

The Feature Core of Dynamic Reduct*

Jiayang Wang

College of Information Science and Engineering, Central South University, Changsha, China

Email: csuwjy@mail.csu.edu.cn

Received April 1, 2012; revised May 1, 2012; accepted May 8, 2012

ABSTRACT

To the reduct problems of decision system, the paper proposes the notion of dynamic core according to the dynamic reduct model. It describes various formal definitions of dynamic core, and discusses some properties about dynamic core. All of these show that dynamic core possesses the essential characters of the feature core.

Keywords: Decision System; Reduction; Dynamic Core

1. Introduction

KNOWLEDGE reduct is an important step in knowledge discovery, and also a favorable method to extract the more generalized rules. There are a number of research papers [1-3] about reduct problems, but most of them are just for static reduct.

The reduct methods based on standard rough set are effective to some extent, but there are also some problems to be solved in practice. Dynamic data, incremental data and noise data make the analysis results instable and uncertain. All of these limit the application of rough set theory [1].

Standard rough set methods are not always sufficient for extracting laws from decision system. One of the reasons is that these methods are not taking into account the fact that part of the reduct set is chaotic, in other words it is not stable.

Firstly, it gives the definition of dynamic core. Then the properties of dynamic core is certified. All of these show dynamic core possesses the essential characters of the feature core.

2. Dynamic Reduct

To the problems of standard rough set reducts or static reducts, dynamic reducts can put up better performance in very large dataset, and also enhance effectively the ability to accommodate noise data.

The rules calculated by means of dynamic reducts are better predisposed to classify unseen cases, because these reducts are in some sense the most stable reducts, and they are the most frequently appearing reducts in sub-

decision system created by random samples of a given decision system.

Definition 1: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $B = (U', C \cup D)$, $U' \subseteq U$, B is called the sub-decision system of S . Let $\rho(S)$ denote the set of all sub-decision system of S , $F \subseteq \rho(S)$ is called a F family of decision system S .

Definition 2: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $RED(S)$ denotes the set which contains all reducts of decision system S , and $RED(B)$ denotes the set which includes all reducts of sub-decision system B .

There is at least one reduct in a decision system, which is just itself, so the set of reduct is not empty. In many cases, there may be several reducts in a given decision system. Each reduct can product a rule set, and it is difficult to justify which is the best rule set. Therefore it is important to search the most stable reduct, dynamic reduct is proposed in this case.

Definition 3: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, dynamic reduct of decision system S is denoted as $DR(S, F)$ [4], and

$$DR(S, F) = RED(S) \cap \bigcap_{B \in F} RED(B).$$

Any element of $DR(S, F)$ is called an F -dynamic reduct, which describes the most stable reducts in decision system. From the definition of dynamic reduct, it follows that a relative reduct of S is dynamic if it is also a reduct of all sub-decision system from a given family F by random sampling.

*The project supported by National Natural Science Foundation of China, Grant No. 61173052.

The project supported by Hunan Provincial Natural Science Foundation of China, Grant No. 11JJ5040.

The reducts in a decision system are not stable, sensitive for sample data [5]. Bazan gives the concept and method about dynamic reduct [4], which grounds the most stable reduct of decision system in theory, then the dynamic core is put forward in the paper.

3. Dynamic Core

Attributes reduct is the basic problem in rough set theory, and the computation of feature core is especially important for resolving this problem. All attributes in the feature core will be present in any reduct, otherwise discernible relation in decision system can not be ensured. Many reduct algorithms are based on the feature core. According to the feature core one can construct reduct heuristically, and the efficiency of reduct can be improved greatly.

Many references discuss about the feature core of reduct [2,3], but it is just static reduct. For dynamic reduct the feature core still need to be probed in a further step.

Given decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, we know that the feature core of decision system S in static reduct is

$$CORE(S) = \bigcap_{R \in RED(S)} R,$$

the feature core of sub-decision system B is

$$CORE(B) = \bigcap_{R \in RED(B)} R.$$

Definition 4: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, the dynamic core of S based on family F is defined by

$$DCORE(S, F) = CORE(S) \cap \bigcap_{B \in F} CORE(B),$$

$DCORE(S, F)$ is called F -dynamic core of decision system S .

Theorem 1: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, then the intersection of F -dynamic reduct contains F -dynamic core, that is

$$\bigcap_{R \in DR(S, F)} R \supseteq DCORE(S, F).$$

Proof: According to the definition of dynamic reduct [5], it follows

$$DR(S, F) = RED(S) \cap \bigcap_{B \in F} RED(B).$$

For any $B \in F$, $DR(S, F) \subseteq RED(B)$,

$$DR(S, F) \subseteq RED(S),$$

then,

$$\bigcap_{R \in DR(S, F)} R \supseteq \bigcap_{R \in RED(S)} R,$$

$$\bigcap_{R \in DR(S, F)} R \supseteq \bigcap_{R \in RED(B)} R$$

$$\bigcap_{R \in DR(S, F)} R \supseteq CORE(S),$$

$$\bigcap_{R \in DR(S, F)} R \supseteq CORE(B),$$

$$\bigcap_{R \in DR(S, F)} R \supseteq CORE(S) \cap \bigcap_{B \in F} CORE(B),$$

therefore,

$$\bigcap_{R \in DR(S, F)} R \supseteq CORE(S, F)$$

Theorem 1 means that each attribute in F -dynamic core is included by all F -dynamic reducts. Dynamic reducts is the most stable reducts of decision system S , then dynamic core is the most stable core of decision system S , which represents a stable unreduced attribute set.

4. (F - λ)-Dynamic Core

F -dynamic core can be sometimes too much restrictive so here applies a generalization of F -dynamic core. It will be more suitable for noise data.

Definition 5: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, $\lambda \in (0.5, 1]$, the (F - λ)-dynamic core of decision system S based on family F is defined by

$$DCORE_\lambda(S, F) = \left\{ \alpha \in CORE(S) \mid \frac{|\{B \in F : \alpha \in CORE(B)\}|}{|F|} \geq \lambda \right\}.$$

λ is precision coefficient, and the value of λ decides which attribute belongs to (F - λ)-dynamic core $DCORE_\lambda(S, F)$.

λ approaches 1, $DCORE_\lambda(S, F)$ will be closed to $DCORE(S, F)$, while λ approaches 0.5, $DCORE_\lambda(S, F)$ is more rough compared with $CORE(S, F)$, but $DCORE_\lambda(S, F)$ will comprise more attributes.

Theorem 2: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, $F' \subseteq \rho(S)$, we have the following propositions.

1) If $F = \{S\}$, then $DCORE(S, F) = CORE(S)$

In this case dynamic core is just the feature core of decision system S .

$$2) DCORE_1(S, F) = DCORE(S, F)$$

Actually, when λ increases from 0.5 to 1, the dynamic core will change from $DCORE_\lambda(S, F)$ to $DCORE(S, F)$.

3) If $\lambda \leq \lambda_1$, then

$$DCORE_{\lambda_1}(S, F) \subseteq DCORE_\lambda(S, F)$$

Obviously, for any $\lambda \in (0.5, 1]$, there will be

$$DCORE(S, F) \subseteq DCORE_\lambda(S, F).$$

Theorem 3: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, $\lambda \in (0.5, 1]$, then the intersection of $(F-\lambda)$ -dynamic reduct contains $(F-\lambda)$ -dynamic core.

Proof: According to the definition of $(F-\lambda)$ -dynamic core, it follows

$$CORE(S) \subseteq DCORE_\lambda(S, F)$$

$(F-\lambda)$ -dynamic reduct is

$$DR_\lambda(S, F) = \left\{ Q \in RED(S) \mid \frac{|\{B \in F : Q \in RED(B)\}|}{|F|} \geq \lambda \right\}$$

It is obviously that

$$DR_\lambda(S, F) \subseteq RED(S), = S$$

then,

$$\bigcap_{R \in DR_\lambda(S, F)} R \supseteq \bigcap_{R \in RED(S)} R = CORE(S),$$

therefore,

$$\bigcap_{R \in DR_\lambda(S, F)} R \supseteq DCORE_\lambda(S, F).$$

5. Generalized Dynamic Core

According to the definition of dynamic core, if some feature attributes of any sub-decision system in F family are comprised by dynamic core, then it is certainly a feature attribute of decision system S . This notion sometimes can not be convenient because we are interested in useful sets of attributes which are not necessarily reducts of the decision system. Therefore we have to generalize the notion of a dynamic core.

Definition 6: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, then

$$GDCORE(S, F) = \bigcap_{B \in F} CORE(B).$$

$GDCORE(S, F)$ is called the generalized dynamic core of decision system S .

Definition 7: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, $\lambda \in (0.5, 1]$, then

$$GDCORE_\lambda(S, F) = \left\{ \alpha \in C \mid \frac{|\{B \in F : \alpha \in CORE(B)\}|}{|F|} \geq \lambda \right\}.$$

The $GDCORE_\lambda(S, F)$ is called the $(F-\lambda)$ -generalized dynamic core of decision system S .

Theorem 4: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, $\lambda \in (0.5, 1]$, we have the following propositions.

$$1) DCORE(S, F) \subseteq GDCORE(S, F)$$

By definitions 6 and 7 we know that it is obviously true.

$$2) DCORE_\lambda(S, F) \subseteq GDCORE_\lambda(S, F)$$

By 1) it is also obviously true.

3) If $S \in F$, then

$$GDCORE(S, F) = DCORE(S, F).$$

when F family contains decision system S , then generalized dynamic core will be just dynamic core.

Theorem 5: Decision system $S = (U, C \cup D)$, where U is a non-empty, finite set called the universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, $\lambda \in (0.5, 1]$, then

1) The intersection of F -generalized dynamic reduct contains F -generalized dynamic core

Proof: According to the definition of F -dynamic reduct [5], it follows

$$GDR(S, F) = \bigcap_{B \in F} RED(B),$$

for any $B \in F$, $GDR(S, F) \subseteq RED(B)$,

$$\bigcap_{R \in GDR(S, F)} R \supseteq \bigcap_{R \in RED(B)} R,$$

$$\bigcap_{R \in GDR(S, F)} R \supseteq CORE(B) \supseteq \bigcap_{B \in F} CORE(B),$$

then,

$$\bigcap_{R \in GDR(S, F)} R \supseteq GDCORE(S, F).$$

2) The intersection of $(F-\lambda)$ -generalized dynamic reduct contains $(F-\lambda)$ -generalized dynamic core

Proof: For any attribute $\alpha \in GDCORE_\lambda(S, F)$, it satisfies $\frac{|\{B \in F : \alpha \in CORE(B)\}|}{|F|} \geq \lambda$.

It might as well suppose that attribute α exists in core of k sub-decision systems,

$$k = \left| \left\{ B \in F : \alpha \in \text{CORE}(B) \right\} \right|,$$

We have

$$\begin{aligned} &\alpha \in \text{CORE}(B_1), \dots, \alpha \in \text{CORE}(B_k), \\ &\alpha \in \bigcap_{R \in \text{RED}(B_1)} R, \dots, \alpha \in \bigcap_{R \in \text{RED}(B_k)} R. \end{aligned} \quad (1)$$

According to the definition of $(F-\lambda)$ -generalized dynamic reduct [5],

$$\text{GDR}_\lambda(S, F) = \left\{ R \subseteq S \mid \frac{\left| \{ B \in F : R \in \text{RED}(B) \} \right|}{|F|} \geq \lambda \right\},$$

because of $\lambda > 0.5$, any reduct R in $(F-\lambda)$ -generalized dynamic reduct satisfies:

$$R \in \text{RED}(B_1) \cup \text{RED}(B_2) \cup \dots \cup \text{RED}(B_k).$$

By (1), for any attribute α , we can infer:

$$\alpha \in \bigcap_{R \in \text{GDR}_\lambda(S, F)} R,$$

then,

$$\bigcap_{R \in \text{GDR}_\lambda(S, F)} R \supseteq \text{GDCORE}_\lambda(S, F).$$

Actually, supposes $\alpha \notin \bigcap_{R \in \text{GDR}_\lambda(S, F)} R,$

therefore it exists $R' \in \text{GDR}_\lambda(S, F)$, and $\alpha \notin R'$.

By (1), it follows

$$R' \notin \text{RED}(B_1), \dots, R' \notin \text{RED}(B_k). \quad (2)$$

If $\lambda > 0.5$, then (2) is contradiction with hypothesis, then

$$\alpha \in \bigcap_{R \in \text{GDR}_\lambda(S, F)} R,$$

only when $\lambda \leq 0.5$, it may be true for the hypothesis, but we have fixed $\lambda > 0.5$, so it is impossible.

6. Quality of Family F

We can not calculate the family of all subtables of universal decision information table, therefore we construct family F to estimate universal decision information table. General speaking, if family F contains more elements then the stability of the calculated reduct will be better.

Extracting too much subtables inevitably cause enormous computation. From the thought of *Bernoulli's theorem* we know that if the extracted subtables is adequate, continuing to extract other subtables has little relate with reduct stability coefficient. So a minimum value of $|F|$ can be determined.

However the computation of sample of dynamic reduct should emphasize the other important aspect, namely the influence of sampling distribution to the validity of the dynamic reduct. We generally assume that the sample is

random. The basic thought of evaluate the quality of family F is proposed. It can be used to weight the equalization of family F and reflect the validity of the dynamic reduct.

Definition 7 (entire sample): Decision system $S = (U, C \cup D)$, where U is universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, and

$$UF = \bigcup \{ U_B \mid B = (U_B \subset U, C \cup D) \in F \}.$$

UF is named entire sample set.

Definition 8 (central sample): Decision system $S = (U, C \cup D)$, where U is universe, C is condition attributes set, D is decision attributes set, $F \subseteq \rho(S)$, and

$$UCF = \bigcap \{ U_B \mid B = (U_B \subset U, C \cup D) \in F \}.$$

UCF is named central sample set.

Taking simple random sampling as a precondition, firstly we will consider the equalization of family F and give two radios to weight it.

1) Entire Sampling Rate

The entire sampling rate is denoted by $|U_F|/|U|$. It is easy to observe that $0 \leq |U_F|/|U| \leq 1$. If $|U_F|/|U|$ takes a too small value then the sample is too centralized. So it can not reflect the character of the original decision information system, and reduce the confidence degree of dynamic reduct.

We think the sample of family F is deficient when the value of $|U_F|/|U|$ less than 0.5. A threshold value can be established. The sample conforms to the requirement and the result of dynamic reduct is effective if $|U_F|/|U|$ is bigger than the threshold value.

2) Entire Balanced Degree

The entire balanced degree is denoted by $|U_{CF}|/|U_F|$, and It is easy to observe that $0 \leq |U_{CF}|/|U_F| \leq 1$. If $|U_{CF}|/|U_F|$ takes a too big value then the sample may too repetitious. So it can not reflect the character of object of the original decision information system, and reduce the confidence degree of dynamic reduct.

We think the sample of family F is unbalanced when the value of $|U_{CF}|/|U_F|$ is bigger than 0.5. A threshold value can be established. The uniformity of sample distribution is better and the result of dynamic reduct is effective if $|U_{CF}|/|U_F|$ is lower than the threshold value.

Just evaluating the performance of entire sample is not enough, we should also avoid the abnormal phenomenon of individual decision table in family F . The influence of individual subtables to the distribution must be not too much. For example, we extract 100 subtables. If the radio of central sample set constructed by 99 subtables of them is larger than 0.5, when a subtable is brought into the central sample set, the radio decline sharply. For another

example, if the ratio of entire distribution is acceptable, but the difference among half of the subtables is small. It indicates the distribution of sampling is not balanced.

We can calculate the degree that central sample set contains each of the subtables and the degree that entire sample set contains each of the subtables. The two parameters can be used to evaluate the equalization of subtables.

3) Partial Sampling Rate

The partial sampling rate is denoted by $|U_B|/|U_F|$. It is easy to observe that $0 \leq |U_B|/|U_F| \leq 1$. If $|U_B|/|U_F|$ of some subtables take a higher value then the sample may be unbalanced. Even the entire balanced degree is permissible, but most of the subtables overlap, this reduces the confidence degree of dynamic reduct.

4) Partial Balanced Degree

The partial balanced degree is denoted by $|U_{CF}|/|U_B|$. It is easy to observe that $0 \leq |U_{CF}|/|U_B| \leq 1$. If $|U_{CF}|/|U_B|$ of some subtables takes a higher value then the sample is too centralized. Even the entire balanced degree is permissible, but most of the subtables overlap.

7. Conclusion

The notion of dynamic core is actually based on dynamic reducts in the paper. Dynamic core consists with the most stable attributes which can not be reducted, and describes the set of feature attributes. At the same time it is proved that the intersection of dynamic reducts contains dynamic core, which is suitable for all kinds of de-

finitions about dynamic reducts. All of these show that the dynamic core possesses the essential properties of the feature core in deed. We can say that dynamic core expresses the feature attributes in a more general way.

REFERENCES

- [1] Z. Pawlak, "Rough Sets (Periodical Style)," *International Journal of Computer and Information Sciences*, Vol. 11, No. 5, 1982, pp. 341-356. [doi:10.1007/BF01001956](https://doi.org/10.1007/BF01001956)
- [2] X. H. Hu and N. Cercone, "Learning in Relational Databases: A Rough Set Approach (Periodical Style)," *International Journal of Computational Intelligence*, Vol. 11, No. 2, 1995, pp. 323-338.
- [3] J. Jelonek, K. Krawiec and R. Slowinski, "Rough Set Reduction of Attributes and Their Domains for Neural Networks (Periodical Style)," *Computational Intelligence*, Vol. 11, No. 2, 1995, pp. 339-347. [doi:10.1111/j.1467-8640.1995.tb00036.x](https://doi.org/10.1111/j.1467-8640.1995.tb00036.x)
- [4] J. G. Bazan, "A Comparison of Dynamic and Non-Dynamic Rough Set Methods for Extracting Laws from Decision Tables (Published Conference Proceedings Style)," In: L. Polkowski and A. Skowron, Eds., *Rough Sets in Knowledge Discovery 1: Methodology and Applications*, Physica-Verlag, Heidelberg, 1998, pp. 321-365.
- [5] J. G. Bazan, A. Skowron and P. Synak, "Dynamic Reducts as a Tool for Extracting Laws from Decision Tables, in: Methodologies for Intelligent System (Published Conference Proceedings Style)," *Proceedings of 8th International Symposium ISMIS'94*, Vol. 869, Springer-Verlag, Berlin, 1994, pp. 346-355.