Heuristics-based Modelling of Human Decision Process

Sarang Balasaheb Bhasme, MS[1]; Raman Saurabh, MS[2]; Muhammad Salman Saeed, Alexey Nazarov, Professor,[1,2]

[1]Moscow Institute of Physics and Technology, [2]Federal Research Center Computer Science and Control of RAS

Abstract

Attitudinal Choquet integral (ACI) is a recent aggregation operator that considers in the aggregation process the criteria interaction and the DM's attitude, both of which are specific to the decision-maker. However, this capability comes at the cost of increased complexity that hinders its applicability in big data analytics. To address the same, in this paper, we explore some heuristics-based forms of the ACI operator, so as to somehow overcome its complexity. We devise new and efficient forms of ACI, and test their validity in the real world datasets, against the backdrop of preference learning.

Index Terms Attitudinal Choquet integral; efficiency; complexity reduction; attitudinal character; multi criteria decision making.

Introduction

In human decision making, the aggregation operation is typically required to obtain a representative value. For instance, different criteria evaluations for an alternative are aggregated to obtain a net score that is used to compare the alternatives. There are several aggregation operators that have been proposed in the literature, for example, weighted averaging and geometric mean are the conventional aggregation operators existing since long. Ordered weighted averaging (OWA) operator24, power averaging 24, and power geometric operators 23 are relatively recent operators that find applications in multi criteria decision making (MCDM). The discrete Choquet integral 6,18,21 is an aggregation function that considers varying degree of interaction among the criteria. The criteria can interact positively, indicating synergic or desirable combination; or negatively indicating redundancy or undesirable combination. For instance, while making a choice of house, the criteria size and location, are complementary, and hence interact positively, luxury and size may interact negatively for a buyer who is looking for a luxurious but small house. It is shown in 11,14 that most of the extant aggregation operators are the special cases of it. It has been applied in policy capturing in strategic decision making 16, analysis of root dispersal with competition among wood species in forests 19, computation of the number of citations22, clinical diagnosis 20, monitoring of the improvement of an overall industrial performance5, selection of groups of genes with high classifying power in gene expression data analysis 9, evaluation of discomfort in sitting position when driving a car 13, feature selection 12, and MCDM 10,11. The human aggregation process is typically characterized by criteria interaction. However, besides the criteria interaction, the individualistic attitude of a DM adds to the complexity. In 25, Zimmermann empirically demonstrated the effect of criteria interaction and the attitude in experiments involving human subjects. The real subjects were first asked to evaluate different tiles against multiple criteria, and then also asked the overall aggregated evaluation. It was observed that each subject displayed a different compensatory effect in his aggregation process. A tolerant DM stresses on meeting only some criteria. In contrast, a perfectionist DM would like all criteria to be met. This difference in attitudes also affect their aggregation processes. A tolerant DM has an OR-like or disjunctive aggregation, in which the aggregated value is towards the best of the values to be aggregated. Opposite to this, a perfectionist displays an AND-like or conjunctive aggregation, with the aggregated value towards the minimum of the arguments of aggregation. To portray such attitudinal tendencies, Zimmermann combined a T -norm and a T -conorm in a controllable proportion. The attitudinal tendencies of a DM are empirically validated in 8,15,17. Yager 24 developed ordered weighted averaging (OWA) operator with the ability to adjust the DM's attitude through the weight vector. Among the recent operators in this regard are compensative weighted mean1, generalized compensative weighted mean 2, and attitudinal Choquet integral (ACI) 3.

Methods and Materials

We use the real datasets in our experiments, which are available in UCI and Weka repositories. Since monotonicity is a basic property in all the aggregation operators, all the chosen datasets are monotone in the sense that "the more the better".

Experiment Steps

We implement weighted averaging (WA), Choquet integral (CI), ACI and finally our heuristic methods on a set of nine datasets shown in Table I by performing the following steps:

1)Given a set of alternatives A, two halves are created, namely A train and A test, for training and testing, respectively.

2)The training information is provided in terms of the pairwise preference pairs of the form an > bn are identified through the standard method of random sampling from Atrain The number of preferences are taken as N = 1000. For a healthy comparison, the same set of N random preferences is provided to different methods, in each iteration.

3)With different approaches, different learning models are induced on the given preference pairs.

4)The ranks of the given alternatives in Atest are predicted through WA, CI, and ACI methods.

5)The corresponding prediction accuracies are evaluated by comparing the predicted ranking for Atest through C-index.

6)For sanctity, the process is repeated 100 times for each dataset and for each approach.

7)The average of the 100 accuracy values, so obtained, is taken as the accuracy value. The respective standard deviation is also determined in the multiple accuracy values. The figures are shown in Table II.

Results

The scoring function U (a) helps to predict the rank of an alternative a ∈ Atest. The predicted rankings are compared with the ground truth ranking for the given alternatives in Atest. The performance of a method is determined by computing the degree of match between the predicted and ground truth ranks.

Based on the ordinal ground-truth rankings, an ordered partition of Atest can be formed shown as Atest = {A1, A2, . . . , Ak}, where Ai = {a ∈ Atest | L(a) = Li}, i = 1, . . . , p. C-index is a measure of such degree of match (or mismatch) between the true and the predicted rankings. It is shown as follows:

Table II presents the comparison in terms of gain between the original ACIμ,λ and its surrogate which is assigned by ∗. To this end, we have considered 2-additive, 3-additive and finally full ACI heuristic versions. Table II illustrates a significant improvement over the original version in many of the datasets considered. This can be explained by the fact that the heuristic version requires less complexity during optimization process, which obviously reduces the complexity of efforts and therefore quickly leading to a precise solution. This is especially interesting from an optimization point of view. Another advantage of this setting is in its efficient run-time. We have simultaneously measured run-time of each of the methods. Table III1 shows the average run-time of the given methods in milliseconds. For datasets with large number of attributes, the proposed heuristic approach is superior in terms of the runtime. However, for datasets with not so large number of attributes this difference is not noticeable. Note that, the original setting of ACI for k-additivity has almost the same runtime like the full-ACI.

	#instances	#criteria	#classes	source
Employee Selection (ESL)	488	4	4	WEKA WEKA
EmployeeRejection\Acceptance(ERA) Lecturers Evaluation (LEV)	1000	4	4	WEKA UCI UCI
Mamographic (MMG)	1000	4	3	UCI UCI
CPU	830	5	2	7
Car Evaluation (CEV) Breat-Cancer(BCC) DenBosch	209	6	2	UCI
Auto MPG	1728	6	4	
	286	7	2	
	120	8	2	
	398	8	6	

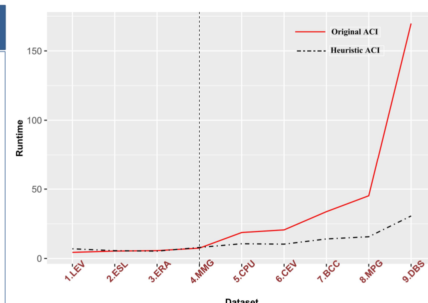
Table I. Data sets and their properties

methods / datasets	WA	CI	$ACI_{\mu,\lambda}$	$ACI^*_{\mu,\lambda}$-2-add	$ACI^*_{\mu,\lambda}$-3-add	$ACI^*_{\mu,\lambda}$ - full
ESL	.054 ± .022	.048 ± .026	.054 ± .022	.073 ± .037	.059 ± .033	.057 ± .028
ERA	.304 ± .054	.306 ± .051	.290 ± .056	.337 ± .053	.304 ± .055	.309 ± .051
LEV	.165 ± .038	.169 ± .039	.162 ± .036	.149 ± .028	.145 ± .024	.145 ± .024
MMG	.165 ± .139	.156 ± .027	.155 ± .022	.082 ± .068	.082 ±.068	.081 ± .067
CPU	.049 ± .040	.012 ± .014	.010 ± .014	.012 ± .006	.028 ± .023	.025 ± .021
CEV	.132 ± .014	.063 ± .019	.061 ± .022	.162 ± .015	.061 ±.003	.051 ± .010
BCC	.242 ± .107	.362 ± .079	.331 ± .084	.147 ± .090	.171 ±.152	.147 ± .151
DBS	.031 ± .062	.056 ± .065	.027 ± .032	.072 ± .040	.075 ± .034	.077 ± .035
MPG	.039 ± .032	.070 ± .015	.070±.015	.021 ± .018	.025 ± .019	.029 ± .024
Avg. rank						

TABLE II: Error in terms of the average C-Index ± standard deviation.2

methods / datasets	$ACI_{\mu,\lambda}$	$ACI^*_{\mu,\lambda}$-2-add	$ACI^*_{\mu,\lambda}$-3-add	$ACI^*_{\mu,\lambda}$-full
ESL	5.33	5.02	5.32	5.61
ERA	5.64	4.53	4.91	5.28
LEV	4.38	6.19	6.59	6.96
MMG	7.44	6.89	7.24	7.94
CPU	18.71	8.65	9.15	10.61
CEV	20.60	8.41	8.90	10.32
BCC	33.74	11.08	11.63	14.07
DBS	169.57	23.57	24.52	30.62
MPG	45.20	12.50	13.11	15.66

TABLE III: Runtime reported in milliseconds



Conclusions

The proposed version heuristic form of ACI offers an unique provision to obtain optimal parameters quickly while preserving the basic advantages of the ACI operator. The proposed approach is developed so as to reduce the complexity in each step, and to have an overall more favorable optimization settings which together could lead to finding a better solution efficiently overcoming the complexity of the ACI in its original form. It is also notable that in the original ACI setting, the solver is burdened with an exponential number of constraints, which naturally affects the quality of the solution obtained. The proposed form imparts simplicity to the optimization process that can be carried out conveniently resulting in an overall better performance both in terms of the results as well as the time required. It is also interesting to note that despite a large number of attributes (as the case in a few of the chosen datasets), our approach displays almost linear characteristics. Lastly but perhaps more importantly, the weights derived from the proposed approach obey monotonicity property that makes it very much desirable in most of human decision making applications.

As a future work, it would be useful to develop the measures such as Shapley index, interaction index and specifically fuzzy measure for the proposed form of ACI. Besides, it would be worth to investigate the application of the proposed form of ACI and such measures in a human decision making problem.

Contact

Sarang Balasaheb Bhasme
Moscow Institute of Physics and Technology
Email: sarangbhasme@phystech.edu
Phone: +79015366090

References

[1]M. Aggarwal. Compensative weighted averaging aggregation operators. Applied Soft Computing, 28:368– 378, 2015.

[2]M. Aggarwal. Generalized compensative weighted averaging aggregation operators. Computers & Industrial Engineering, 87:81 – 90, 2015.

[3]M. Aggarwal. Attitudinal choquet integrals and applications in decision making. International Journal of Intelligent Systems, 33(4):879–898, 2018.

[4]M. Aggarwal and A. Fallah Tehrani. Modelling human decision behaviour with preference learning. INFORMS Journal on Computing, 31(2):318–334, 2019.

[5]L. Berrah, G. Mauris, and J. Montmain. Monitoring the improvement of an overall industrial performance based on a choquet integral aggregation. Omega, 36(3):340–351, 2008.

[6]G. Choquet. Annales de l Institut Fourier, chapter Theory of capacities, pages 5: 131–295. 1953. [7]H. Daniels and B. Kamp. Applications of MLP networks to bond rating and house pricing. Neural Computation and Applications, 8:226–234, 1999.

[8]H. Dyckhoff and W. Pedrycz. Generalized means as a model of compensation connectives. Fuzzy Sets and Systems, 14:143–154, 1984.

[9]V. Fragnelli and S. Moretti. A game theoretical approach to the classification problem in gene expression data analysis. Computers & Mathematics with Applications, 55 (5):950–959, 2008.

[10]M. Grabisch. Fuzzy integral in multicriteria decision making. Fuzzy Sets and Systems, 69(3):279–298, 1995.

[11]M. Grabisch. The application of fuzzy integrals in multicriteria decision making. European Journal of Operational Research, 89:445–456, 1996.

[12]M. Grabisch. Fuzzy Measures and Integrals - Theory and Applications, chapter Fuzzy integral for classification and feature extraction, pages 415–434. Physica Verlag, 2000.

[13]M. Grabisch, J. Duchene, F. Lino, and P. Perny. Subjective evaluation of discomfort in sitting position. Fuzzy Optimization and Decision Making, 287-312:1(3), 2002.

[14]M. Grabisch, J. L. Marichal, R. Mesiar, and E. Pap. Aggregation functions: means. Information Sciences, 181:1–22, 2011.

[15]R. Krishnapuram and J. Lee. Fuzzy-connective-based hierarchical aggregation networks for decision making. Fuzzy Sets and Systems, 46 (1):11–27, 1992.

[16]D. Liginlal and T. T. Ow. On policy capturing with fuzzy measures. On policy capturing with fuzzy measures, 167:461–474, 2005.