

Attribute Reduction in Utility-Based Decision-Theoretic Rough Set Models

Nan Zhang^{1,2(✉)}, Lili Jiang^{1,2}, and Caihui Liu³

¹ School of Computer and Control Engineering,
Yantai University, Yantai 264005, Shandong, China

zhangnan0851@163.com, ytjianglili0919@163.com

² Key Lab for Data Science and Intelligent Technology of Shandong Higher
Education Institutes, Yantai University, Yantai 264005, Shandong, China

³ Department of Mathematics and Computer Science,
Gannan Normal University, Ganzhou 341000, Jiangxi, China
liu_caihui@163.com

Abstract. Decision-theoretic rough set (DTRS) model, proposed by Yao in the early 1990's, introduces Bayesian decision procedure and loss function in rough set theory. Considering utility function in decision processing, utility-based decision-theoretic rough set model (UDTRS) is given in this paper. The utility of the positive region, the boundary region and the negative region are obtained respectively. We provide a reduction definition which can obtain the maximal utility in decisions. A heuristic reduction algorithm with respect to the definition is proposed. Finally, experimental results show the proposed algorithm is effective.

Keywords: Attribute reduction · Utility theory · Decision-theoretic rough sets

1 Introduction

Decision-theoretic rough set (DTRS) model was firstly introduced by Yao et al. [1] in the early 1990's. As a probabilistic rough set model, it has been successfully used in many research areas, such as knowledge presentation [2–4], data mining [5], machine learning [6], artificial intelligence [7, 8] and pattern recognition.

Attribute reduction [9–14] aims to remove the unnecessary attributes from the information system while keeping the particular property, and becomes one of the hottest issues in rough set theory. Yao and Zhao [9] studied attribute reduction in decision-theoretic rough set models with respect to the different classification properties, confidence, coverage, decision-monotonicity, generality and cost, they also gave a general definition of probabilistic attribute reduction. Jia et al. [10] provided a minimum cost attribute reduction in decision-theoretic rough set model, and decision cost induced from the reduct is minimum. Dou et al. proposed a parameterized decision-theoretic rough set model in the paper [11]. In the proposed model, the smallest possible cost and the largest possible cost are computed respectively. Li et al. [12] introduced a non-monotonic attribute reduction for decision-theoretic rough set model. The expanded positive region can be kept by the non-monotonic attribute

reduction in an information system. To extend classical indiscernibility relation in Yao's decision-theoretic rough sets, Ju et al. [13] gave the δ -cut decision-theoretic rough set. In the proposed decision-theoretic rough set model, attribute reduction of the decision-monotonicity criterion and the cost minimum criterion are proposed respectively in the paper. By constructing variants of conditional entropy in decision-theoretic rough set model, Ma et al. [14] proposed solutions to the attribute reduction based on decision region preservation.

The remaining parts of this paper are arranged as follows. Some basic notions with respect to utility-based decision-theoretic rough set (UDTRS) model are briefly recalled in Sect. 2. Definition of attribute reduction in UDTRS and relative heuristic reduction algorithm are investigated respectively in Sect. 3. We give the experimental analysis in Sect. 4. The paper is summarized in Sect. 5.

2 Utility-Based Decision-Theoretic Rough Sets

By considering the subjective factors in risk decision, Zhang et al. [15] proposed utility-based decision-theoretic rough set (UDTRS) model based on Yao's decision-theoretic rough set model [1]. In this section, we briefly recall some basic notions about utility-based decision-theoretic rough set model. Detailed information about UTRS can be found in the paper [15].

A decision system is defined as the 3-tuple $S = (U, AT = C \cup D, V_a)$. Universe U is the finite set of the objects; AT is a nonempty set of the attributes, such that for all $a \in AT$; C is the set of conditional attribute and D is the set of decision attribute; V_a is the domain of attribute a .

For each nonempty subset $A \subseteq AT$, the indiscernibility relation $IND(A)$ is defined as: $IND(A) = \{(x, y) \in U^2, a(x) = a(y), \forall a \in A\}$. Two objects in U satisfy $IND(A)$ if and only if they have the same value in $\forall a \in A$. U is divided into a family of disjoint subsets $U/IND(A)$ defined a quotient set of U as $U/IND(A) = \{[x]_A : x \in U\}$, where $[x]_A = \{y \in U : (x, y) \in IND(A)\}$ denotes the equivalence class determined by x with respect to A . The set of states is given by $\Omega = U/D = \{X, X^c\}$ indicating that an object is in state X or X^c .

Utility is an important economic concept, and it reflects degree of one's satisfaction related to the cost or profit in decision procedure. For $\forall x \in U$ and $[x] \in U/\pi$, $u(\lambda)$ is utility function, λ denotes the cost of taking action. The expected utilities for different actions can be expressed as:

$$\begin{aligned}\Psi(a_P|[x]) &= u(\lambda_{PP})P(X|[x]) + u(\lambda_{PN})P(X^c|[x]) \\ \Psi(a_N|[x]) &= u(\lambda_{NP})P(X|[x]) + u(\lambda_{NN})P(X^c|[x]) \\ \Psi(a_B|[x]) &= u(\lambda_{BP})P(X|[x]) + u(\lambda_{BN})P(X^c|[x])\end{aligned}$$

According to maximal utility in Bayesian procedures, we have the following as:

- If $\Psi(a_P|[x]) \geq \Psi(a_N|[x])$ and $\Psi(a_P|[x]) \geq \Psi(a_B|[x])$, then $[x] \subseteq POS_\pi(X)$,
- If $\Psi(a_N|[x]) \geq \Psi(a_P|[x])$ and $\Psi(a_N|[x]) \geq \Psi(a_B|[x])$, then $[x] \subseteq NEG_\pi(X)$,
- If $\Psi(a_B|[x]) \geq \Psi(a_P|[x])$ and $\Psi(a_B|[x]) \geq \Psi(a_N|[x])$, then $[x] \subseteq BND_\pi(X)$.

If $P(X|[x]) = P$ then $P(X^c|[x]) = 1 - P$, then we derived the following decision rules:

- If $P(X|[x]) \geq \alpha_u$, then $[x] \subseteq POS_\pi(X)$,
- If $P(X|[x]) \leq \beta_u$, then $[x] \subseteq NEG_\pi(X)$,
- If $\beta_u < P(X|[x]) < \alpha_u$, then $[x] \subseteq BND_\pi(X)$;

where

$$\alpha_u = \frac{u(\lambda_{BN}) - u(\lambda_{PN})}{(u(\lambda_{BN}) - u(\lambda_{PN})) + (u(\lambda_{PP}) - u(\lambda_{BP}))},$$

$$\beta_u = \frac{u(\lambda_{NN}) - u(\lambda_{BN})}{(u(\lambda_{NN}) - u(\lambda_{BN})) + (u(\lambda_{BP}) - u(\lambda_{NP}))}.$$

Since $u(\lambda_{PP}) \geq u(\lambda_{BP}) > u(\lambda_{NP})$, $u(\lambda_{NN}) \geq u(\lambda_{BN}) > u(\lambda_{PN})$, $\alpha_u \in (0, 1]$, $\beta_u \in [0, 1)$ Then, we can obtain

$$\alpha_u = \frac{1}{1 + \Delta(\alpha_u)} = \frac{1}{1 + \frac{u(\lambda_{PP}) - u(\lambda_{BP})}{u(\lambda_{BN}) - u(\lambda_{PN})}},$$

$$\beta_u = \frac{1}{1 + \Delta(\beta_u)} = \frac{1}{1 + \frac{u(\lambda_{BP}) - u(\lambda_{NP})}{u(\lambda_{NN}) - u(\lambda_{BN})}}.$$

For $\forall X \subseteq U$, (α_u, β_u) -upper and lower approximations in utility-based decision-theoretic rough set model are presented as:

$$\underline{apr}_{(\alpha_u, \beta_u)}(X) = \{x \in U | P(X|[x]) \geq \alpha_u\},$$

$$\overline{apr}_{(\alpha_u, \beta_u)}(X) = \{x \in U | P(X|[x]) > \beta_u\}.$$

Based on the definition of rough approximations in UDTRS, the positive, boundary and negative regions are defined as

$$POS_\pi(X) = \underline{apr}_{(\alpha_u, \beta_u)}(X),$$

$$BND_\pi(X) = \overline{apr}_{(\alpha_u, \beta_u)}(X) - \underline{apr}_{(\alpha_u, \beta_u)}(X),$$

$$NEG_\pi(X) = U - \overline{apr}_{(\alpha_u, \beta_u)}(X).$$

3 Attribute Reduction in UDTRS

In this section, we will give the definition of attribute reduction based on maximal utility in UDTRS. By attribute reduction, the maximal utility will be obtain in decisions. According to the proposed definition of reduction, a heuristic algorithm with respect to the maximal utility will be investigated in this section.

Similar to the Bayesian expected cost [10] in decision-theoretic rough set model, the Bayesian expected utility [15] of each rule is expressed as:

$$\begin{aligned} \text{Utility of the positive rule} &: p \cdot u(\lambda_{PP}) + (1 - p) \cdot u(\lambda_{PN}); \\ \text{Utility of the negative rule} &: p \cdot u(\lambda_{BP}) + (1 - p) \cdot u(\lambda_{BN}); \\ \text{Utility of the boundary rule} &: p \cdot u(\lambda_{NP}) + (1 - p) \cdot u(\lambda_{NN}). \end{aligned}$$

From above, we can easily get the Bayesian expected utility of decision rules:

Utility of positive rules:

$$Utility_A^{POS} = \sum_{x_i \in POS_{(z_u, \beta_u)}(\pi_D/\pi_A)} (p_i \cdot u(\lambda_{PP}) + (1 - p_i) \cdot u(\lambda_{PN}));$$

Utility of boundary rules:

$$Utility_A^{BND} = \sum_{x_j \in BND_{(z_u, \beta_u)}(\pi_D/\pi_A)} (p_j \cdot u(\lambda_{BP}) + (1 - p_j) \cdot u(\lambda_{BN}));$$

Utility of negative rule:

$$Utility_A^{NEG} = \sum_{x_k \in NEG_{(z_u, \beta_u)}(\pi_D/\pi_A)} (p_k \cdot u(\lambda_{NP}) + (1 - p_k) \cdot u(\lambda_{NN})).$$

For any subset $A \subseteq AT$, the whole utility is composed of three parts: utility of positive region, utility of boundary region and utility of negative region. Then, we have the whole utility of all decision rules in decision systems as follows [15]:

$$\begin{aligned} Utility_A &= Utility_A^{POS} + Utility_A^{BND} + Utility_A^{NEG} \\ &= \sum_{x_i \in POS_{(z_u, \beta_u)}(\pi_D/\pi_A)} (p_i \cdot u(\lambda_{PP}) + (1 - p_i) \cdot u(\lambda_{PN})) \\ &\quad + \sum_{x_j \in BND_{(z_u, \beta_u)}(\pi_D/\pi_A)} (p_j \cdot u(\lambda_{BP}) + (1 - p_j) \cdot u(\lambda_{BN})) \\ &\quad + \sum_{x_k \in NEG_{(z_u, \beta_u)}(\pi_D/\pi_A)} (p_k \cdot u(\lambda_{NP}) + (1 - p_k) \cdot u(\lambda_{NN})) \end{aligned}$$

In real applications, it is better to obtain more utility in decision procedures. Thus, according to “non-decreasing” principle, we define attribute reduction in utility-based decision-theoretic rough set model as follows:

Definition 1. A decision system $S = (U, C \cup D, V_a)$, $R \subseteq C$ is a reduct of C with respect to D if it satisfies the following two conditions:

- (1) $Utility_R \geq Utility_C$;
- (2) $\forall Rh' \subset R, Utility_{R'} < Utility_R$.

From Definition 1, the decision utility will be increased or unchanged by the reduction. Condition (1) is the jointly sufficient condition and condition (2) is the individual necessary condition. Condition (1) guarantees that the utility induced from the reduct is maximal, and condition (2) guarantees the reduct is minimal.

The fitness function, which shows the significance of an attribute, is usually used to construct a heuristic algorithm in rough set theory. In UTRS model, the fitness function is defined as:

Definition 2. A decision system $S = \{U, C \cup D, V_a\}$, $A \subseteq C$ The utility fitness function of attribute $a_i \in A$ is defined as:

$$Sig_{Utility}(A, a_i) = \frac{Utility_{A-\{a_i\}} - Utility_A}{Utility_A}.$$

The three strategies in heuristic algorithm is summarized in paper [9]. In this paper, we take deletion strategy to give an algorithm in UDTRS. The heuristic algorithm (The algorithm of maximal-utility attribute reduction, **AMUAR**) based on the utility fitness function is described as follows:

Algorithm: The Algorithm of Maximal-Utility Attribute Reduction, **AMUAR**

Input: A decision system $S = \{U, C \cup D, V\}$

Output: A reduct R

1 $B = C, A = C$;

2 While $A \neq \emptyset$

 Compute the fitness for each attribute $a_i \in A$ using fitness function $Sig_{Utility}(A, a_i)$; Find the minimal

$Sig_{Utility}(B, a_i) \geq 0$ and the corresponding attribute a_i ;

$A = A - \{a_i\}$; If $Utility_{B-\{a_i\}} \geq Utility_C$, $B = B - \{a_i\}$;

3 Output $B = R$.

The fitness function shows the significance of an attribute. In the processing of deleting attributes, if $Utility_B \geq Utility_C$, the algorithm will stop the deleting procedure and output reduct of decision systems.

4 Experimental Analysis

In this section, we will verify effectiveness of the algorithm AMUAR and the monotonicity of utility with attributions by experiments. All the experiments have been carried out on a personal computer with Windows 7, Intel (R) Pentium (R) CPU G640 (2.8 GHz) and 6.00 GB memory. The programming language is Matlab 2010b.

We take $u(\lambda) = a(-\lambda + c)^b$ as the utility function. If $0 < b < 1$, then UDTRS model is risk aversion; If $b = 1$, UDTRS model is risk neutrality; If $b > 1$, UDTRS model is risk loving; Six data sets from the UCI Machine Learning Repository are used. For each data set, the utility functions are randomly generated in interval value [100, 1000]. Their values meet the following constraint conditions: $u(\lambda_{BP}) > u(\lambda_{NP})$, $u(\lambda_{BN}) > u(\lambda_{PN})$, $u(\lambda_{PP}) = u(\lambda_{NN}) = 1$. 10 different groups of utility functions are randomly generated. Table 1 shows the average length of the derived reduct with different data sets.

Table 1. Average length of a reduct

| Data sets | Attributes | Samples | Risk aversion | Risk neutrality | Risk loving |
|---------------|------------|---------|---------------|-----------------|-------------|
| credit_a | 15 | 690 | 7.0 ± 0.0 | 7.0 ± 0.0 | 7.0 ± 0.0 |
| forestfires | 12 | 517 | 4.0 ± 0.0 | 4.0 ± 0.0 | 4.0 ± 0.0 |
| german | 20 | 1000 | 12.0 ± 0.0 | 12.0 ± 0.0 | 12.0 ± 0.0 |
| heart_statlog | 13 | 270 | 6.0 ± 0.0 | 6.0 ± 0.0 | 6.0 ± 0.0 |
| lymph | 18 | 148 | 8.0 ± 0.0 | 8.0 ± 0.0 | 8.0 ± 0.0 |
| vote | 16 | 435 | 11.9 ± 0.3 | 11.9 ± 0.3 | 11.8 ± 0.4 |
| breast_cancer | 9 | 286 | 7.9 ± 0.3 | 8.0 ± 0.0 | 7.9 ± 0.3 |
| fertility | 9 | 100 | 6.9 ± 0.3 | 6.9 ± 0.3 | 6.8 ± 0.4 |

To validate the monotonicity of utility with attributes, utility is calculated with the increasing number of attributes from 1 to the total attribute number in each data set. In Fig. 1, the x-coordinate represents the number of attributes, and the y-coordinate represents the utility of three models. Figure 1 shows the utility of three models do not strictly increase with the increasing of attribute numbers. For example, the utility decrease with adding an attribute in data set credit_a, forestfires and vote. The utility with the number of attributes increasing do not present monotonicity strictly.

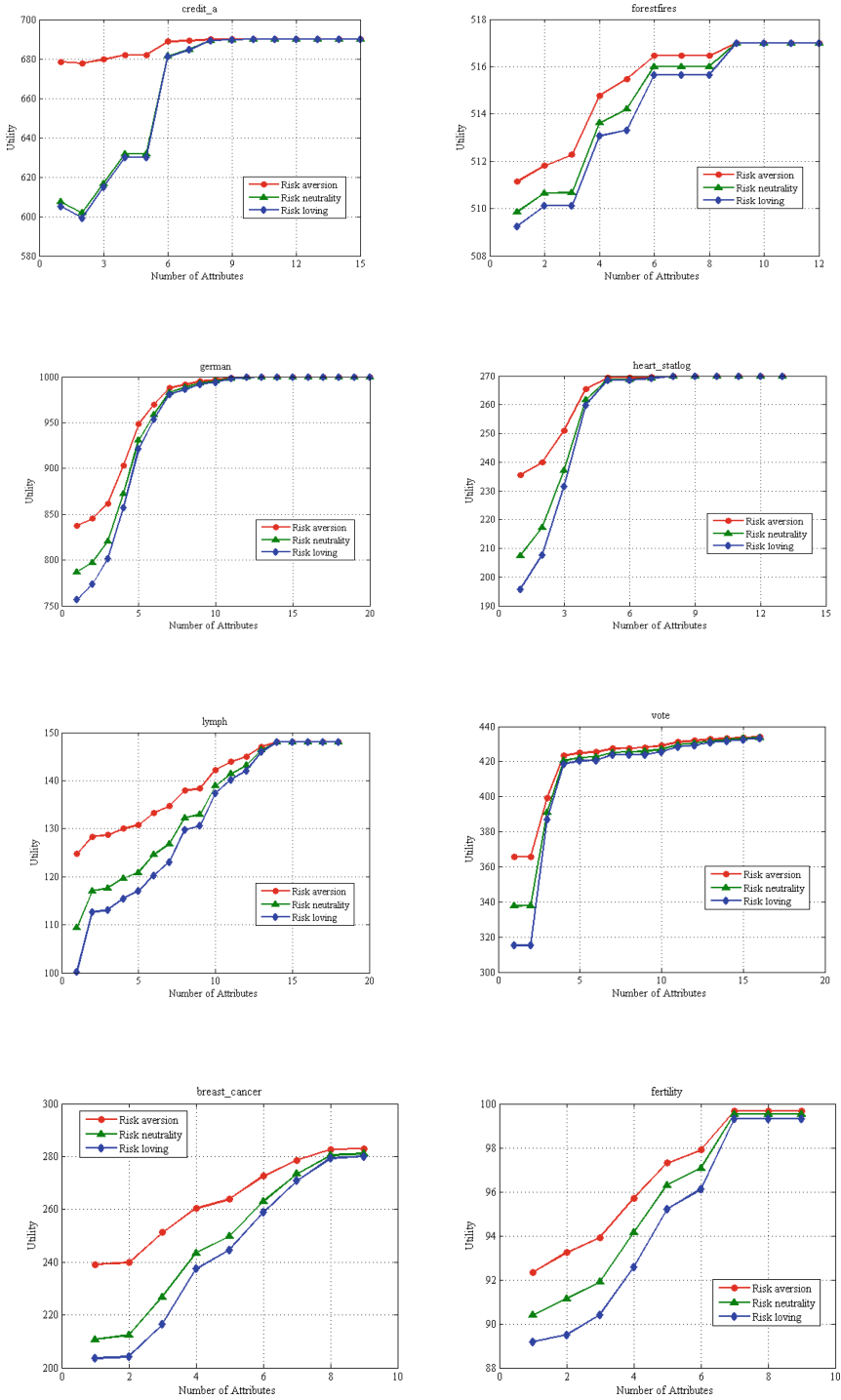


Fig. 1. Utility comparison of three attribute reductions

5 Conclusions

Utility-based decision-theoretic rough set model is introduced in this paper. The utility of the positive region, the boundary region and the negative region are given respectively. We provide a definition of reduction which aims to obtain the maximal utility in decisions. A heuristic reduction algorithm with respect to the definition is proposed. Finally, experimental results show the proposed algorithm is effective.

Acknowledgements. This work was partially supported by the National Natural Science Foundation of China (Nos. 61403329, 61572418, 61663002, 61502410, 61572419), the Natural Science Foundation of Shandong Province (No. ZR2013 FQ020).

References

1. Yao, Y.Y., Wong, S.K.M., Lingras, P.: A decision-theoretic rough set model. In: Ras, Z.W., Zemankova, M., and Emrichm M.L. (eds.) Proceedings of the 5th International Symposium on Methodologies for Intelligent Systems, 25–27 October 1990. North-Holland, New York (1990)
2. Liu, D., Liang, D.C., Wang, C.C.: A novel three-way decision model based on incomplete information system. *Knowl.-Based Syst.* **91**, 32–45 (2016)
3. Deng, X.F., Yao, Y.Y.: Decision-theoretic three-way approximations of fuzzy sets. *Inf. Sci.* **279**, 702–715 (2014)
4. Herbert, J.P., Yao, J.T.: Game-theoretic rough sets. *Fundamenta Informaticae* **108**, 267–286 (2011)
5. Yu, H., Liu, Z.G., Wang, G.Y.: An automatic method to determine the number of clusters using decision-theoretic rough set. *Int. J. Approx. Reason.* **55**, 101–115 (2014)
6. Zhang, H.R., Min, F.: Three-way recommender systems based on random forests. *Knowl.-Based Syst.* **91**, 275–286 (2016)
7. Zhou, B., Yao, Y.Y., Luo, J.G.: Cost-sensitive three-way email spam filtering. *J. Intell. Inf. Syst.* **42**, 19–45 (2014)
8. Liang, D.C., Pedrycz, W., Liu, D., Hu, P.: Three-way decisions based on decision-theoretic rough sets under linguistic assessment with the aid of group decision making. *Appl. Soft Comput.* **29**, 256–269 (2015)
9. Yao, Y.Y., Zhao, Y.: Attribute reduction in decision-theoretic rough set model. *Inf. Sci.* **178**, 3356–3373 (2008)
10. Jia, X.Y., Liao, W.H., Tang, Z.H., Shang, L.: Minimum cost attribute reduction in decision-theoretic rough set models. *Inf. Sci.* **219**, 151–167 (2013)
11. Dou, H.L., Yang, X.B., Song, X.N., et al.: Decision-theoretic rough set: a multicost strategy. *Knowl.-Based Syst.* **91**, 71–83 (2016)
12. Li, H.X., Zhou, X.Z., Zhao, J.B.: Non-monotonic attribute reduction in decision-theoretic rough sets. *Fundamenta Informaticae* **126**, 415–432 (2013)
13. Ju, H.R., Dou, H.L., Qi, Y., Yu, H.L., Yu, D.J., Yang, J.Y.: δ -cut decision-theoretic rough set approach: model and attribute reductions. *Sci. World J.* **2014**, 1–12 (2014)
14. Ma, X.A., Wang, G.Y., Yu, H., Li, T.R.: Decision region distribution preservation reduction in decision-theoretic rough set model. *Inf. Sci.* **278**, 614–640 (2014)
15. Zhang, N., Jiang, L.L., Yue, X.D., Zhou, J.: Utility-based three-way decisions model. *CAAI Trans. Intell. Syst.* **11**, 459–468 (2016)