Research on the Construction and Implementation of Soil Fertility Knowledge Based on Ontology

Li Ma, Helong Yu, Guifen Chen, Liying Cao, Yue Wang

HAL Id: hal-01220822
https://hal.inria.fr/hal-01220822

Submitted on 27 Oct 2015

Research on the Construction and Implementation

of Soil Fertility Knowledge Based on Ontology

Li Ma, Helong Yu, Guifen Chen, Liying Cao, Yue Wang

College of Information and Technology Science, Jilin Agricultural University,

Chang Chun, Jilin, China

mary19801976@sohu.com

Abstract.Soil fertility is the comprehensive reflection of related factors and the related factors. Soil fertility evaluation knowledge is stored by relational database as usually, and it is difficult to show the correlation and constraints among attributes .In this paper, Nongan county farmland productivity data is as the research object, Using rough set approach to do attribute reduction, using ontology method to establish the soil fertility level knowledge base, using multi Agent technology to implement the prototype system, and complete the reuse and sharing of knowledge, lay the foundation for semantic level reasoning.

Key words: data mining; ontology; soil fertility; multi-Agent

1   Introduction

The cultivated land fertility is integrated by soil characteristics, natural conditions and farm management and other elements of the productive capacity of cultivated land, is the comprehensive reflection of related factors and the influence of the nature of the soil properties [1-3]. In soil fertility evaluation, the description of the knowledge and rules is mostly used relational database. And is less considering the relationship among properties of the soil, and there are no sound constraints on attributes, Can't describe the implicit relationships between concepts, Cannot carry out the sharing and reuse of the data, cause a large amount of data redundancy. Rough set can reduce the redundant attribute in database under the premise of guarantee the same classification and decision, simplify the knowledge representation, and improve the efficiency of system processing [4]. Ontology is a new method of data description, Can be a clear description of concept in concept level, and expression association and constraints between concepts [5].

   In this paper, on the basis of the research of data mining and ontology technology, the cultivated land of NongAn of Jilin province is as the research object, make extraction and classification of the knowledge in the field, use of ontology technology for knowledge representation, Construct of soil fertility of ontology library, and realize knowledge sharing using Multi-agent Technology, present a new method for data storage.

2 The Research Data and Related Technologies

2.1 The Data Source

The data this paper used is from the cultivated land fertility survey data, the NongAn (2006), offers by agricultural technology extension center of NongAn. The data includes 25 attributes, such as soil humidity, groundwater depth, light radiation intensity, soil irrigation capacity, annual rainfall, soil drought resistance and soil erosion degree, soil texture, crop rotation suitability, topography, soil parent material, part into layer thickness, salt concentration, humus soil pH value, effective copper, iron, effective slowly available k, effective k, effective fierce, total nitrogen, phosphorus, organic matter and cationic content, effective zinc and productivity grade. Some data are shown in table 1.

Table 1. Some fertility data

ground water depth	soil irrigation capacity	annual rainfall	soil drought resistance	soil texture	soil parent materia	part into layer thickness	salt concen- tration	pH value	effective copper	effective iron
3-5m	no	400-450mm	strong	Light clay	alluvial	10-20cm	<0.1	6.6	1.25	8.72
<3m	strong	400-450mm	strong	Light clay	alluvial	10-20cm	>0.1	6.7	1.26	8.47
3-5m	strong	>450mm	strong	loam	alluvial	10-20cm	<0.1	6.7	1.39	4.70
3-5m	strong	400-450mm	strong	sandy loam	alluvial	10-20cm	<0.1	6.5	1.21	9.45
<3m	strong	400-450mm	strong	sandy loam	alluvial	10-20cm	<0.1	6.7	1.27	9.71
5-8m	no	400-450mm	weak	sandy loam	loess	0-10cm	<0.1	6.6	1.20	8.30
3-5m	no	400-450mm	weak	sandy loam	diluvial	20-30cm	<0.1	6.6	1.30	8.50
<3m	strong	400-450mm	strong	sandy loam	alluvial	10-20cm	>0.1	6.7	1.28	8.48
<3m	strong	400-450mm	strong	sandy loam	alluvial	10-20cm	>0.1	6.7	1.30	0.10
3-5m	no	400-450mm	strong	sandy loam	alluvial	10-20cm	<0.1	6.6	1.27	8.28

2.2 The Rough Set Theory

The rough set (RS) theory is a new mathematical tool which can process fuzzy and uncertainty knowledge, its characteristic is that it does not need to assign quantity descriptions of some characteristics and attributes in advance, starts from the description of given problems, discovers the inherent laws, its basic philosophy is closer to the realistic situation[6].

The basic philosophy of is rough set is that $S=(U, A,\{V_a\}, a)$ is called the knowledge expression system, and U is a non-spatial finite set，called the universe ; A is a non-spatial finite set, called attributes set; V_a is the range of $a \in A$, $a:U \rightarrow V_a$ is a injective maps. If A composes of condition attribute set C and conclusion attribute set D, C and D satisfy $C \cup D=A$，$C \cap D = \Phi$, then S is called the decision system.

In a decision system, the dependence or the connection are existence in certain degrees between each condition attribute, the reduce may be considered that under the premise of losing no information ,using simple description to express the dependence and connection between conclusion attributes to condition attributes in a decision system(Z.Pawlak，1995). Indiscernibility relation ind(C) divide U into t indiscernibility classes $X1$，$X2$，…，Xt, makes $D(Xi)$ is a set of All values of conclusion attribute d of Xi, that is $D(Xi)=\{v=d(x):x \in Xi\}$, if $D([Xi] ind(C-\{a\}))=D(Xi)$，then said that the condition attribute $a \in C$ can be removed compared with indiscernibility classes Xi. $C' \subseteq C$ is called   the reduce of   C relative to indiscernibility classes Xi , if $a \in C '$, then a cannot be   removed relative to Xi.All reduce sets compared with Xi is recorded SRED(C, Xi), Score(C，Xi)=∩SRED(C，Xi) is called core of Xi.

2.3 The Ontology

The concept of ontology is from the field of philosophy. In the 1960 s, computers started to use ontology. At present, the definition of ontology is a shared standard conceptual model explicit formal specification[7].

As to how to use ontology to organize knowledge, Perez et al. summed up the 5 basic modeling primitives using classification: class, relation, function, the axiom and examples.

(1) classes or Concepts. Classes can be any transaction, such as job, function, behavior, strategies, and reasoning process, and so on. Semantically, the meaning of class is a collection of objects, Its definition generally adopts frame structure, including the name of a concept, with the rest of the concept of the relationship between the collections, as well as in natural language description of the concept, etc.

(2) relations. Formal definitions for the n dimension of a subset of the Cartesian product: $R：C1 \times C2 \times … \times Cn$. For example, the subclass relationship (subclass-of), which represents the interaction between domain concepts.

(3) functions. Function is a kind of special relationship. The formal definition of F is $C1 \times C2 \times … \times Cn-1 \rightarrow Cn$ .The n element can be uniquely determined by n-1 elements

before of the relationship, such as is-a is a function, is-a (m, n) means that n is an instance of m.

(4) axioms. Axiom represents the eternal truth assert, is defined in the "concept" and "property" on the limit and rules, such as the concept of A belongs to the scope of the concept of B.

(5) the instance (instances). Instances represent elements. Semantically representation is the object, is a concrete entity refers to a concept class.

3 Construction of Soil Fertility Ontology

3.1 Attribute Reduction in Rough Set

This paper uses the genetic algorithm for attribute reduction, the process is: (1) A randomly generated population. (2) Evaluation of the merits of each chromosome, then choose the excellent chromosome, format a new species. (3) Operators (crossover and mutation) to the new population genetic, then get new specie. (4) Repeat the genetic operation chose the best chromosomes as a solution [8].

According to rough set theory, data set is divided into attribute set and decision set. Attribute set includes 24 attributes, those are soil humidity, groundwater depth, light radiation intensity, soil irrigation capacity, annual rainfall, soil drought resistance and soil erosion degree, soil texture, crop rotation suitability, topography, soil parent material, part into layer thickness, salt concentration, humus soil pH value, effective copper, iron, effective slowly available k, effective k, effective fierce, total nitrogen, phosphorus, organic matter and cationic content, effective zinc. Decision set contains 1 attributes that is fertility level.

Using genetic algorithms to do the reduction, attribute set contains soil humidity, groundwater depth, light radiation intensity, soil irrigation capacity, annual rainfall, soil drought resistance and soil erosion degree, soil texture, crop rotation suitability, topography, soil parent material, part into layer thickness, salt concentration, humus soil pH value, iron, effective slowly available k, effective k, total nitrogen, phosphorus, organic matter and cationic content, A total of 21 properties. Reduces 3 attributes including effective copper, effective fierce, effective zinc The decision set contains an attribute {soil fertility level}.

3.2 Construct The Soil Ontology

Using the seven step of ontology construction as the guidance, taking fertility evaluation based on the series of elements, Soil data based on the actual situation, according to the actual situation of the soil data, establishing the ontology classes and attributes. Soil ontology construction process is as follows:

Define classes and class rating system: (1)to establish the section and physical and chemical properties of physical attributes of, ph value, soil moisture, salt concentration, humus layer thickness, soil texture, cationic content and groundwater

depth; (2)The site conditions, including attribute of site topography, soil erosion degree and parts into soil organic; (3)Meteorological conditions of weather, contain attributes rainfall and light radiation intensity; (4)Establish the nutrient content of nutrient, contain attributes effective iron, slowly available k, effective, total nitrogen, phosphorus, effective organic matter; (5) Establish the soil management class management, including attribute soil drought resistance, irrigation and crop rotation suitability; (6)Set up the soil, the soil as a kind of all kinds of the father; (7)Established the kind of equipment, including productivity grade attributes.

Define a relationship class_is, used to determine a specific example belonging to which productivity grade, its domain of definition is soil, its range is class.

The exact makeup of soil ontology is：

Class soilclass｛

　　　Class physical （ph， moisture， salinity， humus， textrue， cation， groundwater）

　　　Class site（position，erosion，material）

　　　Class weather（rainfall，light）

　　　Class nutrient（fe，slow k，k，n，p，om）

　　　Class management（irrigation，drought，crop）

　　｝

Class class（attribute is soil_class）

Class relations attribute：class_is

The domain of definition of these attributes including pH，fe，slow k，k，n，p，om，cation is soil，and the range is int. The domain of definition of these attributes including moisture，groundwater，light，irrigation，rainfall，drought，erosion，textrue，crop，position，material，humus，salinity, is soil，and the range is float. The domain of definition of soil_class is class，and the range is int. Ontology construction result is shown in figure 1.
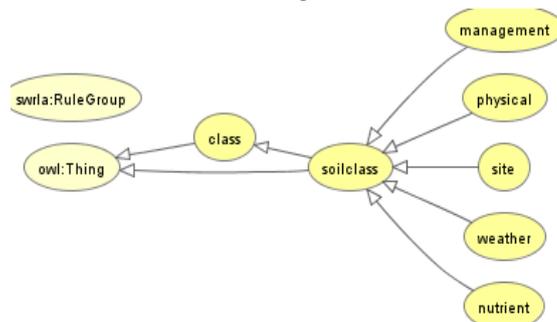

Fig.1. Ontology Construction

4 Realization System Model

Multi - Agent technology with the development of distributed artificial intelligence can effectively solve the problem of distributed data integration. This paper uses the Agent oriented software development method， designs and develops data integration and exchange system for the soil fertility data. System development tools is VC++, it is used to achieve the Agent procedures for the preparation and encapsulation. The main function of Agent consists of information exploration of Agent, Agent, intelligent evaluation task decomposition and scheduling Agent, ontology evaluation Agent and ontology information processing Agent and Ontology learning Agent [9]. Using KQML as the message protocol Agent interaction, Enables Agent to exchange knowledge and information to other Agent and Agent running environment. System model is shown as figure 2
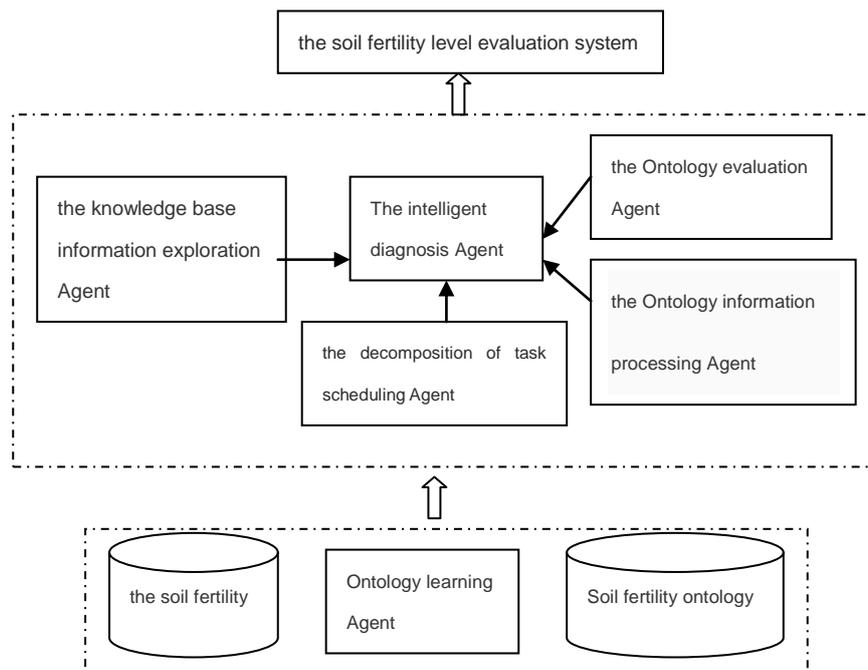
Fig.2. System Model

Soil fertility level evaluation process can be described as: User submitted data remotely through the client, after the system receives the data, the system preliminary analysis, using the evaluation model for data that is stored in advance，if no match is found in the model, then uses of ontology learning Agent, to study the corresponding evaluation model. Task scheduling of Agent will evaluate the task decomposition,

Uses intelligent evaluation Agent call the evaluation module algorithm, then ontology evaluation evaluates the algorithm Agent, Or the ontology information processing Agent deals with the corresponding evaluation results. The final conclusion is reached, and the evaluation results will be output to the user.

5 Summary

In this paper, we study the rough set theory and ontology technology, propose a new method to describe the soil fertility data, realize the sharing and reuse of heterogeneous system knowledge. Using rough set approach to attribute reduction of data, can remove redundant attribute data, reduced data set; Using ontology can make the concept in the field of pattern be understood more accurately, can more clearly express the relationships and rules among the concepts. In this paper, the application of multi Agent technology, realized the sharing of knowledge of soil fertility preliminary. Future research will combine the multi platform interoperability, in-depth development and establish standards ontology of the application domain, and further promote the knowledge sharing and information integration.

Acknowledgment

References

[1] CHEN Gui-fen,MA Li,DONG Wei,XIN Min-gang.The combination of clustering and rough set and decision tree algorithm applied in the evaluation of soil fertility[J]. Scientia Agricultura Sinica,2011,12.
[2] Ma Li.Research on Rough Set and Decision Tree Application in Evaluation of Soil Fertility Level[D]. Jilin agricultural university,2010.
[3]Tian Jian Hu Yueming, Wang Changwei, et al. Clustering support decision tree model in the evaluation of cultivated land application [J]. Journal of agricultural engineering, 2007, 23 (12) : 58-63.
[4]Yanbo Huang, Yubin Lan, Steven J. Thomson. Development of soft computing and applications in agricultural and biological engineering[J]. Computers and Electronics in Agriculture. 2010 (71) :107–127.

[5]Ma Li,Chen Guifen.The Knowledge Representation and Semantic Reasoning Realization of Productivity rade Based on Ontology and SWRL[J], Computer and Computing Technologies in Agriculture,2011,11

[6]Chen Guifen. For the spatial data mining technology research and application of precision agriculture [D]. Jilin university. 2009.

[7]Du Xiaoyong, lee, king of bashan. Ontology learning research review [J]. Journal of software, 2006, 12:1837-1847.

[8]Cui Guangcai, Yin qiao true. A kind of attribute reduction algorithm based on genetic algorithm [J]. Journal of changchun university of science and technology, 2003, 9:4 ~ 7

[9]Li Chune Chen Quangong. Multi Agent technology applications in disease diagnosis system [J]. Computer engineering, 2008 (33) : 182-184