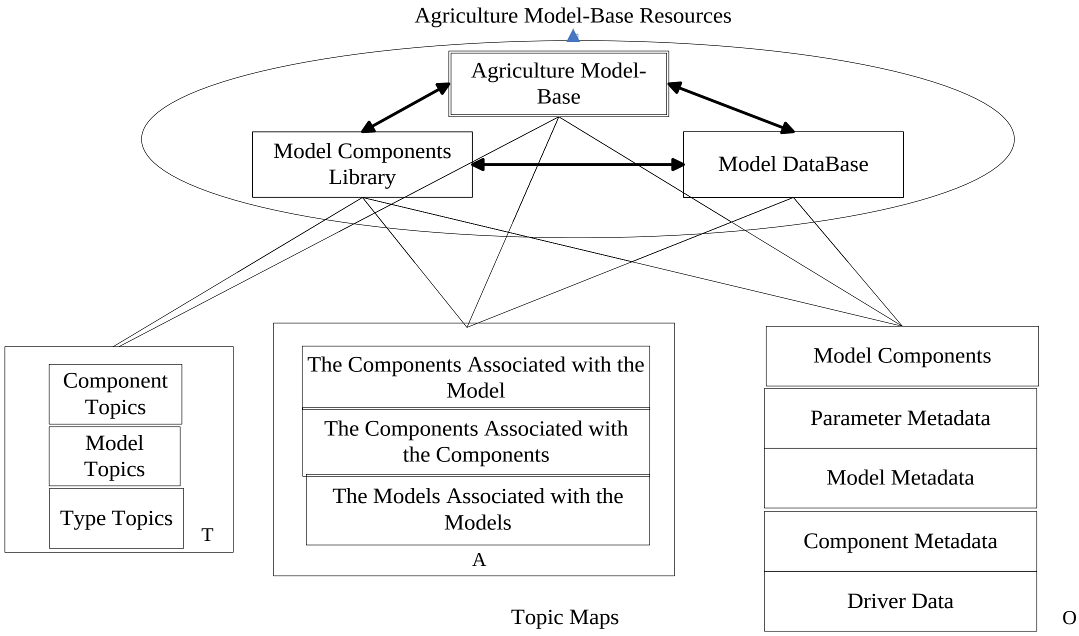

Fig.1. The mapping of agricultural model-base resources and topic maps

3 Description model of agriculture model

By using Jade Bird Component Model [15], the domain description tree and association description diagram for agriculture model component library are formed.

3.1 The component-based analysis on agriculture model size level

Agricultural model is composed of the model description, model algorithms and model structure three parts [16]. The agriculture model is composed of the logic model and simulation model, the logic model uses mathematical modeling tools to reflect the objective elements of things and their relationship. The simulation model is a software realization of the logic model in the computer, the mapping relation between the logic model and simulation model is one-many. The model description describes the model static characteristic in detail. Model algorithm and model structure describe the behavioral characteristics of agricultural objects. Model structure is a set of constraint relations between concepts of model components and sub-model components, reflects the combination rules between the model and the sub-model. Model algorithm uses a series of calculations and formulas to represent the model's solving process including the model input variables, output variables, state transition variables and transition rules and calculation equations, etc., is a concretion of abstract concept model, and can be mapped to model components. Among them, input / output variables, including model parameters and the driving variable, reflect the temporal and spatial characteristics of driven model computing. For example, field scale crop growth models influenced by weather, soil, varieties and cultivation conditions, the daily temperature value are driven variable of running model. In the application process at the regional scale, crop production and management knowledge model needs to interpolate and process the daily temperature data of simulation regional meteorological stations to be used as driven variable of field scale model, this reflects the spatial characteristics of model input variables. On the whole, agriculture model is often a complex model, and it involves multiple sub-model, the input / output variables is complex and is of significant spatial and temporal characteristics. Therefore, the component size level is associated with the model size level. The sectional capacity of model can be mapped to a combination of ability of model components.

According to the characteristic analysis of component and agriculture model size level, it is known that agriculture model components are composed of atomic model components, composite model components and model framework component, the specific algorithm can be. Net components, JavaBean or web Service and other component entities. Atomic model component is a model algorithm component to complete single-function, and its internal is composed of class, conditions and equations. For example, crop growth model component has only one single-function for calculating the crop growth phenology, belongs to business components. Composite model component can guide the assembly, calls and combinations of sub-model component , realize the specific functions of application, belongs to application or service component. For example, field scale crop growth model is composed of a number of sub-model components, can complete the different levels of field scale crop systems simulation, it belongs to system-level model components. Model framework components can guide atomic model components to form composite model components.

3.2 Component-based description of the agricultural model

The semantic and grammar information of agricultural model can be expressed by the description model of the agriculture model component library (DM-AMCL). DM-AMCL is composed of Facet Description Model (FDM) and Associate Description Model (ADM).

3.2.1 Facet Description

FDM describes the model components concept space faceted terms, and is composed of model property, model parameters, the model framework and the resource files of four domains (Formula 1).

$$FDM = \{ModelExplain, ModelFramework, ModelParameter, ResourceFiles\} \quad (1)$$

Each domain contains a number of key term nodes, the key term node is divided into many groups according to the model and component, "domain + model group" and "Domain + component group" are seen as an abstract object, the characteristic facet terms of these abstract objects are property terms, which form 8 facet term trees in domain unit. Table 1 lists the FDM classification domain and property term contents. Among them, ModelExplain.ME and ModelExplain.CE are separately the descriptive property information of model and component. ModelParameter.MF describes input and output parameter's semantic information of a single atomic model. ModelParameter.CI describes syntax specification of atom model component interface. ModelFramework domain specifies a set of portfolio reference standards of composite model, and sub-model components interface adapter standard, and can guide the assembly of sub-model components. Among this, ModelFramework.δ describes composite model component hierarchy, and the combination or collaboration rules among components. ModelFramework.λ describes the interactive processes, rules and interface specification among model components composed by composite model. For example, crop growth model is a composite model, contains sub-calculation models of growth period, biomass, dry matter allocation, organ built, quality formation, soil - crop water, soil - crop nutrition etc. in its ModelFramework.δ . Wheat, rice and other crop growth model has similar simulation framework, and can simulate different levels of cereal crop systems through a combination of sub-models. Further use of component interface and component portfolio reference standard in the ModelFramework.λ, can complete the assembly of sub-model components and construct the system-level model components.

Table1. The Facet Description Model Details Of Agriculture Model Component Library.

Faceted classification domain	Property glossary
ModelExplain.ME	model number, model name, plant type, model class, model type, model scale, simulation level, functional description, application fields, modeling agencies, etc.
ModelExplain.CE	model number, model name, plant type, model class, model type, model scale, simulation level, functional description, application fields, modeling agencies, etc.
ModelParameter.MF	model number, model name, plant type, model class, model type, model scale, simulation level, functional description, application fields, modeling agencies, etc.
ModelParameter.CI	component number, component name, component size, encapsulation forms,

	representation forms, development languages, application fields, deployment environment, instructions, etc.
ModelFramework.δ	model component sequence, combination or collaboration rules, reference input / output, etc
ModelFramework.λ	composable interface sequence, call interface rules, reference Interface Specification, etc
ResourceFiles.MR	model description file,etc
ResourceFiles.CR	component description file, component entity file,etc

3.2.2 Association Description

Although the FDM can describe the semantic and syntactic information of agricultural model faceted term space, but there is no associations between objects of the domain. By refining the association between model objects and component objects in the same domain, this study associated the model concept term space with the component concept term space and physical resource, and formed the association description model (ADM). ADM includes Atom Association Description Model (AADM) and Complex Association Description Model (CADM). AADM is composed of model - component properties, the model parameters - component interface, model attribute - model description file, component attributes – component file, and composite model structure – component combinations statute, a total of five kinds of association types. Formula (2) reflects the mapping relationship set involved in atomic model, each mapping is of one-to-many relationship. Formula (3) reflects that the composite model has added the part of association on the basis of atomic component mapping relationship set, and the association of complex model structure - component combinations statute has a one-to-one relationship. Integrating FDM and ADM can describe semantic and grammatical information of agriculture model base totally.

$$ADM = (\Pi_{ME \rightarrow CE}, \Pi_{MP \rightarrow CI}, \Pi_{\delta \rightarrow \lambda}, \Pi_{ME \rightarrow MR}, \Pi_{CE \rightarrow CR}) \quad (2)$$

$$AADM = \{\Pi_{ME \rightarrow CE} \cap \Pi_{MP \rightarrow CI} \cap \Pi_{ME \rightarrow MR} \cap \Pi_{CE \rightarrow CR}\} \quad (3)$$

$$AADM(i) \cap \Pi_{\delta \rightarrow \lambda} \rightarrow CADM \quad i \in atomic\ model, 1, 2, \ldots, n \quad (4)$$

4 Topic map framework oriented Agriculture Model Component Library

The construction of topic maps includes extraction, annotation and merging of topic map elements, and saves it using XTM file storage.

4.1 Topic map formation

By using the information of DM-AMCL Faceted description domain and association description domain as metadata, combined with the specific model types and component models, and entity file, this study formed themes, associations and events of agriculture model component library, and established the Topic Maps of Agricultural Model Component Library (TM-AMCL) by using XTM representation. It consists of model’s meta topic maps (MTM), atomic topic maps (ATM) and complex topic maps (CTM) together, and followed by a meta-data - instances of relations, and a total share of model information is realized by using topic map fusion algorithm. MTM is the basic mode for generating other thematic maps, and come directly from DM-AMCL . It is composed of meta-attributes for topic maps (AMTM) and meta-framework for topic maps (FMTM), reflects the hierarchical relationship between the generic concepts of domain model, component domain and resource domain in the model component library, and the association relationship between the domains. MTM is the abstract representation of concept and relationship of meta-model of component library, composed of local atom topic maps (LATM) and global atomic topic maps (GATM), and can reflect the different model component entity concept and association information in same type of model. Each model component corresponds to an LATM, the same type of model has more than LATM, and GATM is formed by using the topic maps merging technology, the global component level share of similar model is realized. CTM is the abstract representation of concept and relationship of composite model components, and can fully reflect sub-models involved in the composite model, and combination rule of sub-models etc., the global model-based system level share is realized.

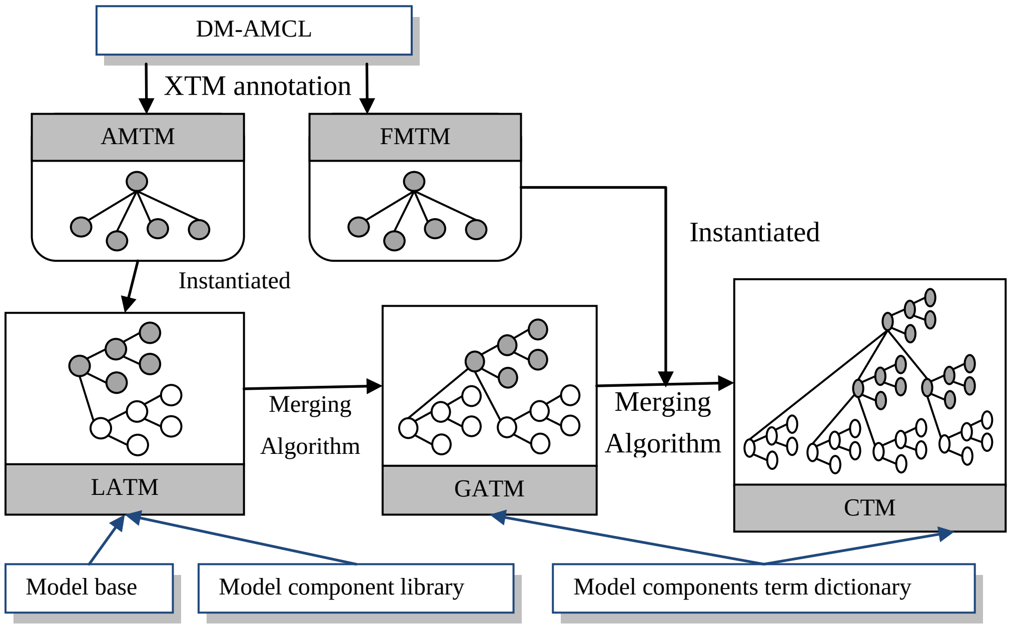

Fig.2. The structure and the merging process of TM-AMCL

Figure 2 show that TM-AMCL in turn are divided into three steps: ①MTM reflects "meta-data relationship", maps the faceted classification, terminology and association types of DM-AMCL to AMTM topic type, association type and event type separately, all represented by the topics without the relation of specific model types. The term value in the Model Framework domain is mapped to the topic type of FMTM, is related to specific model types. One type model has a set of reference standard of common model framework. ②AMTM reflects "abstract - concrete - merging relations". Based on the meta-property topic map mode, all instances of the topic are extracted from the model component database, these instances are seen as new topic and are added to the MTM structure to form new LATM. Further, under the guidance of model components term dictionary, the LATM is merged to GATM by using topic map merging algorithm, and the different crop model algorithm components of the same agriculture model type is shared. ③CTM reflects "model composable relationship". Based on the meta-framework topic map mode, the GATM is classified and merged to form CTM, and the system-level model information is shared. In figure 2, model components term dictionary defines the related field terms and their synonyms in agriculture model component library, including number, the term name, symbol, synonyms, description and scope of use. For example, the dictionary of growth stage model includes ‘M001, development model. DEV, growth development and phenology, "Simulation of crop growth and development, computing phenological time", crop model’.

4.2 XTM Annotation Of Topic Maps

The study used the XTM elements such as <topic>, <association>, <mergeMap>, <topicRef>, <scope>, <instanceOf>, <topic> and <baseName> etc., to notate the faceted and the associated term in the DM-AMC, notated the model component metadata according to level name mode[16], notated the ATM by using classified and instantiated method, and formed the agriculture model XTM file. The figure 3 list the Model Explain domain, model-component attribute association domain and the XTM fragment for AMTM document. Other facet description domain can refer to the method to notate, because it is difficult to describe in here.

```
<?xml version="1.0" encoding="gb2312"?>
<topicMap>
   <topic id=" 1_ModelExplain ">
      <baseName> <baseNameString> Model Description </baseNameString> </baseName>
   </topic>
   <topic id="2_ME. ModelID ">                <! Model Number >
   <instanceOf> <topicRef xlink:href="#1_ ModelExplain "/></instanceOf>
   <baseName> <baseNameString>2_ME/ModelID</baseNameString></baseName>
   <baseName> <baseNameString>ModelID</baseNameString></baseName>
   </topic>
<topic id="2_ME. ModelType ">                <! Model Type >
<instanceOf><topicRef xlink:href="#1_ ModelExplain "/></instanceOf>
   <baseName> <baseNameString>2_ME/ModelType</baseNameString></baseName>
   <baseName><baseNameString> ModelType </baseNameString> </baseName>
   <scope>
   Crop model, meteorological model, soil model, hydrological model, economic model, plant protection model
   </scope>
   </topic>
```

Fig.3. The XTM fragment of AMTM

4.3 A working example：The wheat development model component’s LATM building

The research team saved the descriptive information of the wheat growing model components in the agricultural model component database, wrote and used C # programs to access Microsoft SQL Server database, and according to the XTM standard format, converted the field data of database table to the XTM elements and property information, such as < topicMap>, < instanceOf>, < topicRef > and < Scope>. Then, by using the leaf node faceted topics of MTM as the topic type of LATM, the appropriate topic instances are established. Each topic includes the base name and extension name. Base name is the basic name of the current topic, and extension name records the topic path information in the MTM. If the current topics are in the ModelExplain domain, then its extension name is the current model number / base name. The other topic name definition is similar. Figure 4-6 separately lists the computer-generated model class topic, model number and component number of the wheat growing model components, and XTM description fragment of component entity file information, and among this, the wheat growing model number is M001, its corresponding software component number is C001, and its corresponding binary entity file name is CDeve-Wheat_1.1.3.dll.

```
<topic id="M001/ModelClass">
  <instanceOf>
    <topicRef xmlns:xlink="#2_ME/ModelClass" />
  </instanceOf>
  <baseName> <baseNameString>Development Class</baseNameString> </baseName>
  <baseName> <baseNameString>M001/ Development Clsaal </baseNameString> </baseName>
</topic>
```

Fig.4. The topic of wheat development model class for the M001

```
<association id="M001_C001_association">
  <instanceOf> <topicRef xmlns:xlink="#100_Model-Component Explain Associate" /> </instanceOf>
  <member> <roleSpec> <topicRef xmlns:xlink="#Model_Role"/> </roleSpec>
    <topicRef xmlns:xlink="#M001/ModelID">
  </member>
  <member> <roleSpec> <topicRef xmlns:xlink="#Com_Role"/> </roleSpec>
    <topicRef xmlns:xlink="#C001/ComID">
  </member>
</association>
```

Fig.5. The association between the M001 and C001

```
<topic id="C001/ComID">
<occurrence>
    <instanceOf> <topicRef xmlns:xlink="#ComFile_4"/> </instanceOf>
    <resourceData>CDeve-Wheat_1.1.3.dll</resourceData>
</occurrence>
</topic>
```

Fig.6. The occurrence of component entity for the C001

5 Conclusion

This study proposed a new approach for agriculture model description and representation with coordination of software components and topic maps. initially converted the agriculture model-base to topic map library, formed the data document by using XTM semantic notation.

(1) Combined with Jade Bird Component Model, the study formed DM-AMCL that is metadata of TM-AMCL. The faceted description is four domains. Associated description includes five kinds of association types. The unified description of agriculture model base information is initially realized.

(2)The organization framework and build process of TM-AMCL are raised. The DM-AMCL information is mapped to MTM, the LATM is created through the instances of DM-AMCL, the GATM and CTM are formed by using emerging algorithms. The study supply reference mode to different agricultural models developed by using different technologies in the network environment.

(3) The agriculture model is converted to the XTM document data, and provided the network information exchange means and the uniform access to data for developers of intelligent decision support system, managers of agriculture models, agricultural research organizations, and the computing system.

Acknowledgments

We acknowledge the funding of this work by National Natural Science Foundation of China (No:30971697), Science and Technology Support Project Plan of Jiangsu-Agriculture section (No:BE2009342) and Social Science Foundation project of Nanjing Agricultural University in China (No:SK08017).

References

[1] GAO L.Z. Foundation of Agricultural Modeling Science. Hong Kong：Pegasus Books Ltd Publisher (2004)

[2] Jouven M., Baumont R.. Simulating grassland utilization in beef suckler systems to investigate the trade-offs between production and floristic diversity. Agricultural Systems,2008, (96)，260-272

[3] Jonesa J.W., Keatingb B.A., Portera C.H.. Approaches to modular model development. Agricultural Systems ,2001,(17),421-443

[4] Argent R.M. An overview of model integration for environmental applications-components, frameworks and semantics. Environmental Modelling & Software, 2004, 19, 219-234.

[5] Smith F.P., Holzworth D.P., Robertson M.J.. Linking icon-based models to code-based models: a case study with the agricultural production systems simulator． Agricultural Systems, 2005, 83, 135-151．

[6] van Ittersum M, Ewert F, Thomas H, Jacques W, ect. A component-based framework for the European Union (SEAMLESS). Agricultural Systems,2008, 96, 150–165

[7]Zhao C.J.,Wu H.R.,Yang. Development platform for agricultural intelligent system based on techno-componentware model. Transact ions of the CSAE,2004,20(2),140-143(in Chinese).

[8] Zhu Y，Cao W.X.，Wang S.H.，Pan J. Application of soft component technology to design of intelligent decision-making system for crop management. Transact ions of the CSAE, 2003，19(1), 132-136(in Chinese).

[9] Cao W.X., Zhu Y, Tian Y.C., Yao X, Liu X.J． Research progress and prospect of digital farming techniques. Scientia Agricultura Sinica, 2006, 39(2), 281-288. (in Chinese)

[10] ISO/IEC13250: Topic Maps–Information Technology, Document Description and Processing Languages. http://www.y12.doe.gov/sgml/sc34/document/0129.pdf

[11] Ma W.F.，Du X.Y. Digital Resources integration：theory,method and application. Beijing Library Press Publisher(2007)

[12] Ellouze N, Lammari N, Métais E,etc. Usage-Oriented Topic Maps Building Approach. Metadata and Semantic Research, Springer Berlin Heidelberg.2009, 13-23

[13]Wu X.F., Zhou L, Zhang L,etc. Knowledge warehouse based on XML topic maps. Engineering Journal of WUHAN University , 2006,39(1),pp.121-125

[14]Herzum,P.,Sims,O. Business Component Factory:A Comprehensive Overview of Component-Based Development for the Enterprise. Machinery Industry Press Publisher(2005)

[15] Chang J.C., Li K.Q. Representing and Retrieving Reusable Software Components in JB (Jadebird) System. ACTA Electronica Sinica, 2000, 28(8), 20-24. (in Chinese)

[16] Jiang H.Y., Zhu Y,Xu H.L.,etc. Development of crop model resource construction platform. Transactions of the CSAE,2008,24(2), 170-175.