Accepted Manuscript
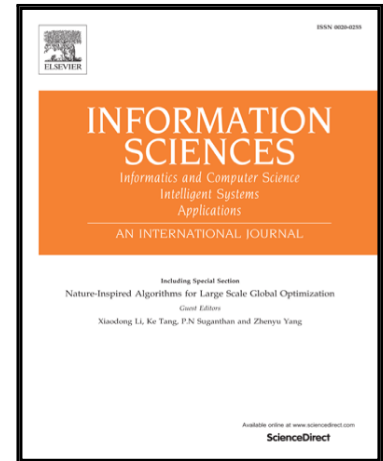
Semi-Monolayer Cover Rough Set: Concept, Property and Granular Algorithm

Zhengjiang Wu, Ning Chen, Yan Gao

Please cite this article as: Zhengjiang Wu, Ning Chen, Yan Gao, Semi-Monolayer Cover Rough Set: Concept, Property and Granular Algorithm, *Information Sciences* (2018), doi: 10.1016/j.ins.2018.04.066

Semi-Monolayer Cover Rough Set: Concept, Property and Granular Algorithm

Zhengjiang Wu[a], Ning Chen[a], Yan Gao[a]

[a]College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo, China 454003

Abstract

Rough set theory is known as a powerful tool for dealing with incomplete, inconsistent and insufficient information in decision making. With the constraint that each set of cover has at least one reliable element, the semi-monolayer cover (SMC) is a special neighborhood system defined on the universe between the partition and regular cover. More similar to Pawlak approximation operators than cover approximation operators, SMC approximation operators have some interesting properties, such as the equivalent granular forms, the cross duality, etc.

In this paper, we define four pairs of SMC rough approximation operators, namely GC, GU, DE and DA, and then discuss their properties and relations. It has been proven that (1) GU is the equivalent granular-form approximation operator of DE. (2) DE and DA are two pairs of cross dual approximation operators. That is to say that upper DE and lower DA, lower DE and upper DA are two pairs of dual approximation operators. The equivalent relation and cross dual relation serve as a bridge connecting granular-form approximation operators with the dot-form ones. (3) A DE or DA approximation set can be rebuilt by the corresponding GC approximation set and an SMC-Mapping. The equivalent granular form of DE and DA gives rise to some advanced algorithms that improve the efficiency of calculating DE and DA approximation sets. Experimental results on a UCI data set show the granular algorithm achieves better performance than the basic dot algorithm, and the speedup of efficiency tends to be exponential with the size of universe.

Keywords: semi-monolayer cover rough set, granular-form approximation operator, dot-form approximation operator, granular computing

Rough set theory, introduced by Pawlak as an extension of classic set theory [26], is known as a powerful tool for dealing with incomplete, inconsistent and insufficient information in decision making. For Pawlak rough set, the equivalent binary relation serves as the basis of the application space on the universe set. The information system can be formally divided into several categories corresponding to the equivalent relation. Rough set theory has abroad applications in a variety of domains including bioinformatics, signal and image processing, finance, medicine, multimedia, Web and text mining and so forth.

In all types of rough set theory, the lower and upper approximation operators play a vital role [1, 2, 5, 7]. To the best of our knowledge, there are at least three approaches to research approximation operators, namely constructive approach, axiomatic approach and basis algebra choosing approach. In those approaches, binary relation on a universe, neighborhood system, basis algebra, the form and the properties are four fundamental concepts to build the approximation operators [12, 23, 24] . Binary relation and neighborhood system are the result of distributing the elements of universe, such as partition, cover, etc. If the border of the equivalent class of partition extends from a line to a set, the distributing result will be a cover. A cover contains some elements which are named as covering sets in this paper. Semi-monolayer cover (SMC) proposed in this paper is a special case of cover. In SMC, each covering set has at least one reliable element. The reliable elements play an important role to identify a covering set from others in SMC and give SMC some potential to replace the partition. SMC generalizing from the partition, which is the direct result of the consistency in system, has abroad applications. The consistency is uncommon and sometimes expensive in a real system. The exorbitant cost forces us to accept the data with tiny inconsistence and mining information in it.

- (Scene I.) A low myopia person can not see a girl in the distance definitely, but he can still identify the some character, such as black hair, white skin and khaki wind-coat, and recognize his girlfriend accurately. However

Email addresses: `wuzhengjiang@hpu.edu.cn` (Zhengjiang Wu), `18839192353@163.com` (Ning Chen), `gaoyan@hpu.edu.cn` (Yan Gao)