Jian-hua Dai, Hu Hu, Guo-jie Zheng, Qing-hua Hu, Hui-feng Han, Hong Shi, 2016. Attribute reduction in interval-valued information systems based on information entropies. *Frontiers of Information Technology & Electronic Engineering*, **17**(9):919-928.
http://dx.doi.org/10.1631/FITEE.1500447

Attribute reduction in interval-valued information systems based on information entropies

Key words: Rough set theory, Interval-valued data, Attribute reduction, Entropy

Corresponding author: Jian-hua Dai
E-mail: david.joshua@qq.com

ORCID: http://orcid.org/0000-0003-1459-0833

Motivation

- Interval-valued data appear as a way to represent the uncertainty affecting the observed values. Dealing with interval-valued information systems is helpful to generalize the applications of rough set theory.
- Attribute reduction is a key issue in analysis of interval-valued data.
- Existing attribute reduction methods for single-valued data are unsuitable for interval-valued data. So far, there have been few studies on attribute reduction methods for interval-valued data.

Main idea

- Some information theory concepts, including entropy, conditional entropy, and joint entropy, are given in interval-valued information systems.
- The paper provides an information theory view for attribute reduction in interval-valued information systems.

Method

1. Introduce the concept of relative bound difference similarity degree between two interval values;
2. Introduce the concepts of entropy, conditional entropy, and joint entropy;
3. Construct algorithms for interval-valued information systems and interval-valued decision systems;
4. Carry out various experiments based on the proposed framework.

Major results

- Our method can measure the importance of conditional attributes.

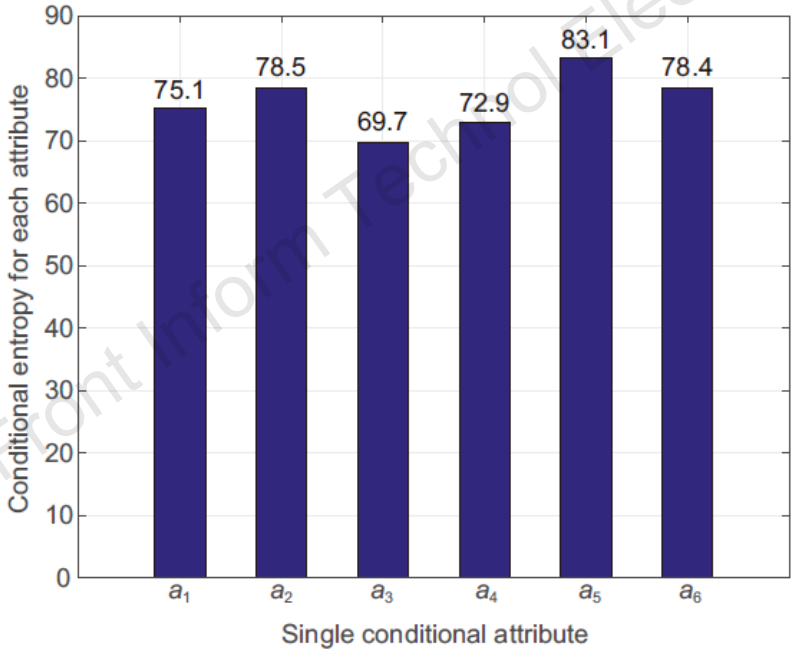

Fig. 2 Conditional entropy for the single attribute of the Face Recognition dataset

Major results (Cont'd)

- Compared with intersection-union similarity (IU), possible degree similarity (PD), and uncertainty measurement attribute reduction (UMAR), our method obtains satisfactory results.

Table 2 Performance on the Fish dataset by KNN and PNN

Threshold	RBD		IU		PD		UMAR	
α	KNN	PNN	KNN	PNN	KNN	PNN	KNN	PNN
Full attributes	0.5833	0.6667	0.5833	0.6667	0.5833	0.6667	0.5833	0.6667
0.10	**0.5833**	**0.6667**	0.5000	0.5000	0.5000	0.5000	**0.5833**	0.5833
0.15	**0.5833**	**0.6667**	0.5000	0.5000	0.5000	0.5000	**0.5833**	0.5833
0.20	0.5833	**0.6667**	**0.6667**	**0.6667**	0.5000	0.5000	0.5833	0.5833
0.25	0.5833	**0.6667**	**0.6667**	**0.6667**	**0.6667**	**0.6667**	0.5000	0.5000
0.30	0.5833	**0.6667**	0.5000	0.5833	**0.6667**	**0.6667**	0.5000	0.5000
0.35	0.5833	**0.6667**	0.5000	0.5833	**0.6667**	**0.6667**	0.5000	0.5000
0.40	0.3333	0.5833	0.5000	0.5000	**0.6667**	**0.6667**	0.5000	0.5000
0.45	0.5833	**0.7500**	0.5000	0.5000	**0.6667**	0.6667	0.5000	0.5000
0.50	**0.5833**	**0.8333**	0.3333	0.4167	0.5000	0.5833	0.5000	0.5000
0.55	**0.5833**	**0.5833**	0.3333	0.4167	0.5000	**0.5833**	0.5000	0.5000
0.60	**0.5000**	**0.5000**	0.3333	0.4167	**0.5000**	**0.5000**	0.3333	0.4167
0.65	**0.5833**	**0.5833**	0.1667	0.3333	0.3333	0.4167	0.3333	0.4167
0.70	**0.5000**	**0.5833**	0.1667	0.3333	0.3333	0.4167	0.3333	0.4167
0.75	**0.5000**	**0.5000**	0.3333	0.4167	0.3333	0.4167	0.3333	0.4167

Conclusions

- By introducing the concept of relative bound difference similarity degree between two interval values, this paper constructs the concepts of entropy, conditional entropy, and joint entropy.
- Algorithms for interval-valued information systems and interval-valued decision systems are proposed.
- Experiments show that the proposed framework is effective for attribute reduction in interval-valued information systems.