Rules mining in hierarchical rough sets

Wei Yang, Duoqian Miao, Qinrong Feng
Department of Computer Science and Technology, Tongji University
Shanghai, P.R. China 201804
yangw312@yahoo.com.cn, miaoduoqian@163.com, fengqr72@163.com

Abstract

A novel algorithm is proposed in this paper, which can process the higher level of conditional attribute and the decision attribute at the same time. When building the higher level decision table for simplifying the decision table based on concept hierarchies tree, the inconsistent objects may be appeared because of exchanging lower level attributes values with higher level attributes values. For solving this problem, we propose that the attribute values in different levels may be used in a decision table. In order to substantiate the conceptual arguments numerical examples are given.

Keywords: Rough set; Multi-hierarchy; Reduction; Rules mining;

1. Introduction

Zadeh[1-3] proposed the information granularity and fuzzy information granulation. The idea of information granulation has been applied many fields such as rough set, fuzzy set and so on. Several relative research works on granular computing have been proposed [4-8]. Granular computing is a world outlook and methodology. Rough set theory was proposed by Pawlak [9] in 1982, which is used as a mathematical tool to deal with imprecise, vague and uncertain data sets in many fields. Several extensions on rough set have been proposed [10-22].

In granular computing, data are analyzed and processed from multi-view and multi-level. Many real-world applications have the hierarchical value such as day, month, quarter and year. In the traditional Rough set theory or its generalized model, data are processed only on a single level of attribute domains. In essence, there exists hierarchical structure impliedly in every attribute domain. So, the traditional rough set can not process data at multiple abstract levels. The data mining in hierarchical rough set [23] is proposed, which introduced one algorithm to mine the rules in different levels.

In this paper, a novel algorithm is proposed for the hierarchical rough sets. The algorithm may give different levels rules, and can process the conditional attribute and the decision attribute at the same time. Without using the multidimensional data model [23], we not only change the lower level attribute valve into higher attribute value, but also change the higher level attribute valve to lower attribute value when the inconsistent objects are appeared. So, the attribute values in different levels may be used in a decision table. For a consistent decision table, this algorithm can give the consistent decision rules.

The rest of this paper is organized as follows. The concept hierarchy and some notions of hierarchical rough sets [23] are introduced in Section 2. The algorithm is presented and illustrated with an example in Section 3. The whole paper is summarized in Section 4.

2. Preliminaries

In this section, concept hierarchy and some notions of hierarchical rough sets [23] are introduced.

Let $\zeta=(U,AT,V,f)$ be an information system, where U is a non-empty finite set of objects, called the universe set, AT is a non-empty finite set of attributes, $V=\bigcup_{a\in AT}V_a$, where V_a is the value set of the attribute a, f is an information function from U to V.

Given a decision table $\zeta=(U,AT\bigcup D,V,f)$, where D is a decision attribute, and AT is called conditional attributes set.

For each attribute in a decision table, correspond to a concept hierarchy tree with the prior knowledge. For each concept hierarchy tree, each level can be labeled by a number from leaf node to root node.

Example 1. Table 1 is a decision table. The concept hierarchy tree can be given for each attribute in fig.1.

Given a decision table $\zeta=(U,AT\bigcup D,V,f)$, $A_1,A_2,\ldots A_m\subseteq AT$, all concept hierarchies tree of

978-1-4673-2311-6/12/$31.00 ©2012 IEEE

4. Summary

In this paper, for processing the higher level of conditional attribute and the decision attribute at the same time, an algorithm is provided. When building the higher level decision table, the inconsistent objects may be appeared for exchanging lower level attributes values with higher level attributes values, and we propose that the attribute values in different levels may be used in a decision table. So, when the decision is a consistent decision table, this algorithm can give the consistent decision rules. In order to substantiate the conceptual arguments numerical examples are given.

Acknowledgements.

The research is supported by the National Natural Science Foundation of China under grant No: 60970061,61075056.

5. References

[1] Zadeh L A. Fuzzy sets and information granularity, Advances in Fuzzy Set Theory and applications, M. Gupta, R. Ragade and R. Yager (eds.), Amstrdam: North-Holland publishing, 1979,3-18.

[2] Zadeh L A. Fuzzy logic = computing with words. IEEE Transaction on Fuzzy System, 1996,Vol.4, Issue 2, pp.103-111.

[3] Zadeh L A. Towards a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. Fuzzy Sets and Systems, 1997, 90:111-121.

[4] T.Y. Lin, Granular computing, announcement of the BISC special interest group on granular computing, 1997.

[5] Yao, Y.Y., Granularity computing: basic issues and possible solutions, In: Proceedings of the 5th Joint Conference on Information Sciences, 2000, Vol I, pp. 186-189.

[6] Yao, Y.Y., The art of granular Computing, In: Proceedings of the International Conference on Rough Sets and Emerging Intelligent Systems Paradigms, 2007,pp.101-112.

[7] Zadeh L A., Granular computing and rough set theory, 2007, LNCS 4585, 1-4.

[8] LI D.G, Miao D.Q, Zhang D.X, Zhang H.Y, An Overview of Granular Computing, Computer Science, 2005,32(9):1-12.

[9] Pawlak, Z. Rough sets, *International Journal of Information and Computer Sciences*, 11, 341-356, 1982.

[10] Wojciech Ziarko, Variable precision rough set model, *Journal of Computer and System Sciences*, Volume 46, Issue 1, Pages 39-59, February 1993.

[11] Kryszkiewicz, M. Rough set approach to incomplete information systems, *Information Sciences*, 112, 39-49, 1998.

[12] Guoyin Wang, extension of rough set under incomplete information systems [J], *Chinese Journal of Computer Research and Development*, 39 (10) :1238-1243, 2002

[13] Leung Yee, Li Deyu. Maximal consistent block technique for rule acquisition in incomplete information systems [J], *Information Sciences*, 153 (1): 85-106, 2003

[14] Guan Yanyong , Wang Hongkai. Set-valued information systems [J], *Information Sciences*, 176(17): 2507 -2525, 2006

[15] Leung , Y. ,Fischer , M.M. , Wu , W. Z.and Mi , J. S.A rough set approach for the discovery of classification rules in interval-valued information systems, *International Journal of Approximate Reasoning*, 47 , 233-246, 2008.

[16] Qian, Y.H., Liang, J.Y. and Dang, C.Y. Interval ordered information systems, *Computer and Mathematics with Applications*, 56, 1994-2009, 2008.

[17] Duoqian Miao, Nan Zhang, Xiaodong Yue. Know Reduction in Interval-valued Information Systems, *Proceedings of the 8th IEEE International Conference on Cognitive Informatics*, pp.320-327, 2009

[18] Wu, Q. and Liu, Z.T. Real formal concept analysis based on grey-rough set theory, *Knowledge-based systems*, 22, 38-45, 2009.

[19] Miao, D.Q. and Hu, G.R., A heuristic algorithm for reduction of knowledge, *Chinese Journal of Computer Research and Development*, 36, 681-684, 1999.

[20] Wang, G.Y. ,Yu, H. and Yang, D.C. Decision table reduction based on conditional information entropy , *Chinese Journal of Computers*, 25, 759-766 , 2002.

[21] Yao, Y.Y. and Zhao, Y. Attribute reduction in decision-theoretic rough set models, *Information Sciences*, 178, 3356-3373, 2008.

[22] Skowron, A. and Rauszer, C. The discernibility matrices and functions in information systems, in: Slowinski, R. (Eds.), Intelligence Decision Support: *Handbook of Applications and Advances of Rough Set Theory*, Kluwer Academic Publisher, Dordrecht , 1992.

[23] Feng, Q.R, Miao, D.Q, Cheng, Y, Hierarchical decision rules mining, Expert Systems with Applications, 2010, pp.2081-2091.