Discovering Ordinal Attributes Through Gradual Patterns, Morphological Filters and Rank Discrimination Measures

Christophe Marsala, Anne Laurent, Marie-Jeanne Lesot, Maria Rifqi, Arnaud Castelltort

Discovering Ordinal Attributes through Gradual Patterns, Morphological Filters and Rank Discrimination Measures

C. Marsala[1], A. Laurent[2], M.-J. Lesot[1], M. Rifqi[3], and A. Castelltort[2]

[1] Sorbonne Université, CNRS, LIP6, F-75005 Paris, France
{firstname.lastname}@lip6.fr
[2] LIRMM, University of Montpellier, CNRS, Montpellier, France
{firstname.lastname}@umontpellier.fr
[3] LEMMA, University Panthéon-Assas, Paris, France
maria.rifqi@u-paris2.fr

Abstract. This paper proposes to exploit heterogeneous data, i.e. data described by both numerical and categorical features, so as to gain knowledge about the categorical attributes from the numerical ones. More precisely, it aims at discovering whether, according to a given data set, based on information provided by the numerical attributes, some categorical attributes actually are ordinal ones and, additionally, at establishing ranking relations between the category values. To that aim, the paper proposes the 3-step methodology OSACA, standing for Order Seeking Algorithm for Categorical Attributes: it first consists in extracting gradual patterns from the numerical attributes, to identify rich ranking information about the data; it then applies mathematical morphology tools, more precisely alternated filters, to induce an associated order on the categorical attributes. The third step evaluates the quality of the candidate rankings through measures derived from the rank entropy discrimination.

Keywords: Heterogeneous Data · Ordinal Attributes · Gradual Patterns · Rank Discrimination Measure · Mathematical Morphology.

1 Introduction

By definition, heterogeneous data are described using several types of features, including numerical and categorical attributes. Research issues about such data usually aim at simultaneously exploiting all the attributes, raising questions about how to combine the information they respectively provide. For instance, clustering tasks can then be addressed using relational approaches and appropriate distance measures, classification can be performed using decision trees, that can efficiently process both types of attributes.

This paper proposes another point of view, aiming at exploiting the information provided by the numerical attributes so as to gain knowledge about the categorical ones. More precisely, the goal is to determine whether some of the

Table 1. Illustrative toy data set.

Id	X	Y	colour
o_1	0	1	blue
o_2	1.2	1.5	blue
o_3	1.8	1.6	red
o_4	2.3	9.3	yellow
o_5	2.5	9.8	red
o_6	3.0	2.1	blue
o_7	4.8	3.2	yellow
o_8	5.0	8.5	yellow

categorical features actually are ordinal, and additionally identify the associated order.

For an illustration, consider the toy data set given in Table 1 where the data points are described by two numerical attributes and a categorical ones. It can be considered that, for this data set, the numerical attributes lead to the knowledge that the categorical attribute is ordinal, with partial order blue $\prec$ yellow.

To that aim, the paper proposes an original method called OSACA, standing for Order Seeking Algorithm for Categorical Attributes, that combines three tools, as detailed in Section 2, namely gradual patterns, mathematical morphology and rank entropy discrimination: gradual patterns are patterns of the form *the higher/lower a_1, ..., the higher/lower a_k* where a_i are numerical attributes. In the proposed approach, they are extracted to discover how the objects of the data base can be ranked. These rankings are then processed by mathematical morphology tools, more precisely alternated filters, to induce candidate orders on the categorical attributes. The candidates are finally evaluated through measures derived from the rank entropy discrimination.

The paper is organized as follows: Section 2 recalls some preliminaries, for the three types of tools, Section 3 describes the proposed approach and Section 4 presents illustrative results showing the relevance of OSACA. Section 5 concludes the paper and gives some directions for future works.

2   Preliminaries

This section provides some reminder about the three major tools used in the proposed method, successively gradual patterns, rank entropy measures and mathematical morphology tools.

Throughout the paper, $\Omega = \{o_1, \ldots, o_n\}$ denotes a set of n objects, or data points, described by a set $m + p$ attributes made of the union of the set $\mathcal{N}$, containing m numerical attributes, and the set $\mathcal{C}$, containing p categorical attributes. The value of attribute a for object o is denoted by $o[a]$.

2.1   Gradual Patterns

Gradual patterns extract linguistic knowledge from data described by numerical attributes that can be expressed by patterns of the form *the higher/lower a_1, ..., the higher/lower a_k*, where $a_i \in \mathcal{N}$, e.g.*the higher the budget, the higher the number of Champion's cup wins*. Initially introduced in the fuzzy implication formalism [5, 6, 9], gradual itemsets have then been interpreted as expressing constraints on the attribute covariations. Within this framework, several interpretations of the constraints have in turn been proposed, as regression [10], correlation of induced order [2, 11] or identification of compatible object subsets [3, 4]. Each interpretation is associated with the definition of a support to quantify the validity of gradual itemsets and with methods for the identification of the itemsets that are frequent according to these support definitions.

This section focuses on the approach based on the identification of compatible object subsets [3, 4], which is the one exploited in OCASA.

Definition 1 (Gradual Item). *A gradual item i is a pair $(a, *)$ where $a \in \mathcal{N}$ is a numerical attribute and $* \in \{\uparrow, \downarrow\}$ is the variation direction.*

For instance *budget* $\uparrow$ is a gradual item meaning *the higher the budget*.

Definition 2 (Gradual Pattern). *A gradual pattern P of size k is a set of k gradual items $\{(a_1, *_1), \ldots, (a_k, *_k)\}$, interpreted as their conjunction.*

For instance $\{(X, \uparrow), (Y, \uparrow)\}$ is a gradual pattern meaning *the higher X, the higher Y*.
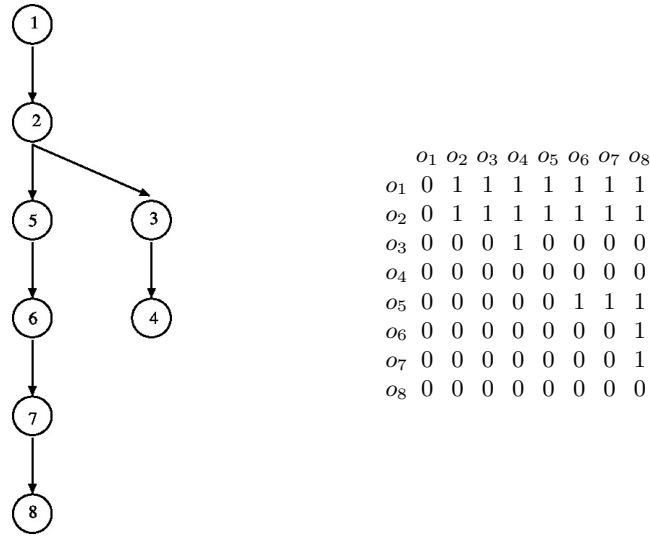
The main question is then the evaluation of the quality of candidate gradual patterns, with respect to the considered data set. One approach [3, 4] proposes to consider that a pattern is all the truer as it occurs frequently in the data, measuring the truth degree by its *support*. In regular frequent patterns and association rules, computing the support amounts to counting the number of objects containing the pattern [1]. For gradual pattern, counting the support requires to rank the objects with respect to the pattern. This relies on the definition of the pattern-induced precedence relation, which defines a partial order on the objects:

Definition 3 (Object Precedence w.r.t. a Gradual Pattern). *Given a data set Ω, two objects $o, o' \in \Omega$ and a gradual pattern $P = \{(a_1, *_1), \ldots, (a_k, *_k)\}$, the* precedence relation $o \prec_P o'$ *in Ω holds if and only $\forall j \in [1, k]$:*

- *if $*_j = \uparrow$ then $o[a_j] < o'[a_j]$*
- *if $*_j = \downarrow$ then $o[a_j] > o'[a_j]$*

For instance, considering the data set given in Table 1 and for the pattern $P = \{(X, \uparrow), (Y, \uparrow)\}$, it holds that $o_3 \prec_P o_4$, as $o_3[X] = 1.8 < 2.3 = o_4[X]$ and $o_3[Y] = 8.3 < 8.8 = o_4[Y]$.

The precedence relation leads to the definition of a precedence graph, where nodes represent objects and the directed edges the existence of a precedence

	o_1	o_2	o_3	o_4	o_5	o_6	o_7	o_8
o_1	0	1	1	1	1	1	1	1
o_2	0	1	1	1	1	1	1	1
o_3	0	0	0	1	0	0	0	0
o_4	0	0	0	0	0	0	0	0
o_5	0	0	0	0	0	1	1	1
o_6	0	0	0	0	0	0	0	1
o_7	0	0	0	0	0	0	0	1
o_8	0	0	0	0	0	0	0	0

Fig. 1. Precedence graph and matrix for pattern $\{(X,\uparrow),(Y,\uparrow)\}$ and data from Table 1.

relation, as illustrated on Figure 1. The graph can equivalently be represented by its adjacency matrix, also shown on Figure 1.

The precedence relation then leads to the definition of a support path:

Definition 4 (Support Path). *Given a data set Ω and a gradual pattern $P = \{(a_1, *_1), \ldots, (a_k, *_k)\}$, a support path p of length l is an ordered subset $\{o_{\rho_1}, \ldots, o_{\rho_l}\}$ containing l objects from Ω such that $\forall j \in [1, l-1]$, $o_{\rho_j} \prec_P o_{\rho_{j+1}}$.*

We denote $\mathcal{P}$ the set of all the paths supporting P.

Definition 5 (Support by Longest Path). *Given a data set Ω and a gradual pattern P, the* support *of P in Ω is defined by the length of the longest support path, relative to the total number of objects: $supp(P) = \dfrac{1}{|\Omega|} \max_{p \in \mathcal{P}}(length(p))$.*

For instance, for the pattern $P = (X \uparrow, Y \uparrow)$, we have the chains $\langle o_3, o_4 \rangle$ and $\langle o_1, o_2, o_3, o_5, o_6, o_7, o_8 \rangle$ (and all sub-chains from these chains). The support of P is thus $supp(P) = \frac{6}{8} = 0.75$.

Based on this quality criterion, one can then define the gradual patterns of interest, which must be both frequent and maximal:

Definition 6 (Frequent Gradual Pattern). *Given a data set Ω and a minimal support value minsup, a gradual pattern P is said to be* frequent *if we have $supp(P) \geq minsup$.*

Definition 7 (Maximal Frequent Gradual Pattern). *Given a data set Ω and a minimal support value minsup, a frequent gradual pattern P is said to be* maximal *if there does not exist any frequent pattern P' such that $P \subset P'$.*

The GARE algorithm [4] proposes an efficient method to mine the maximal frequent gradual patterns according to these definitions.

2.2   Rank Discrimination Measures

Rank discrimination measures [14, 15] have been studied in the setting of a specific classification task, when the class (or label) to be predicted is ordinal, and not categorical as it is usually the case. We propose to describe these data as *ordinally labelled*. The aim in such a task is then to preserve the rank knowledge on the class in the trained classifier and to highlight a gradual relation between the numerical attributes and the class.

Monotonic decision trees [15] for instance allow to address this task. Such decision trees are built from attributes that are the most gradually related to the class. They rely on the notion of rank discrimination measures: the latter are an extension of discrimination measures used in the attribute selection step of the tree building, so as to favour attributes whose behaviour satisfies the monotonicity aim.

It applies to ordinally labelled numerical datasets: first, the data description only contains numerical attributes, i.e. $\mathcal{C} = \emptyset$. Second, each data point is associated to a class by the labelling function $\lambda : \Omega \to C$ where $C = \{c_1, \ldots c_k\}$ is an ordered set of class, i.e. associated with a ranking $\prec_C$.

Rank discrimination measures are based on the concept of *dominance* [7]: any object $o \in \Omega$ is associated to the sets $[o]_{\lambda}^{\leq}$ and $[o]_{a}^{\leq}$ for any attribute a, respectively defined as

$$[o]_{\lambda}^{\leq} = \{o' \in \Omega \: : \: \lambda(o) \preceq_C \lambda(o')\}$$
$$[o]_{a}^{\leq} = \{o' \in \Omega \: : \: o[a] \leq o'[a]\}$$

In this paper, we focus on the rank Shannon entropy that corresponds to a rank version of the classical Shannon entropy. It is obtained by substituting the conditional probabilities with the corresponding measures obtained by means of the dominant sets, as shown in the next definition. The rank Shannon entropy can also be considered as a measure comparing the orders of two ordered sets: Ω_{λ} the ordered set of objects from Ω according to λ, and Ω_a the ordered set of objects from Ω according to a. We thus denote $H_S^*(\Omega_{\lambda}|\Omega_a)$ the rank Shannon entropy of attribute a with respect to the class labelling λ.

Definition 8 (Rank Shannon Entropy [8]). *For an ordinally labelled numerical dataset Ω, the* rank Shannon entropy *of attribute a with respect to the class labelling λ is computed as*

$$H_S^*(\Omega_{\lambda}|\Omega_a) = -\frac{1}{|\Omega|} \sum_{o \in \Omega} \log_2 \left(\frac{|[o]_{\lambda}^{\leq} \cap [o]_{a}^{\leq}|}{|[o]_{a}^{\leq}|} \right)$$

2.3   Mathematical Morphology to Induce Order

Mathematical morphology [17] defines a set of tools for the identification of spatial structures as the shape and size of objects, it has been extensively used for image processing and functional analysis. One-dimensional mathematical morphology [12, 13] applies to symbolic words, obtained as data transcriptions on a set of symbols. In particular, it has been proposed as an efficient approach to perform an automatic fuzzy partition method to discretise a numerical attribute: it makes it possible to identify class homogeneous intervals, tolerating some noise in the intervals, i.e. to highlight class homogeneous class kernels. One-dimensional mathematical morphology has also been applied to enrich gradual patterns through characterising clauses exploiting categorical attributes [16].

Considering a numerical attribute a whose universe X_a has to be partitioned, first, a "word" is built, made by the values of the class for all $o \in \Omega$ ordered by the corresponding values of a for o. Then, a smoothing of this word is done to highlight consistent kernels of class values. During this smoothing, maximally homogeneous sequences of letters are searched for to define kernels for the fuzzy sets. To obtain such a sequence, a *morphological filter* is applied to the word using a particular class as structuring element.

A morphological filter is defined as the composition of *opening* and *closure* operators, themselves defined as composition of *dilation* and *erosion* operators: according to the class c used as structuring element, the dilation operator enables the merging of two sequences of letters c separated by a "small" number of other classes; the erosion operator enables the deletion of very small sequences of letters c. The strength of the filter is defined as the number of repetitions of each operator in the sequence [13].

3   Proposed Method

This section describes the proposed OSACA approach which consists in three steps, described in turn in the following subsections: the first one exploits the numerical attributes to highlight rich ordering information, inducing an object ranking, based on gradual patterns extraction and their supporting paths. The second step considers these paths from the point of view of the categorical attributes. and processing the induced class words using a morphological alternated filter, so as to smooth it and identify kernels of homogeneous values and leads to a candidate (partial) ranking on the categorical attribute. The third step evaluates the candidate rankings and outputs the relevant ones, based on quality measures regarding both the ranking compatibility to the order induced by the gradual pattern and the number of involved objects.

3.1   Identification of Compatible Object Subsets

The first step exploits the numerical attributes to extract rich knowledge through gradual patterns using the covariation interpretation: they allow to maximally

combine the attributes so as to derive rich ordering information, inducing an object ranking.

We propose to apply the GARE algorithm [4], whose theoretical principles are recalled in Section 2.1, as it offers the additional advantage of providing a subset of objects satisfying the induced ranking.

Given the data set Ω, this step outputs a set of gradual patterns from the numerical attributes $\mathcal{N}$, each of them being associated to a set of supporting paths $\mathcal{P}$.

3.2   Construction of Candidate Rankings

In the second step, for each gradual pattern P extracted by GARE in the previous step, each associated supporting path p is considered in turn to build a tentative ranking on each categorical attribute $c \in \mathcal{C}$. In the following, we consider the categorical attribute as a labelling function: for any $o \in \Omega$, $\lambda(o) = o[c]$.

Given a path $p \in \mathcal{P}$, i.e. an ordered set of object $p = \langle o_1, \ldots, o_{|p|} \rangle$, and λ, the construction of a candidate ranking on C first builds the word made of the corresponding sequence of categorical values $w_p = \langle \lambda(o_1), \ldots, \lambda(o_{|p|}) \rangle$, denoted $w = \langle w_1, \ldots w_p \rangle$ below.

The principle then consists in smoothing w in order to highlight kernels for each class. As recalled in Section 2.3, morphological alternated filters are appropriate tools to perform such an operation. In the particular setting considered here, a specific dilation operator needs to be defined: it must indeed enable the deletion of any class distinct from the structuring element, which is not possible in the classical approaches [12, 13]. We thus introduce a *strong dilation* operator to achieve this aim and apply the induced alternated filter F.

The obtained filtered word $w'_p = F(w_p)$ then allows to define a partial order of the values taken by the considered attribute c: for each value c_i, the largest kernel associated to c_i in w' is retained (the first one in case of ties), and the order of these kernels provide the ranking of the associated values.

It can occur that no kernel for c_i survives after the filtering. In this case, the class is not concerned by the ranking and the result is a partial order $\prec_p$ of the values of C induced by p. We denote $\lambda^* \subseteq \lambda$ the subsets of labels that are comparable according to $\prec_p$. In the extreme case where $\lambda^* = \emptyset$, no order can be induced for this class: the corresponding attribute is only categorical and not, even partially, ordinal.

In the case of the toy data given in Table 1, considering the gradual pattern $P = \{(X, \uparrow), (Y, \uparrow)\}$, and the maximal support path $p = \langle o_1, o_2, o_5, o_6, o_7, o_8 \rangle$, the obtained word is $w_p = \langle$blue, blue, red, blue, yellow, yellow$\rangle$, which can be filtered to $w'_p = \langle$blue, blue, ·, ·, yellow, yellow$\rangle$, suggesting the ranking blue $\prec_p$ yellow.

3.3   Evaluation of a Candidate Ranking

In this step, any order $\prec_p$ identified in the previous step is evaluated by combining several quality criteria.

First, a notion of support is considered so as to quantify the proportion of data points concerned by the candidate ranking. This support is defined as the sum of the sizes of the kernels considered for building the ranking, divided by the path length. For the considered example, the support of the candidate partial ranking blue $\prec_p$ yellow is therefore $4/6 = 0.666$.

Second, the quality of the compatibility between the candidate ranking of categorical values and the pattern ranking is considered. It corresponds to a rank Shannon entropy applied to the subset of all data points Ω that are present in the considered path p with the orders $\prec_p$ on the categorical attribute and $\prec_P$ on the numerical attributes involved in the considered pattern.

It must be underlined that the induced ranking $\prec_p$ may not apply to all points in p, as it can be only partial, as is for instance the case for the considered toy example. Therefore, we introduce, as a second quality criterion of the candidate ranking, the partial rank Shannon entropy defined as follows:

Definition 9 (Partial Rank Shannon Entropy).
Given P a gradual pattern, $\prec_P$ the induced data ranking,
p a support path associated to P,
c a categorical attribute, $\prec_p$ the categorical value ranking induced by p on c,
and λ^* the subset of c values comparable according to $\prec_p$,
the partial rank Shannon entropy is defined as:

$$H_S^*(\Omega_{\lambda^*}|p) = -\frac{1}{|p|} \sum_{o \in p} \left(\log_2 \left(\frac{|[o]_{\lambda^*}^{\prec_p} \cap [o]^{\prec_P}|}{|[o]^{\prec_P}|} \right) \right)$$

It can be noted that, contrary to the rank Shannon entropy, the intersection $[o]_{\lambda^*}^{\prec_p} \cap [o]^{\prec_P}$ can be empty. In this case, the corresponding term in the sum is set to 0.

For instance, for the considered example, the partial rank Shannon entropy equals $-\frac{1}{6}\left(\log_2(\frac{5}{6}) + \log_2(\frac{5}{5}) + 0 + \log_2(\frac{3}{3}) + \log_2(\frac{2}{2}) + \log_2(\frac{1}{1}) \right) = 0.0659$.

4   Illustrative Results

A set of experiments has been conducted to evaluate the OCASA approach on a set with 100 randomly generated data. Each data point is described by two numerical attributes X and Y and is associated with a categorical attribute, also called class, among the unordered set {blue, red, yellow, green}.

4.1   Gradual Relation on X and Y but no Order on the Class

In the first experiment, represented on Fig. 2, the class values are not correlated with the numerical attributes: it is expected that no ranking is identified.

As there is a gradual relation between the two attributes X and Y, the gradual pattern extraction provides several paths for the gradual pattern $P =$
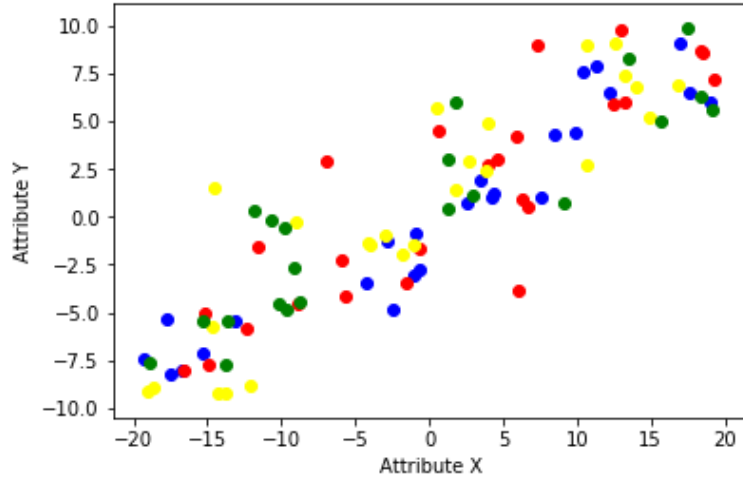
Fig. 2. Gradual relation on X and Y but no order on the classes.

$\{(X, \uparrow), (Y, \uparrow)\}$. There are 16 longest paths, each containing 34 objects. Paths are rather long as they contain one third of the objects of the dataset, which highlights the gradual relation.

When applying the morphological filtering on each of these paths, no kernel can be found and thus, as expected, no order on the class can be identified.

4.2   Gradual Relation on X and Y and Partial Order on the Class

In the second experiment, three of the four class values are generated with an underlying ranking (see Fig. 3): smaller values of the numerical attributes are associated with class value blue, middle values with yellow, and bigger values with green. The red class is uniformly spread among the universes.

As in the previous experiment, the gradual relation between the two numerical attributes X and Y is highlighted by the fact that several paths for the gradual pattern $\{(X, \uparrow), (Y, \uparrow)\}$ are found: 2160 longest paths, each containing 24 objects are found. They are still rather long as they contain one quarter of the objects of the dataset, which highlights the existing gradual relation between X and Y.

In this case, the ranking of the class values is successfully identified: when applying the morphological filtering on each of these paths, 1800 paths among the 2160 highlight the order blue $\prec$ yellow $\prec$ green, with a support ranging from 0.75 to 0.875 and a partial rank Shannon entropy very small, ranging from 0.0 to 0.008.

The other paths highlight the order blue $\prec$ green with a support ranging from 0.583 to 0.625.
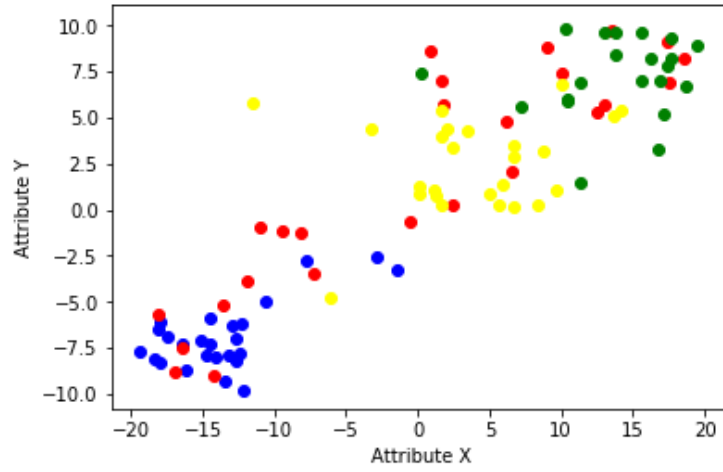
Fig. 3. Gradual relation on X and Y and partial order on the class.
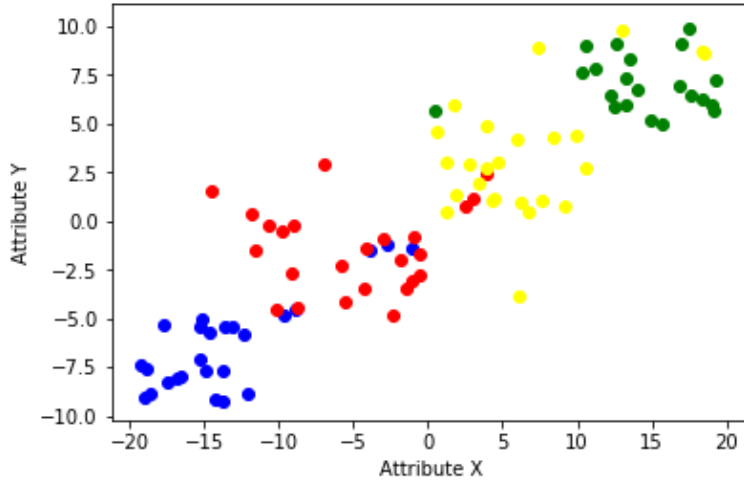

Fig. 4. Gradual relation on X and Y and order on the class

4.3   Gradual Relation on X and Y and Order on the Class

In the third experiment, all 4 class values are ordered (see Fig. 3): the expected ranking should indicate blue before red, before yellow, before green.

As in the previous experiment, the gradual relation between the two attributes X and Y is highlighted by the fact that several paths for the gradual pattern $\{(X, \uparrow), (Y, \uparrow)\}$ are found: 16 longest paths, each containing 34 objects.

As expected, after applying the morphological filtering on each of these paths, all paths induce the ranking blue $\prec$ red $\prec$ yellow $\prec$ green. In all cases, the support is 1 and the partial rank Shannon entropy is very small, ranging from 0.02 to 0.03.

5   Conclusion and Future Works

In this paper, we consider gradual patterns and use them to suggest ordinal attributes within the categorical ones. When such an attribute is pointed out, the OCASA method suggests a tentative ranking over its values and asses the quality of this suggestion with a ranking measure. When applied to toy data sets, the method appears to provide relevant results, and opens many directions for future works, beside experiments on real data sets.

One question is to determine which gradual patterns should be considered: OSACA considers all gradual patterns whose support is greater than the minimum support. It may be relevant to consider only patterns with maximum number of attributes or only patterns with maximum support value, knowing that the longer a pattern, the lower the support, due to the anti-monotonicity property. Besides, it may be relevant to choose a subset of categorical attributes to be studied, or even only one. Indeed, OSACA scans all categorical attributes from the database while some of them may not be interesting for end-users. Finally, we may consider evaluating the quality of the method by computing a kind of precision and recall describing on the one hand to which extent ordinal attributes have been retrieved, and on the other hand to which extent the orders suggested are relevant.

References

1. Agrawal, R., Imieliński, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data. pp. 207–216. SIGMOD '93, ACM, New York, NY, USA (1993)
2. Berzal, F., Cubero, J.C., Sanchez, D., Vila, M.A., Serrano, J.M.: An alternative approach to discover gradual dependencies. Int. Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **15**(5), 559–570 (2007)
3. Di Jorio, L., Laurent, A., Teisseire, M.: Fast extraction of gradual association rules: A heuristic based method. In: Proc. of CSTST'08 (2008)
4. Di Jorio, L., Laurent, A., Teisseire, M.: Mining frequent gradual itemsets from large databases. In: Proc. of IDA'09 (2009)
5. Dubois, D., Prade, H.: Gradual inference rules in approximate reasoning. Information Sciences **61**(1-2), 103–122 (1992)
6. Galichet, S., Dubois, D., Prade, H.: Imprecise specification of ill-known functions using gradual rules. Int. Journal of Approximate Reasoning **35**(3), 205–222 (2004)
7. Greco, S., Matarazzo, B., Slowinski, R.: Rough approximation by dominance relations. Int. J. of Int. Sys. **17**(2), 153–171 (2002)
8. Hu, Q.H., Guo, M.Z., Yu, D.R., Liu, J.F.: Information entropy for ordinal classification. Science China Information Sciences **53**, 1188–1200

9. Hüllermeier, E.: Implication-based fuzzy association rules. In: Proc. of PKDD'01. pp. 241–252 (2001)
10. Hüllermeier, E.: Association rules for expressing gradual dependencies. In: Proc. of PKDD'02. pp. 200–211 (2002)
11. Laurent, A., Lesot, M.J., Rifqi, M.: Graank: Exploiting rank correlations for extracting gradual dependencies. In: Proc. of FQAS'09 (2009)
12. Marsala, C.: Fuzzy partitioning methods. In: Pedrycz, W. (ed.) Granular Computing: an Emerging Paradigm, pp. 163–186. Springer-Verlag (2001)
13. Marsala, C., Bouchon-Meunier, B.: Fuzzy partioning using mathematical morphology in a learning scheme. In: Fifth IEEE Int. Conf. on Fuzzy Systems. vol. 2, pp. 1512–1517. New Orleans, USA (1996)
14. Marsala, C., Petturiti, D.: Hierarchical model for rank discrimination measures. In: European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty (ECSQARU). pp. 412–423. Springer (2013)
15. Marsala, C., Petturiti, D.: Rank discrimination measures for enforcing monotonicity in decision tree induction. Information Sciences **291**, 143–171 (Jan 2015)
16. Oudni, A., Lesot, M.J., Rifqi, M.: Characterisation of gradual itemsets through" especially if' clauses based on mathematical morphology tools. In: Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2013) (2013)
17. Serra, J.: Introduction to mathematical morphology. Computer vision, graphics, and image processing **35**(3), 283–305 (1986)