International Journal of Approximate Reasoning 187 (2025) 109548

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Flexible categorization using formal concept analysis and Dempster-Shafer theory

Marcel Boersma [e,f], Krishna Manoorkar [a,f,*,1], Alessandra Palmigiano [a,b], Mattia Panettiere [a,f], Apostolos Tzimoulis [a], Nachoem Wijnberg [c,d]

[a] School of Business and Economics, Vrije Universiteit, Amsterdam, the Netherlands
[b] Department of Mathematics and Applied Mathematics, University of Johannesburg, South Africa
[c] College of Business and Economics, University of Johannesburg, South Africa
[d] Faculty of Economics and Business, University of Amsterdam, the Netherlands
[e] Computational Science Lab, University of Amsterdam, Amsterdam, the Netherlands
[f] KPMG, Amstelveen, the Netherlands

ARTICLE INFO

Keywords:
Formal concept analysis
Dempster-Shafer theory
Learning algorithm
Auditing
Categorization

ABSTRACT

Based on the intuitive idea that sets of objects or entities can be categorized in very different ways, and that some ways to categorise objects are better than others, depending on the purpose of the categorization, in this paper, a formal framework is introduced for parametrically generating a space of possible categorizations of a set of objects, based on the features which individual agents or groups thereof regard as relevant (formally encoded in the notion of *interrogative agenda*). This formal framework accounts both for two-valued (crisp), and for many-valued (fuzzy) judgments about the relevance of given features, and introduces ways to aggregate individual agendas to group agendas. As an application on this framework, we discuss a machine-learning meta-algorithm for outlier detection and classification which provides local and global explanations of its results.

1. Introduction

Categories are cognitive tools used both by humans and machines to organize experience, understand and function in their environment, and understand and interact with each other, by grouping things together which can be meaningfully compared and evaluated. Categorization is the basic operation humans perform e.g. when they relate experiences/actions/objects in the present to those in the past, thereby recognizing them as instances of the same type and being able to compare them with each other. This is what humans do when trying to understand what an object is or does, or what a situation means, and when making judgments or decisions based on experience. Categorization is the single cognitive mechanism underlying *meaning*-attribution, *value*-attribution and *decision-making*, and therefore it is the background environment in which these three cognitive processes can formally be analyzed in their relationships to one another. Nowadays, categories are key to the theories and methodologies of several research areas at the interface of social sciences and AI, not only because categories underlie a wide array of phenomena key to these areas, spanning from the mechanisms of perception [57], to the creation of languages [3], social identities and cultures [49], but also because category-

* Corresponding author.
E-mail address: k.b.manoorkar@vu.nl (K. Manoorkar).
1 Krishna Manoorkar is supported by the NWO grant KIVI.2019.001 awarded to Alessandra Palmigiano.

https://doi.org/10.1016/j.ijar.2025.109548
Received 7 April 2025; Received in revised form 14 July 2025; Accepted 6 August 2025

Available online 12 August 2025
0888-613X/

formation is core to the development of data-analytic techniques based on AI, such as data mining [46,39], text analysis [11], social network analysis [24,30], which have become fundamental tools in empirical sciences.

Contributions, motivation, and related work. Categorization fundamentally shapes how agents interpret and evaluate any given set (of objects or data), yet any such set can be categorized in multiple ways, each tailored to different objectives. For instance, the experienced auditors' evaluation of evidence is rooted in a process of categorization of the various pieces of evidence (e.g. financial transactions) which is possibly very different from the "official" categorization system through which the evidence is presented in the self-reported financial statement of the given firm. These categories provide the context of evaluation in which different pieces of evidence are compared with/against each other.

Motivated by this practice, we introduce a formal, parametric framework that systematically spans the space of categorization systems over a fixed set of entities, and explicitly links each categorization system to an agent's goals, evaluative criteria, and epistemic stance. By doing so, the present framework identifies the categorization schemes that are optimal for specific tasks, such as outlier detection and supervised classification.

Interdisciplinary research in cognitive science and management has long demonstrated that changes in category structures influence reasoning and decision making [62,41,14,4,29]. Computationally, the present approach echoes the goals of feature selection (FS), which seeks the most informative dimensions for predictive tasks [20,43]. However, while traditional FS methods treat subsets of features as purely statistical artifacts, we model the choice of features as the expression of an agent's underlying agenda and priorities. Because FS algorithms often differ in stability and effectiveness, ensemble-based strategies have been proposed to improve robustness [47,64,67]. In the present paper, we adapt this ensemble methodology to derive stable categorization systems from non-crisp agendas, and introduce a meta-learning algorithm (Algorithm 1) that leverages these ensembles for both classification and outlier detection.

Within explainable AI, FS has emerged as a key technique to simplify models and highlight significant attributes (or significant combinations of attributes) for human comprehension [44,45,34]. Building on this technique, in [9], Algorithm 1 was used to produce (possibly fuzzy) interpretable categorization systems suitable for outlier detection tasks. Doing so was shown to yield effective unsupervised and supervised algorithms for outlier detection, providing both local and global explanations for its results. In the present paper, the theoretical background is developed and presented that provides a formal foundation for such algorithms and suggests further substantial avenues of exploration of this agenda-driven categorization framework.

Interrogative agendas. The final outcome of the auditing process is the formation of a qualitative opinion, by expert auditors, on the fairness and completeness of a given firm's financial accounts. Towards the formation of their opinion, the auditors might not attribute the same importance to all the features of the pieces of evidence, and they might also disagree with regard to the relative importance of certain features. In the present framework, the different epistemic attitudes of agents are captured by the notion of *interrogative agenda* [23]. This notion can be understood as the 'questions' that an agent sets to resolve (by gathering information) before making a decision. Interestingly, the level of *expertise* of an agent can be captured by how good the questions they ask are. In the present framework, a simple way for representing an agent's interrogative agenda is as a designated set of features, i.e. the set of features that the agent considers significantly more relevant than others. However, in many cases, an agent's agenda may not be realistically approximated in such a simple way, but might consist of different relevance or importance values assigned to different sets of features. Such agendas will be represented by Dempster-Shafer mass functions. Independently of their mathematical representation, interrogative agendas will induce categorization systems on a given set of entities, and hence parametrize the space of categorization systems.

Background theory. The present framework is set within *Formal Concept Analysis* (FCA) [65,5], and is based on *formal contexts* [65], i.e. structures consisting of domains A (of objects) and X (of features), and binary relations I between them. The running example mentioned above concerns the formal context arising from a network (bipartite graph) of financial transactions (i.e. a *financial statements network*) [10].

Mathematically, bipartite graphs and formal contexts are isomorphic structures; moreover, both types of structures are used to represent databases [65,60,37,48,31,56,69,13]. Setting the present framework on formal contexts allows for access to a mathematically principled way to generate categorization systems, in the form of the construction of the lattice of *formal concepts* of a formal context [7], as well as a suitable base for expanding the present formal framework to represent and support vagueness, epistemic uncertainty, evidential reasoning, and incomplete information [66,16,15,25]. In particular, representing categorization systems as lattices allows for hierarchical, rather than flat, categorizations of objects, as well as for a more structured control of the categorization, based on the generation of categories from arbitrary subsets of objects or features.

Meta-algorithm. Based on the present framework, we discuss an explainable meta-algorithm for outlier detection or classification, introduced in [1], in which the categorization systems are not pre-specified but are 'learned', based on the accuracy of their prediction. This algorithm is better understood in the light of the framework presented in this paper.

Structure of the paper. In Section 2, we collect the basic definitions and facts pertaining to FCA and Dempster-Shafer theory. In Section 3, we introduce the (crisp and non-crisp) formal framework for generating a set of categorization systems, over a given set of entities, parametrized by interrogative agendas, and introduce ways of aggregating (crisp and non-crisp) agendas. In Section 4, we illustrate this framework by means of our running example on financial statements network.

In Section 5, we introduce a methodology for associating single categorization systems to non-crisp agendas (i.e. Dempster-Shafer mass functions), and define a partial order on non-crisp agendas with respect to which this mapping is order-preserving. In Section 6, we compare this partial order to other orderings of mass functions in the literature, and study its properties with respect to the aggregated agendas. In Section 7, we discuss a meta-learning algorithm for outlier detection and classification. We conclude and mention some directions for future research in Section 8.

2. Preliminaries

Formal contexts and their concept lattices A *formal context* [28] is a structure $\mathbb{P} = (A, X, I)$ such that A and X are sets, and $I \subseteq A \times X$ is a binary relation. Formal contexts can be thought of as abstract representations of databases, where elements of A and X represent objects and features, respectively, and the relation I records whether a given object has a given feature. Every formal context as above induces maps $I^{(1)} : \mathcal{P}(A) \to \mathcal{P}(X)$ and $I^{(0)} : \mathcal{P}(X) \to \mathcal{P}(A)$, respectively defined by the assignments

$$I^{(1)}[B] := \{x \in X \mid \forall a(a \in B \Rightarrow aIx)\} \quad \text{and} \quad I^{(0)}[Y] := \{a \in A \mid \forall x(x \in Y \Rightarrow aIx)\}.$$

A *formal concept* of $\mathbb{P}$ is a pair $c = ([\![c]\!], (\![c]\!))$ such that $[\![c]\!] \subseteq A$, $(\![c]\!) \subseteq X$, and $I^{(1)}[[\![c]\!]] = (\![c]\!)$ and $I^{(0)}(\![c]\!) = [\![c]\!]$. A subset $B \subseteq A$ (resp. $Y \subseteq X$) is said to be *closed*, or *Galois-stable*, if $\mathsf{Cl}_1(B) = I^{(0)}[I^{(1)}[B]] = B$ (resp. $\mathsf{Cl}_2(Y) = I^{(1)}[I^{(0)}[Y]] = Y$). The set of objects $[\![c]\!]$ is the *extension* of the concept c, while the set of features $(\![c]\!)$ is its *intension*.[2] The set $\mathrm{L}(\mathbb{P})$ of the formal concepts of $\mathbb{P}$ can be partially ordered as follows: for any $c, d \in \mathrm{L}(\mathbb{P})$,

$$c \leq d \quad \text{iff} \quad [\![c]\!] \subseteq [\![d]\!] \quad \text{iff} \quad (\![d]\!) \subseteq (\![c]\!). \tag{2.1}$$

With this order, $\mathrm{L}(\mathbb{P})$ is a complete lattice, the *concept lattice* $\mathbb{P}^+$ of $\mathbb{P}$. As is well known, any complete lattice $\mathbb{L}$ is isomorphic to the concept lattice $\mathbb{P}^+$ of some formal context $\mathbb{P}$ [7]. Throughout this paper, we will often identify the concept lattice $\mathbb{P}^+$ with the lattice of the extensions of its formal concepts, ordered by inclusion, and will write e.g. $G \in \mathbb{P}^+$ for a subset $G \subseteq A$ to signify that G is the extension of some formal concept of $\mathbb{P}$.

Discretization of continuous attributes and conceptual scaling. The framework discussed above can also be applied to cases in which the relation $I \subseteq A \times X$ can take continuous values, modulo a process known as *conceptual scaling* [27]. Scaling is an important part of most FCA-based techniques and has been studied extensively [27,51,52]. Choosing the correct scaling method depends on the specific task the concept lattice is used for.

Belief, plausibility and mass functions Here we recall standard notation and terminology from Dempster-Shafer theory [58,68]. A *belief function* (cf. [59, Chapter 1]) on a set S is a map $\mathrm{bel} : \mathcal{P}(S) \to [0,1]$ such that $\mathrm{bel}(S) = 1$, and for every $n \in \mathbb{N}$,

$$\mathrm{bel}(A_1 \cup, \cdots, \cup A_n) \geq \sum_{\emptyset \neq I \subseteq \{1,\cdots,n\}} (-1)^{|I|+1} \mathrm{bel}\left(\bigcap_{i \in I} A_i\right).$$

A *plausibility function on* S is a map $\mathrm{pl} : \mathcal{P}(S) \to [0,1]$ such that $\mathrm{pl}(S) = 1$, and for every $n \in \mathbb{N}$,

$$\mathrm{pl}(A_1 \cup A_2 \cup, \cdots, \cup A_n) \leq \sum_{\emptyset \neq I \subseteq \{1,2,\cdots,n\}} (-1)^{|I|+1} \mathrm{pl}\left(\bigcap_{i \in I} A_i\right).$$

For any set X, let $\overline{X}$ be its complete $S \setminus X$. Belief and plausibility functions on sets are interchangeable notions: for every belief function bel as above, the assignment $X \mapsto 1 - \mathrm{bel}(\overline{X})$ defines a plausibility function on S, and for every plausibility function pl as above, the assignment $X \mapsto 1 - \mathrm{pl}(\overline{X})$ defines a belief function on S. Let S be any set.

A *(Dempster-Shafer) mass function* is a map $m : \mathcal{P}(S) \to [0,1]$ such that $\sum_{X \subseteq S} m(X) = 1$.

On finite sets, belief (resp. plausibility) functions and mass functions are interchangeable notions: any mass function m as above induces the belief function $\mathrm{bel}_m : \mathcal{P}(S) \to [0,1]$ defined as

$$\mathrm{bel}_m(X) := \sum_{Y \subseteq X} m(Y) \qquad \text{for every } X \subseteq S, \tag{2.2}$$

and a plausibility function

$$\mathrm{pl}_m(X) := \sum_{Y \cap X \neq \emptyset} m(Y) \qquad \text{for every } X \subseteq S. \tag{2.3}$$

Conversely, any belief function bel as above induces the mass function $m_{\mathrm{bel}} : \mathcal{P}(S) \to [0,1]$ defined as

[2] The symbols $[\![c]\!]$ and $(\![c]\!)$, respectively denoting the extension and the intension of a concept c, have been introduced and used in the context of a research line aimed at developing the logical foundations of categorization theory, by regarding formulas as names of categories (formal concepts), and interpreting them as formal concepts arising from given formal contexts [16,15,25,17,19].

$$m_{\mathrm{bel}}(X) := \mathrm{bel}(X) - \sum_{Y \subseteq X} (-1)^{|X \setminus Y|} \mathrm{bel}(Y) \quad \text{for all } X \subseteq S. \tag{2.4}$$

For any mass function $m : \mathcal{P}(X) \to [0,1]$, its associated *quality function* q_m is

$$q_m(Y) = \sum_{Y \subseteq Z} m(Z) \qquad \text{for all } Y \subseteq X.$$

3. Categorizations induced by interrogative agendas

In this section, we introduce two multi-agent frameworks, each of which represents the interrogative agendas associated with *individual* agents and *groups* of agents, as well as the categorizations induced by these agendas. These two frameworks differ in the way interrogative agendas are represented; namely as *subsets* of features, and as *Dempster-Shafer mass functions* over the set of features, respectively. We refer to the interrogative agendas represented in the former way as the *crisp* ones, and to those represented in the latter way as the *non-crisp* ones.[3]

Crisp framework. We consider tuples $(\mathbb{P}, C, R)$ such that $\mathbb{P} = (A, X, I)$ is a finite (many-valued) formal context, C is a finite set, and $R \subseteq X \times C$ is a binary relation. The formal context $\mathbb{P}$ captures, as usual, the set A of the objects to be categorized, and the set X of features attributed to each object through the incidence relation I; the set C is understood as a set of agents, and the relation R associates any agent $j \in C$ with their *(crisp) interrogative agenda*, represented as the set $X_j := R^{-1}[j] = \{x \in X \mid xRj\}$ of features $x \in X$ which j considers relevant.

Interrogative agendas can also be associated with *coalitions* of agents as follows: for any $c \subseteq C$, we let

$$\Diamond c := \bigcap \{X_j \mid j \in c\} \qquad \text{and } \rhd c := \bigcup \{X_j \mid j \in c\}$$

denote the *common* and the *distributed* interrogative agendas associated with c, respectively.

For any $Y \subseteq X$, let $\mathbb{P}_Y := (A, Y, I_Y)$, where $I_Y := I \cap (A \times Y)$. Hence, if Y is an interrogative agenda, Y induces a categorization system on the objects of A, which is represented by the concept lattice $\mathbb{P}^+_Y$.

Notice that if $Z \subseteq Y \subseteq X$, and $B \in \mathbb{P}_Z$ (i.e. $B = \{a \in A \mid \forall x(x \in Z' \Rightarrow xI_Z a)\}$ for some $Z' \subseteq Z$) then $B \in \mathbb{P}_Y$; indeed, to show that some $Y' \subseteq Y$ exists such that $B = \{a \in A \mid \forall x(x \in Y' \Rightarrow xI_Y a)\}$, it is enough to let $Y' := Z'$. This shows that, for every $B \in \mathbb{P}^+$, the set $\mathcal{V}_B := \{Y \subseteq X \mid B \in \mathbb{P}^+_Y\}$ is upward closed w.r.t. inclusion. Consequently, for any coalition c, and any agent j in coalition c,

$$\mathbb{P}^+_{\Diamond c} \subseteq \mathbb{P}^+_{X_j} \subseteq \mathbb{P}^+_{\rhd c} \tag{3.1}$$

The parametric structure of $\{\mathbb{P}^+_Y \mid Y \subseteq X\}$ allows for the possibility to select those categorization systems with 'meaningful categories' relative to the task at hand: larger agendas induce finer categorization systems, capable of making more distinctions among objects, and smaller agendas induce coarser categorization systems, potentially suitable for e.g. identifying outliers while reducing the number of false positives.

Non-crisp framework. We consider tuples $(\mathbb{P}, C, \mathcal{M})$ such that $\mathbb{P}$ and C are as above, and $\mathcal{M} = \{m_j : \mathcal{P}(X) \to [0.1] \mid j \in C\}$ is a C-indexed set of Dempster-Shafer mass functions. Intuitively, if $j \in C$ and $Y \subseteq X$, then the value $m_j(Y)$ represents agent j's preference to use the concept lattice $\mathbb{P}^+_Y$ associated with Y as a categorization system. The following definition captures particularly interesting cases:

Definition 1. For any Dempster-Shafer mass function $m : \mathcal{P}(S) \to [0,1]$, a set $Y \subseteq S$ is a *focal set* of m if $m(Y) > 0$. If m has at most one focal set $X \subset S$, then m is *simple*.

For instance, if we only have information about agent j's opinion on the relevance of a certain subset $Y \subseteq X$ (quantified as $\alpha \in [0,1]$), this information can be encoded by representing the agenda of j as the simple mass function m_j such that $m_j(Y) = \alpha$ and $m_j(X) = 1 - \alpha$. Assigning the remaining mass to X is motivated by the idea that if an agent is unsure about which categorization is preferable for a given task, then the agent will prefer the finest possible categorization i.e. the categorization generated by all available features, so as to not ignore any features that may be relevant to the given task. For example, the FCA-based outlier detection algorithm used in [9] is more likely to flag an object as an outlier in a finer categorization than in a coarser categorization. However, if an agent believes that a coarser categorization (i.e. one associated with a smaller agenda) is already capable of flagging outliers, then the agent will prefer this one, for the sake of avoiding false positives. However, if an agent is unsure whether a smaller agenda suffices to flag outliers, that agent will prefer the finest categorization system available so to avoid false negatives.

If agent j has opinions on the relevance of each individual feature, the agenda of j can be represented by a mass function $m_j : \mathcal{P}(X) \to [0,1]$, for any $Y \subseteq X$,

[3] While an interrogative agenda modelled as a subset gives rise to a single categorization system, when modelled as a Dempster Shafer mass function, it gives rise to different priority or preference values assigned to different sets of features, which in turn induces a mass function on a whole spectrum of categorization systems.

$$m_j(Y) = \sum_{y\in Y} v_j(y),$$

where $v_j(y)$ is the (normalized) importance value assigned by j to each $y \in X$.

Each $m_j \in \mathcal{M}$ induces a probability mass function $m'_j : \mathcal{R} \to [0,1]$, where $\mathcal{R} := \{\mathbb{P}^+_Y \mid Y \subseteq X\}$, defined by the assignment $m'_j(\mathbb{P}^+_Y) = m_j(Y)$. Intuitively, for any $Y \subseteq X$, the value $m'_j(\mathbb{P}^+_Y)$ represents the extent to which agent j prefers the categorization system induced by Y. Moreover, each m_j induces a probability function $p_j : \mathcal{P}(\mathcal{P}(X)) \to [0,1]$ defined by the assignment $p_j(\mathcal{V}) = \sum_{Y\in\mathcal{V}} m_j(Y)$ for any $\mathcal{V} \subseteq \mathcal{P}(X)$.

For any coalition c, we can associate (non-crisp) interrogative agendas $\oplus c$, $\Diamond c$, and $\rhd c$ with c as follows:

1. For any $Y \subseteq X$, $Y \neq \varnothing$

$$(\oplus c)(Y) = \frac{\sum\{\prod_{j\in c} m_j(Z_j) \mid \bigcap_{j\in c} Z_j = Y\}}{\sum\{\prod_{j\in c} m_j(Z_j) \mid \bigcap_{j\in c} Z_j \neq \varnothing\}} \tag{3.2}$$

and $(\oplus c)(\varnothing) = 0$. This aggregation is the Dempster-Shafer combination [59] of the mass functions m_j for $j \in c$. The normalization in the above rule allows us to ignore completely contradictory agendas (i.e. agendas with empty intersection), thus giving more weight to issues which have consensus of the agents. However, in some scenarios, we may want to allow mass on the empty set of features (which corresponds to a categorization in which all elements are in same category). The value $m(\varnothing)$ describes the agent's preference for categorization with only one class. For such scenarios, rather than taking $\oplus c$ as the aggregated agenda, we may use the unnormalized mass function $\Diamond c$ defined as follows:

$$(\Diamond c)(Y) = \sum\{\prod_{j\in c} m_j(Z_j) \mid \bigcap_{j\in c} Z_j = Y\} \tag{3.3}$$

2. For any $Y \subseteq X$,

$$(\rhd c)(Y) = \sum\{\prod_{j\in c} m_j(Z_j) \mid \bigcup_{j\in c} Z_j = Y\} \tag{3.4}$$

Remark 2. Crisp agendas can be regarded as non-crisp agendas with a single focal element Z_j for every agent j. In this case, for every coalition c, the mass functions $\Diamond c$ and $\rhd c$ also have single focal elements $\bigcap_{j\in c} Z_j$ and $\bigcup_{j\in c} Z_j$, respectively, which coincide with the crisp agendas $\Diamond c$ and $\rhd c$. This justifies our use of the same symbols for these operations in the crisp and non-crisp setting.

Several rules have been used in Dempster-Shafer theory for aggregating preferences of different agents [58,25], and their applicability to the present framework is still widely unexplored. A concrete scenario involving non-crisp agendas is discussed in the next section.

4. Example: financial statements network

In this section, we illustrate the ideas discussed above by way of an example in the context of financial statements networks. These are a type of data structure for organizing the information contained in the journal entries of a company, which record the transfers from one set of financial accounts to another set. These entries are generated by their underlying *business processes*, which can be formally defined as follows [10,8]:

$$a : \sum_{1\le i\le m} \alpha_i x_i \implies \sum_{1\le j\le n} \beta_j y_j \tag{4.1}$$

where m is the number of credited financial accounts, n is the number of debited financial accounts, α_i is the relative amount with respect to the total credited, and β_j the relative amount with respect to the total debited. The arrow represents the flow of money between the accounts.

A *financial statements network* is a bipartite digraph $\mathbb{G} = (A \cup X, E)$ in which A is the set of business processes, X is the set of financial accounts, and the (many-valued) directed edges in $E \subseteq X \times A$ record information on (the share of) a given financial account in a given business process. Clearly, each such $\mathbb{G} = (A \cup X, E)$ can equivalently be represented as a many-valued formal context $\mathbb{P} = (A, X, I)$ where I is the converse of E. For each business process $a \in A$, the (weighted) edges between nodes x_i (resp. y_j) and a are the coefficients α_i (resp. β_j) in (4.1).

Consider the financial statements network presented in Table 1 of Appendix A, with business processes $A := \{a_1, a_2, \dots, a_{12}\}$ and financial accounts $X := \{x_1, x_2, \dots, x_6\}$ specified as follows:

x_1	tax	x_2	revenue	x_3	cost of sales
x_4	personnel expenses	x_5	inventory	x_6	other expenses

Table 2 encodes the many-valued formal context $\mathbb{P} = (A, X, I)$ extracted from this database. Each cell of Table 2 reports the value of the relation $I : A \times X \to [-1, 1]$, which, for any process a and account x, represents the share of x in a.

Using interval scaling,[4] we convert the many-valued formal context $\mathbb{P} = (A, X, I)$ into the 2-valued formal context $\mathbb{P}^{(s)} = (A, X^{(s)}, I^{(s)})$, for $s \in \mathbb{N}$, where $X^{(s)} := \{x_{ik} \mid 1 \le i \le 6, 1 \le k \le s\}$, and $I^{(s)} \subseteq A \times X^{(s)}$ is such that $aI^{(s)}x_{ik}$ iff $I(a, x_i) \in \left[-1 + \frac{2(k-1)}{s}, -1 + \frac{2k}{s}\right]$.

The concept lattice corresponding to $\mathbb{P}^{(s)}$ when $s = 5$ is shown in Fig. 8. Its associated concept lattice represents the categorization system obtained by considering all the features (financial accounts) in the database, and hence, as it faithfully captures all the information of $\mathbb{P}^{(s)}$, it is the finest categorization obtainable by any of our proposed methods for this database.

Consider the set of agents $C = \{j_1, j_2, j_3\}$, and let us assume that agent j_1 is interested in the financial accounts x_1, x_2, and x_5, agent j_2 in x_1, x_2, and x_3, while agent j_3 is interested in x_1, and x_3 with various degrees; their interests can be represented by the following relation $R \subseteq X^{(s)} \times C$:

$$R = \{(x_{1k}, j_1), (x_{2k}, j_1), (x_{5k}, j_1), (x_{1k}, j_2), (x_{2k}, j_2), (x_{3k}, j_2), (x_{1k}, j_3), (x_{3k}, j_3) \mid 1 \le k \le s\},$$

which gives rise to the interrogative agendas $Y_i := R^{-1}[j_i]$ for $1 \le i \le 3$. Their associated categorization systems (i.e. the concept lattices associated with $\mathbb{P}^{(s)}_{Y_i}$) are shown in Fig. 1, 2, and 3, respectively. Let c be the coalition of j_1, j_2, and j_3. Then, the categorization systems induced by the common agenda $\Diamond c$, and the distributed agenda $\rhd c$ of coalition c are shown in Figs. 5, and 4, respectively.

Let us now assume that tax is the most relevant account for j_1 and j_2, while j_2 believes that tax and revenues are also relevant, although somewhat less than tax alone, and j_3 believes that tax, revenues and expenses are very relevant. Then their interrogative agendas can be represented by the following mass functions m_1, m_2, and m_3, respectively, for $1 \le k \le s$:

$$m_1(\{x_{1k}\}) = 0.6 \quad m_1(X) = 0.4,$$

$$m_2(\{x_{1k}\}) = 0.5 \quad m_2(\{x_{1k}, x_{2k}\}) = 0.3 \quad m_2(X) = 0.2,$$

$$m_3\{x_{1k}, x_{2k}, x_{6k}\} = 0.9 \quad m_3(X) = 0.1.$$

The most preferred categorization system (i.e. the one with the highest induced mass) according to m_1 and m_2 is the same and is shown in Fig. 5, while the most preferred categorization according to m_3 is shown in Fig. 6. For the coalition c of agents j_1, j_2, and j_3, the agendas $\oplus c = \Diamond c$ and $\rhd c$ are as follows:

$$(\oplus c)(\{x_{1k}\}) = 0.8 \quad (\oplus c)(\{x_{1k}, x_{2k}\}) = 0.12 \quad (\oplus c)(\{x_{1k}, x_{2k}, x_{6k}\}) = 0.072 \quad (\oplus c)(X) = 0.008,$$

$$(\rhd c)(\{x_{1k}, x_{2k}, x_{6k}\}) = 0.432 \quad (\rhd c)(X) = 0.568$$

The most preferred categorization systems according to the coalition agendas $\oplus c = \Diamond c$, and $\rhd c$ are in Fig. 5, and Fig. 8, respectively.

5. The stability-based method

As discussed in Section 3, non-crisp interrogative agendas do not induce a single categorization, but rather, a probability distribution over a parametrized *set* of possible categorization systems. However, for many applications, it might be desirable to obtain a *single* categorization system associated with this probability distribution. The simplest way to define such categorization system is to choose the concept lattice with the highest preference or probability value attached to it. However, this stipulation ignores a large amount of information of interest in other alternative categorizations. Another possible solution is to estimate the importance attributed to each feature by a given non-crisp interrogative agenda, using methods such as plausibility transformation [12], pignistic transformation [36,61], or decision probability transformation based on belief intervals [21].[5] The values thus obtained, representing the relative importance of individual features, can then be used as weights in computing the proximity or dissimilarity between different objects, based on the features shared and not shared between them. The dissimilarity or proximity data obtained in this way can be used to categorize objects based on any clustering technique [33,32,42]. However, the categorizations obtained with this method are flat (clusterings) rather than hierarchical (concept lattices). In the present section, we propose a novel *stability-based method* for associating a categorization system with any non-crisp agenda.

Throughout the present section, we fix a formal context $\mathbb{P} = (A, X, I)$. Recall that, for any non-crisp agenda $m : \mathcal{P}(X) \to [0, 1]$, we let $m' : \mathcal{R} \to [0, 1]$ denote its associated probability mass function on $\mathcal{R} = \{\mathbb{P}^+_Y \mid Y \subseteq X\}$.

Definition 3. For any $G \subseteq A$ s.t. $G \in \mathbb{P}^+$, the *stability index* of G is defined as follows:

$$\rho_m(G) = \sum \{m'(\mathbb{P}^+_Y) \mid G \in \mathbb{P}^+_Y \text{ for some } Y \subseteq X\}.$$

[4] Interval scaling is one of the methods used commonly for conceptual scaling. For more, see [27].

[5] The pignistic and plausibility transformations of the non-crisp agendas discussed in the previous section are reported in Appendix A.2.

The value of $\rho_m(G)$ can be understood as the extent to which G is a meaningful category according to the non-crisp agenda m. Indeed, by construction (cf. Section 3), $G \in \mathbb{P}^+$ iff $G \in \mathbb{P}^+_Y$ for some $Y \subseteq X$. Hence, intuitively, the higher the weight assigned by m to the categorization systems to which G pertains, the greater the extent to which G is a meaningful category according to m.

For any $\beta \in [0,1]$ and m as above, a *β-categorization system* according to m is the complete $\bigcap$-semilattice (hence complete lattice) $\mathbb{L}(m,\beta)$ of $\mathbb{P}^+$ generated by the set

$$C(m,\beta) := \{G \in \mathbb{P}^+ \mid G \subseteq A \text{ and } \rho_m(G) \geq \beta\}.$$

That is, $\mathbb{L}(m,\beta)$ is the categorization system formed by taking all intersections of the sets of objects with stability index greater than or equal to β. The lattice $\mathbb{L}(m,\beta)$ can be understood as the categorization system which (approximately) represents the given (non-crisp) agenda m given a stability parameter β. Unlike the categorization system which is assigned the highest probability by m (if it exists), this categorization system incorporates information about other possible categorization systems as well. The parameter β is a 'stability threshold' for a concept to be relevant.

Let m_1, m_2, and m_3 be the agendas of agents j_1, j_2, and j_3 discussed in Section 4, and c be the coalition formed by these agents. The categorization systems associated with m_1, m_2, and m_3 using the stability-based method when $\beta = 0.5$ and $s = 5$ are shown in Figs. 5, 7, and 6 respectively, while those associated with the coalition agendas $\oplus c = \Diamond c$ and $\triangleright c$ are shown in Figs. 5 and 8, respectively. Note that the categorization systems (lattices) associated with m_1 and $\oplus c$ are identical.

Proposition 4. *For any non-crisp agenda m and all $\beta_1, \beta_2 \in [0,1]$, if $\beta_1 \leq \beta_2$, then $C(m,\beta_2) \subseteq C(m,\beta_1)$ and $\mathbb{L}(m,\beta_2) \subseteq \mathbb{L}(m,\beta_1)$.*

Proof. If $B \subseteq A$ s.t. $B \in C(m,\beta_2)$ $G \subseteq A$ is such that $\rho_m(G) \geq \beta_2 \geq \beta_1$. Thus, $B \in C(m,\beta_1)$. The second inclusion follows from $\mathbb{L}(m,\beta_i)$ being meet-generated by $C(m,\beta_i)$, for $1 \leq i \leq 2$. □

The proposition above matches with the intuition that lower values of the stability threshold β yield finer-grained categorization systems.

Remark 5. As discussed in Section 3, for any $G \subseteq A$, the set $\mathcal{V}_G := \{Y \subseteq X \mid G \in \mathbb{P}^+_Y\}$ is upward closed w.r.t. inclusion; however, $\mathcal{V}_G$ does not need to be closed under intersection. To see this, consider the formal context $\mathbb{P} = (A, X, I)$ with $A = \{a,b\}$, $X = \{x,y,z\}$, and $I = \{(a,x),(a,y),(a,z),(b,y)\}$ and the set $G = \{a\}$; then $G \in \mathbb{P}^+_Y$ for $Y = \{x,y\}$, and $G \in \mathbb{P}^+_Z$ for $Z = \{y,z\}$, while $G \notin \mathbb{P}^+_{Y\cap Z} = \mathbb{P}^+_{\{y\}}$. Thus, $\{Y \mid G \in \mathbb{P}^+_Y\}$ does not necessarily have a minimum element.

Definition 6. The *upward restricted order* is the partial order $\leq_\uparrow$ on non-crisp agendas defined as follows: $m_1 \leq_\uparrow m_2$ iff for any subset $\mathcal{V} \subseteq \mathcal{P}(X)$ which is upward closed w.r.t. inclusion,

$$\sum_{Y\in\mathcal{V}} m_1(Y) \leq \sum_{Y\in\mathcal{V}} m_2(Y).$$

Lemma 7. *If $m_1 \leq_\uparrow m_2$, then $\rho_{m_1}(G) \leq \rho_{m_2}(G)$ for any $G \in \mathbb{P}^+$.*

Proof. Let $\mathcal{V}_G := \{Y \subseteq X \mid G \in \mathbb{P}^+_Y\}$. Let m'_1 and m'_2 be the mass functions on $\mathcal{R}$ induced by m_1, and m_2. Then,

$$\rho_{m_1}(G) = \sum_{Y\in\mathcal{V}_G} m'_1(\mathbb{P}^+_Y) = \sum_{Y\in\mathcal{V}_G} m_1(Y) \leq \sum_{Y\in\mathcal{V}_G} m_2(Y) = \sum_{Y\in\mathcal{V}_G} m'_2(\mathbb{P}^+_Y) = \rho_{m_2}(G),$$

where the inequality above follows from the fact that $\mathcal{V}_G$ is upward closed, and $m_1 \leq_\uparrow m_2$. □

Proposition 8. *If $m_1 \leq_\uparrow m_2$, then $C(m_1,\beta) \subseteq C(m_2,\beta)$ and $\mathbb{L}(m_1,\beta) \subseteq \mathbb{L}(m_2,\beta)$ for any $\beta \in [0,1]$.*

Proof. If $G \in C(m_1,\beta)$, then, by the assumption and Lemma 7, $\beta \leq \rho_{m_1}(G) \leq \rho_{m_2}(G)$, hence $G \in C(m_2,\beta)$. The second inclusion follows from the fact that $\mathbb{L}(m_i,\beta)$, being meet-generated by $C(m_i,\beta)$ for $1 \leq i \leq 2$. □

Therefore, if $m_1 \leq_\uparrow m_2$, then, for any fixed stability parameter $\beta \in [0,1]$, the categorization system induced by m_1 using the stability-based method is coarser than the one induced m_2. Hence, a $\leq_\uparrow$-smaller agenda considers less information relevant for categorization.

6. Orderings of mass functions and coalition agendas

Several orderings on Dempster-Shafer mass functions have been introduced in the literature, based e.g. on plausibility and quality functions, among others. In the present section, we compare the ordering $\leq_\uparrow$ defined in the previous section with these orderings.

Definition 9 *([22, Section 2.3]).* For any $m_1, m_2 : \mathcal{P}(X) \to [0,1]$,

1. *pl-ordering*: $m_1 \leq_{pl} m_2$ iff $pl_{m_1}(Y) \leq pl_{m_2}(Y)$ for every $Y \subseteq X$.
2. *q-ordering*: $m_1 \leq_q m_2$ iff $q_{m_1}(Y) \leq q_{m_2}(Y)$ for every $Y \subseteq X$.
3. *s-ordering*: $m_1 \leq_s m_2$ iff some square matrix $S : \mathcal{P}(X) \times \mathcal{P}(X) \to [0,1]$ exists such that:
   (a) $\sum_{W \subseteq X} S(W,Y) = 1$ for every $Y \subseteq X$;
   (b) for all $W, Y \subseteq X$, if $S(W,Y) > 0$ then $W \subseteq Y$;
   (c) $m_1(W) = \sum_{Y \subseteq X} S(W,Y) m_2(Y)$ for any $W \subseteq X$.
   In this case, m_1 is said to be a *specialization* of m_2, since the value that m_1 assigns to a given set W is the sum of all the individual fractions $S(W,Y)$ of the values assigned by m_2 to every superset Y of W. The matrix S is called a *specialization matrix*.
4. *Dempsterian specialization ordering*: $m_1 \leq_d m_2$ iff $m_1 = m \cap_m m_2$ for some Dempster-Shafer mass function m, where $m_1 \cap m_2$ denotes the un-normalized Dempster's combination given by

$$m_1 \cap_m m_2(Y) = \sum_{Y_1 \cap Y_2 = Y} m_1(Y_1) m_2(Y_2). \tag{6.1}$$

It is well known that [22, Section 2.3]

$$m_1 \leq_d m_2 \implies m_1 \leq_s m_2 \implies \begin{cases} m_1 \leq_{pl} m_2 \\ m_1 \leq_q m_2. \end{cases} \tag{6.2}$$

Proposition 10. *For all mass functions $m_1, m_2 : \mathcal{P}(X) \to [0,1]$*

$$m_1 \leq_d m_2 \implies m_1 \leq_s m_2 \implies m_1 \leq_\uparrow m_2 \implies \begin{cases} m_1 \leq_{pl} m_2 \\ m_1 \leq_q m_2. \end{cases} \tag{6.3}$$

Proof. We only need to prove the implications involving $\leq_\uparrow$. Let us assume that $m_1 \leq_s m_2$, i.e. a square matrix S exists satisfying conditions (a)-(c) of Definition 9.3. Let $\mathcal{V} \subseteq \mathcal{P}(X)$ be upward closed. Then

$$\begin{array}{rcll} \sum_{W \in \mathcal{V}} m_1(W) & = & \sum_{W \in \mathcal{V}} \sum_{W \subseteq Y} S(W,Y) m_2(Y) & \text{(c) and (b)} \\ & = & \sum_{Y \in \mathcal{V}} \sum_{W \subseteq Y, W \in \mathcal{V}} S(W,Y) m_2(Y) & \mathcal{V} \text{ upward closed} \\ & \leq & \sum_{Y \in \mathcal{V}} m_2(Y), & \text{(a)} \end{array}$$

which shows that $m_1 \leq_\uparrow m_2$. The inclusions $\leq_\uparrow \subseteq \leq_q$ and $\leq_\uparrow \subseteq \leq_{pl}$ follow immediately from the fact that the sets $\{Z \subseteq X \mid Y \subseteq Z\}$ and $\{Z \subseteq X \mid Y \cap Z \neq \emptyset\}$ are upward closed w.r.t. inclusion. □

Remark 11. The following examples show that the inclusions $\leq_\uparrow \subseteq \leq_q$ and $\leq_\uparrow \subseteq \leq_{pl}$ are strict. Let $X = \{y_1, y_2, y_3\}$, and m_1 and m_2 be the following mass functions on X:

$$m_1(\{y_1,y_3\}) = 0.3, \quad m_1(\{y_2,y_3\}) = 0.3, \quad m_1(\{y_1,y_2,y_3\}) = 0.2, \quad m_1(\{y_3\}) = 0.2 \quad \text{and}$$

$$m_2(\{y_1,y_3\}) = 0.1, \quad m_2(\{y_2,y_3\}) = 0.1, \quad m_2(\{y_1,y_2,y_3\}) = 0.5, \quad m_2(\{y_3\}) = 0.3.$$

The following table reports the values of q_{m_1} and q_{m_2} for all $Y \subseteq X$.

	$\emptyset$	$\{y_1\}$	$\{y_2\}$	$\{y_3\}$	$\{y_1,y_2\}$	$\{y_2,y_3\}$	$\{y_1,y_3\}$	$\{y_1,y_2,y_3\}$
q_{m_1}	1	0.5	0.5	1	0.2	0.5	0.5	0.2
q_{m_2}	1	0.6	0.6	1	0.5	0.6	0.6	0.5

Therefore, $m_1 \leq_q m_2$. However, for the upward closed set $\mathcal{V} = \{\{y_1,y_3\},\{y_2,y_3\},\{y_1,y_2,y_3\}\} \subseteq \mathcal{P}(X)$, we have

$$\sum_{Y \in \mathcal{V}} m_2(Y) < \sum_{Y \in \mathcal{V}} m_1(Y).$$

Let m_3 and m_4 be the following mass functions on X:

$$m_3(\{y_1,y_2\}) = 0.3, \quad m_3(\{y_2,y_3\}) = 0.4, \quad m_3(\{y_1,y_3\}) = 0.3, \quad \text{and}$$

$$m_4(\{y_1\}) = 0.1, \quad m_4(\{y_2\}) = 0.2, \quad m_4(\{y_3\}) = 0.2, \quad m_4(\{y_1,y_2\}) = 0.5.$$

The following table reports the values of pl_{m_3} and pl_{m_4} for all $Y \subseteq X$.

	$\emptyset$	$\{y_1\}$	$\{y_2\}$	$\{y_3\}$	$\{y_1,y_2\}$	$\{y_2,y_3\}$	$\{y_1,y_3\}$	$\{y_1,y_2,y_3\}$
pl_{m_3}	0	0.6	0.7	0.7	1	1	1	1
pl_{m_4}	0	0.6	0.7	0.2	0.8	0.9	0.8	1

Therefore, $m_4 \leq_{pl} m_3$. However, for the upward closed set $\mathcal{V} = \{\{y_1, y_2\}, \{y_1, y_2, y_3\}\} \subseteq \mathcal{P}(X)$, we have

$$\sum_{Y \in \mathcal{V}} m_3(Y) < \sum_{Y \in \mathcal{V}} m_4(Y).$$

We leave it as an open question to check if the converse of implication between $\leq_s$ and $\leq_\uparrow$ holds.

As an immediate consequence of Lemma 7 and Propositions 8 and 10, we get the following

Corollary 12. *If $m_1 \leq_s m_2$ or $m_1 \leq_d m_2$, then $\rho_{m_1}(G) \leq \rho_{m_2}(G)$ for any $G \in \mathbb{P}^+$, and $C(m_1, \beta) \subseteq C(m_2, \beta)$ and $\mathbb{L}(m_1, \beta) \subseteq \mathbb{L}(m_2, \beta)$ for any $\beta \in [0, 1]$.*

Lemma 13. *Let C be a set of agents. For any coalition $c \subseteq C$ and any agent $j \in c$,*

$$\Diamond c \leq_s m_j \quad \text{and} \quad m_j \leq_\uparrow \rhd c.$$

Proof. Let us show the first inequality by induction on the cardinality of c. The base case in which c is a singleton is trivial. For the inductive step, note that if $c' = c \cup \{i\}$, then $\Diamond(c') = \Diamond c \cap_m m_i$. Thus, it is enough to show that for any two mass functions m_1 and m_2, $m_1 \cap_m m_2 \leq_s m_1$. The proof follows by setting

$$S(W, Y) = \sum \{m_2(Y') \mid Y' \cap Y = W\}.$$

It is straightforward to check that S satisfies all the conditions required in Definition 9.

Let us show the second inequality by induction on the cardinality of c. The base case in which c is a singleton is trivial.

For any mass functions m_1 and m_2, let $m_1 \cup_m m_2$ be the mass defined as follows: for any $Y \subseteq X$, $m_1 \cup_m m_2(Y) = \sum_{Z_1 \cup Z_2 = Y} m_1(Z_1) m_2(Z_2)$. For the inductive step, if $c' = c \cup \{i\}$, then $\rhd(c') = (\rhd c) \cup_m m_i$. Thus, it is enough to show that $m_1 \leq_s m_1 \cup_m m_2$ for all mass functions m_1 and m_2.

For any $Y \subseteq X$, the mass $m_1(Y)$ is transferred completely to the sets larger than or equal to Y in performing the operation $\cup$. Thus, any mass assigned by m_1 to any set in an up-set $\mathcal{V}$ remains in $\mathcal{V}$ in $m_1 \cup_m m_2$. In particular, for every $Y \in \mathcal{V}$ and every $Y' \subseteq X$, $Y \cup Y' \in \mathcal{V}$. Therefore,

$$\sum_{Y \in \mathcal{V}} m_1(Y) = \sum_{Y \in \mathcal{V}} (\sum_{Y' \subseteq X} m_2(Y')) m_1(Y) \leq \sum_{Y \in \mathcal{V}} \sum_{Y_1 \cup Y_2 = Y} m_1(Y_1) m_2(Y_2) = \sum_{Y \in \mathcal{V}} (m_1 \cup_m m_2)(Y). \quad \square$$

The next corollary follows immediately from Lemma 13 and Proposition 8.

Corollary 14. *Let C be a set of agents. For any coalition $c \subseteq C$, any agent $j \in c$, and any $\beta \in [0, 1]$,*

$$C(\Diamond c, \beta) \subseteq C(m_j, \beta) \subseteq C(\rhd c, \beta) \quad \text{and} \quad \mathbb{L}(\Diamond c, \beta) \subseteq \mathbb{L}(m_j, \beta) \subseteq \mathbb{L}(\rhd c, \beta).$$

Hence, for any coalition c and any parameter β, the categorization system generated by the agendas $\Diamond c$ (resp. $\rhd c$) using the stability-based method with parameter β is coarser (resp. finer) than the categorization system generated by the agenda m_j of any agent $j \in c$. This can further justify understanding the non-crisp agenda $\Diamond c$ (resp. $\rhd c$) as the non-crisp counterpart of $\Diamond c$ (resp. $\rhd c$) as defined in (3.1).

7. Meta-learning algorithm using interrogative agendas

The extant outlier detection and classification algorithms based on FCA [26,54,38,63,70] are typically explainable, being based on data-analytic methods, and determine an appropriate set of relevant features using some feature-selection methods. However, unlike the methods discussed in the previous sections, most feature-selection methods do not allow for different (sets of) features to be assigned different relevance. In [1], a meta-algorithm is introduced which learns the (non-crisp) agenda which is best suited for given outlier detection or classification tasks. Given some outlier detection or classification algorithm in input, which acts on a single 2-valued concept lattice, the meta-algorithm applies it to all the concept lattices induced by the different agendas under consideration. The output of the meta-algorithm is obtained by combining the predictions associated with each of these concept lattices by means of some aggregation function (e.g. weighted average, maximum, minimum) which weights the different lattices according to their relevance values. The best values for these weights relative to a given task can then be learned from training data using gradient descent. The following is an edited version of the algorithm in [1], in which the specific aggregation function used in [1] is replaced by a generic one.

As the cardinality of $\mathcal{R}$ is $2^{|X|}$, only mass functions with a limited number of focal elements are considered, which are determined by the objectives and constraints of the algorithm. For simplicity, the weights in the learning phase are allowed to take values in $\mathbb{R}$. Interrogative agendas (mass functions) can be obtained by normalizing the absolute values of these weights.

Algorithm 1 Meta-Learning Algorithm for Interrogative Agendas.

Input: - A formal context $\mathbb{P} = (A, X, I)$,
- a training set $T \subseteq A$, a map $f : T \to Lab$ assigning labels to the elements of the training set,
- a set $\mathcal{Z} \subseteq \mathcal{P}(X)$ consisting of n different agendas under consideration,
- an algorithm Alg which takes an object $a \in A$ and a concept lattice in $\mathcal{R} = \{\mathbb{P}^+_Y \mid Y \subseteq X\}$ in input, and outputs an element in $\mathbb{R}^{Lab}$ representing, for each $l \in Lab$, the prediction of Alg on the likelihood that l be assigned to a;
- a loss function $loss : (\mathbb{R}^{Lab})^{|T|} \times (\mathbb{R}^{Lab})^{|T|} \to \mathbb{R}$,
- an aggregation function $Agg : (\mathbb{R}^{Lab})^n \times \mathbb{R}^n \to \mathbb{R}^{Lab}$ which combines the outputs of all agendas given some weights $\overline{w} \in \mathbb{R}^n$,
- a number of training epochs $epochs$.

Output: A model that classifies objects in A.

procedure TRAIN($\mathbb{P}$, T, f, Alg, Agg, $loss$, $epochs$)
 $\mathbb{L}_1, \ldots, \mathbb{L}_n \leftarrow$ **compute** the concept lattices induced by the agendas in $\mathcal{Z}$
 let $predictions$ be an empty map from A to $\mathbb{R}^C$
 let $\overline{w}$ be an array of random weights of length n
 for $e = 1, \ldots, epochs$ **do**
  for $a \in A$, $k \in C$ **do**
   $scores \leftarrow \left(Alg_k(a, \mathbb{L}_1), Alg_k(a, \mathbb{L}_2), \cdots, Alg_k(a, \mathbb{L}_n)\right)$
   $predictions[a][k] \leftarrow Agg(scores, \overline{w})$
  end for
  update $\overline{w}$ with an iteration of gradient descent (use $loss$)
 end for
end procedure

The weights learned by the algorithm describe the importance of different sets of features relative to a given task. These weights can be understood as a generalization of Shapely values [45], which estimate the importance of individual features in given machine learning tasks.

As agendas are represented as Dempster-Shafer mass functions, tools from Dempster-Shafer theory are available for furthering their theory and applications. For example, agendas learned by different algorithms can be combined by means of different aggregation rules from Dempster-Shafer theory, so to obtain an ensemble algorithm; or agendas can be compared or aggregated with the agendas of human experts, to allow for human-machine collaboration. This is especially important in fields such as auditing, where human supervision of algorithms is necessary for ethical and legal reasons.

The weights learned by the meta-algorithm arguably provide *global explainability*, since they encode the relevance of different sets of features in the classification decision of the algorithm. Furthermore, the meta-algorithm is *locally explainable*, since its output for any object can be described and tracked back in terms of the individual outputs obtained by running the algorithm Alg on the concept lattices induced by each agenda under consideration with the given input.

An implementation of this meta-algorithm for outlier detection, its comparison with other outlier detection algorithms, and the explanations provided by it are discussed in detail in [9]. In the future, we intend to carry out a similar study for the classification task.

8. Conclusion and further directions

Main contributions. This paper introduces a *multi-agent framework* for representing and reasoning about the space of categorization systems over a fixed set of entities, grounded in the preferences and epistemic stances of individual agents. This framework combines formal tools from FCA (namely formal contexts and their associated concept lattices) which are used to represent categorization systems, and tools from Dempster-Shafer theory (namely mass functions) which serve to encode the agents' epistemic attitudes (non-crisp interrogative agendas).

We illustrate this framework with a *case study* focusing on the problem of categorizing business processes for auditing purposes. We introduce a *stability-based method* for associating a single categorization system (represented as a complete lattice) with a Dempster-Shafer mass function on the set of features of a formal context; we define the *upward restricted order* on mass functions (cf. Definition 6), with respect to which the mapping of categorization systems to mass functions defined by the stability-based method is order preserving (cf. Proposition 8), and compare this new order with alternative mass function orderings defined in the literature. Finally, we introduce a *meta-learning algorithm* for selecting an appropriate agenda for given classification and outlier detection tasks; this meta-learning algorithm can provide global and local explanations of the outcomes of the algorithms to which it is applied.

The generality of these contributions makes them applicable to classification problems far beyond financial transactions. Below, we discuss some directions for future research.

Implementing the meta-learning algorithm. In [9], the meta-algorithm introduced in Section 7 is implemented by using a specific outlier detection algorithm Alg as its input. In the future, we intend to study several such implementations, using various FCA-based outlier detection and classification algorithms such as those in [26,54,38,63,70] and compare their performances.

Extensions to other tasks and fine-tuning. Another interesting direction for future research is to extend the meta-Algorithm 1 to *other tasks* where formal concept analysis has been successfully applied, such as data mining, information retrieval, attribute exploration, knowledge management [53,55,50,28,65].

Specific instantiations of the meta-learning algorithm can often be fine-tuned to more efficient procedures. For example, the outlier detection algorithm proposed in [9] only computes the formal concepts generated by objects, rather than the full concept lattices. This and other similar fine-tunings can mitigate the relatively high computational costs of these algorithms, while still adopting the fundamental idea laid out in the meta-algorithm. In the future, it would be interesting to explore such fine-tunings for the efficient implementation of this meta-algorithm in different tasks.

Further applications of the theoretical framework. The theoretical framework developed in this paper opens up further possibilities for analysis of agenda-driven categorizations and algorithms based on them. For example, it is possible to use it to *compare the weights* [35] learned by the meta-algorithm with different algorithms in input, and hence to understand whether different algorithms assign similar importance to a given set of features while performing the same task.

Moreover, in the present formal framework agendas obtained from different algorithms can be *aggregated* using Dempster-Shafer based methods [6,58], and likewise, agendas of *human experts* can be combined or compared with the agendas learned by the meta-algorithm.

Extending framework to model deliberation and dynamics of categories. In [18], a logico-algebraic framework is proposed for modelling deliberation among agents with different interrogative agendas, aimed at arriving at an aggregated agenda. This framework can be integrated with the agenda-driven categorization model introduced in this paper to capture how agents deliberate to establish a shared categorization system. Such an integrated framework offers a powerful tool for understanding the formation and evolution of social categories through collective reasoning and negotiation.

CRediT authorship contribution statement

Marcel Boersma: Writing – review & editing, Validation, Supervision. **Krishna Manoorkar:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Alessandra Palmigiano:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Conceptualization. **Mattia Panettiere:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Apostolos Tzimoulis:** Writing – review & editing, Formal analysis. **Nachoem Wijnberg:** Writing – review & editing, Supervision.

Disclaimer

The authors alone are responsible for the content and writing of the paper, and the views expressed here are their personal views and do not necessarily reflect the position of their employer.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Financial statements network example

In this section, we describe the (toy) financial statements network used to build the formal context discussed in Section 4, and the estimated importance values of different features according to the different agendas encountered throughout the paper obtained using the pignistic transformation and plausibility transformation.

A.1. Database and its associated formal context

Table 1 is a small database showing different transactions and financial accounts used to obtain a small financial statements network considered in the examples.

Table 2 shows the many valued context obtained by interpreting business processes as objects and financial accounts as features. The value of the incidence relation denotes the share of given financial account in a given business process.

A.2. Pignistic and plausibility transformations

Tables 3 and 4 report the estimated values of the importance of different features (financial accounts) calculated via pignistic and plausibility transformations of the non-crisp agendas (mass functions) discussed in Section 4.

Appendix B. Concept lattices associated with various interrogative agendas

The present appendix section collects the Hasse diagrams, drawn with the help of LatViz [40,2], of various concept lattices associated with the interrogative agendas in Section 4.

Table 1
A small database with 12 business processes and 6 financial accounts. We use same TID to denote all credit and debit activities relating to a single business process.

ID	TID	FA name	Value	ID	TID	FA name	Value
1	1	revenue	-100	17	7	cost of sales	+400
2	1	cost of sales	+100	18	7	revenue	-300
3	2	revenue	-400	19	7	inventory	-100
4	2	personal expenses	+400	20	8	revenue	-150
5	3	other expenses	+125	21	8	cost of sales	+150
6	3	cost of sales	+375	22	9	revenue	-250
7	3	revenue	-500	23	9	cost of sales	+250
8	4	tax	-125	24	10	tax	-250
9	4	cost of sales	+500	24	10	personal expenses	+250
10	4	revenue	-375	26	11	revenue	-250
11	5	tax	-10	27	11	personal expenses	+175
12	5	cost of sales	+200	28	11	other expenses	+75
13	5	revenue	-190	29	12	revenue	-250
14	6	other expenses	+50	30	12	tax	-50
15	6	cost of sales	+450	31	12	personal expenses	+150
16	6	inventory	-500	32	12	other expenses	+150

Table 2
The formal context obtained from transaction database given in Table 1. $I(a,x) = 0$ for any (a,x) pair not present in the table.

Business process (a)	Financial account (x)	Share of value (I(a,x))	Business process (a)	Financial account (x)	Share of value (I(a,x))
1 (a_1)	revenue (x_2)	-1	7 (a_7)	cost of sales (x_3)	+1
1 (a_1)	cost of sales (x_3)	+1	7 (a_7)	inventory (x_5)	-0.25
2 (a_2)	revenue (x_2)	-1	7 (a_7)	revenue (x_2)	-0.75
2 (a_2)	personal expenses (x_4)	+1	8 (a_8)	revenue (x_2)	-1
3 (a_3)	other expenses (x_6)	+0.25	8 (a_8)	cost of sales (x_3)	+1
3 (a_3)	cost of sales (x_3)	+0.75	9 (a_9)	revenue (x_2)	-1
3 (a_3)	revenue (x_2)	-1	9 (a_9)	other expenses (x_6)	+1
4 (a_4)	tax (x_1)	-0.25	10 (a_{10})	tax (x_1)	-1
4 (a_4)	cost of sales (x_3)	+1	10 (a_{10})	personal expenses (x_4)	+1
4 (a_4)	revenue (x_2)	-0.75	11 (a_{11})	revenue(x_2)	-1
5 (a_5)	tax (x_1)	-0.05	11 (a_{11})	personal expenses (x_4)	+0.7
5 (a_5)	cost of sales (x_3)	+1	11 (a_{11})	other expenses (x_6)	+0.3
5 (a_5)	revenue (x_2)	-0.95	12 (a_{12})	revenue(x_2)	-0.83
6 (a_6)	other expenses (x_6)	+0.1	12 (a_{12})	tax(x_1)	-0.17
6 (a_6)	cost of sales (x_3)	+0.9	12 (a_{12})	personal expenses (x_4)	+0.5
6 (a_6)	inventory (x_5)	-1	12 (a_{12})	other expenses (x_6)	+0.5

Table 3
Importance estimates via pignistic transformation.

Agenda	x_1	x_2	x_3	x_4	x_5	x_6
m_1	0.67	0.067	0.067	0.067	0.067	0.067
m_2	0.683	0.183	0.033	0.033	0.033	0.033
m_3	0.317	0.317	0.017	0.017	0.017	0.317
m	0.885	0.085	0.001	0.001	0.001	0.025
m'	0.239	0.239	0.095	0.095	0.095	0.239

Table 4
Importance estimates via plausibility transformation.

Agenda	x_1	x_2	x_3	x_4	x_5	x_6
m_1	0.333	0.133	0.133	0.133	0.133	0.133
m_2	0.435	0.217	0.087	0.087	0.087	0.087
m_3	0.303	0.303	0.030	0.030	0.030	0.303
m	0.767	0.153	0.006	0.006	0.006	0.061
m'	0.213	0.213	0.121	0.121	0.121	0.213

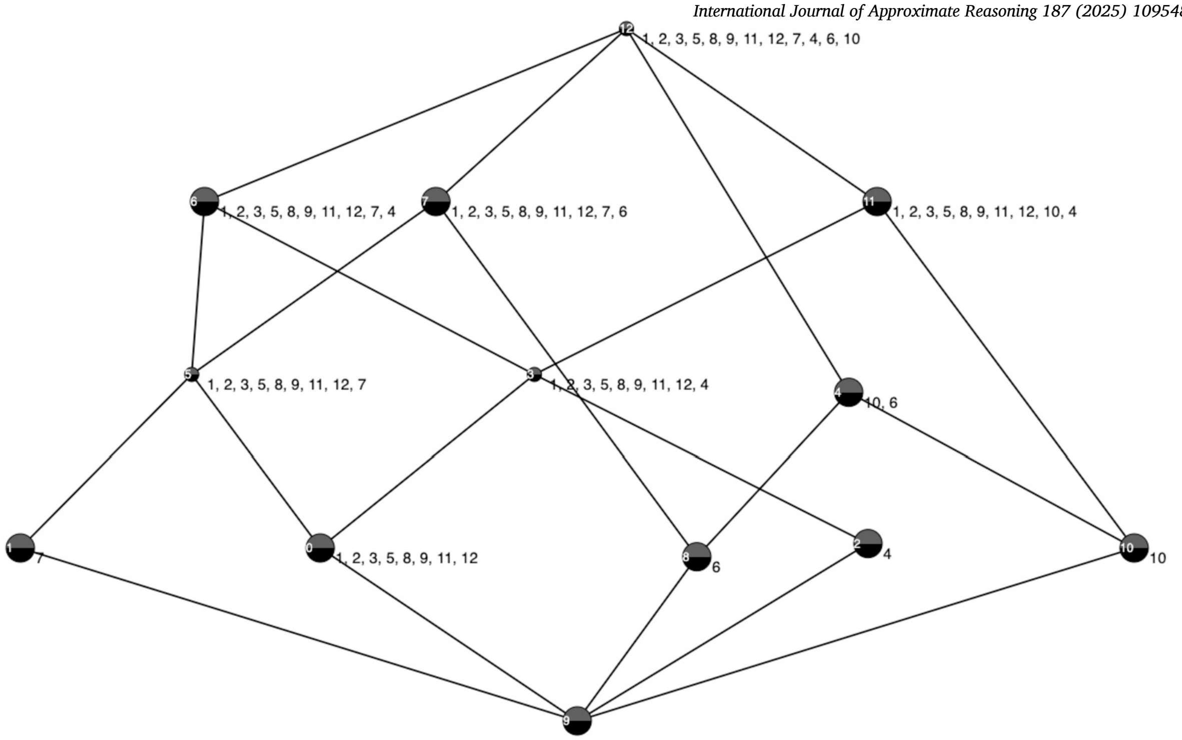

Fig. 1. Concept lattice associated with the crisp agenda $\{x_1, x_2, x_5\}$ of agent j_1.

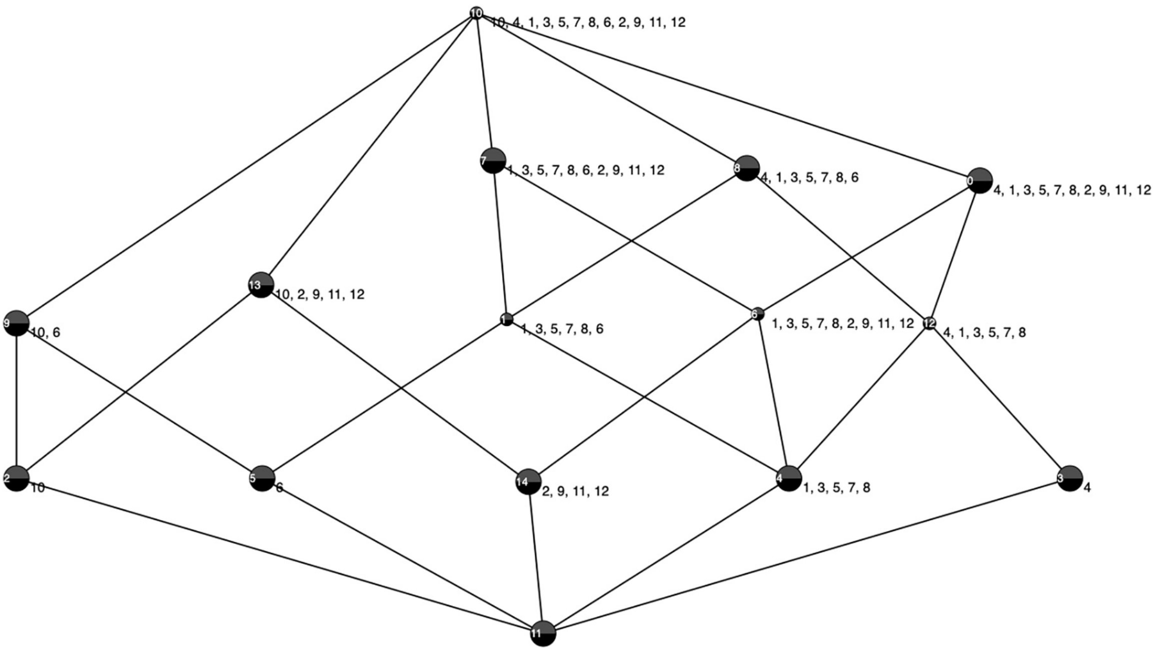

Fig. 2. Concept lattice associated with the crisp agenda $\{x_1, x_2, x_3\}$ of agent j_2.

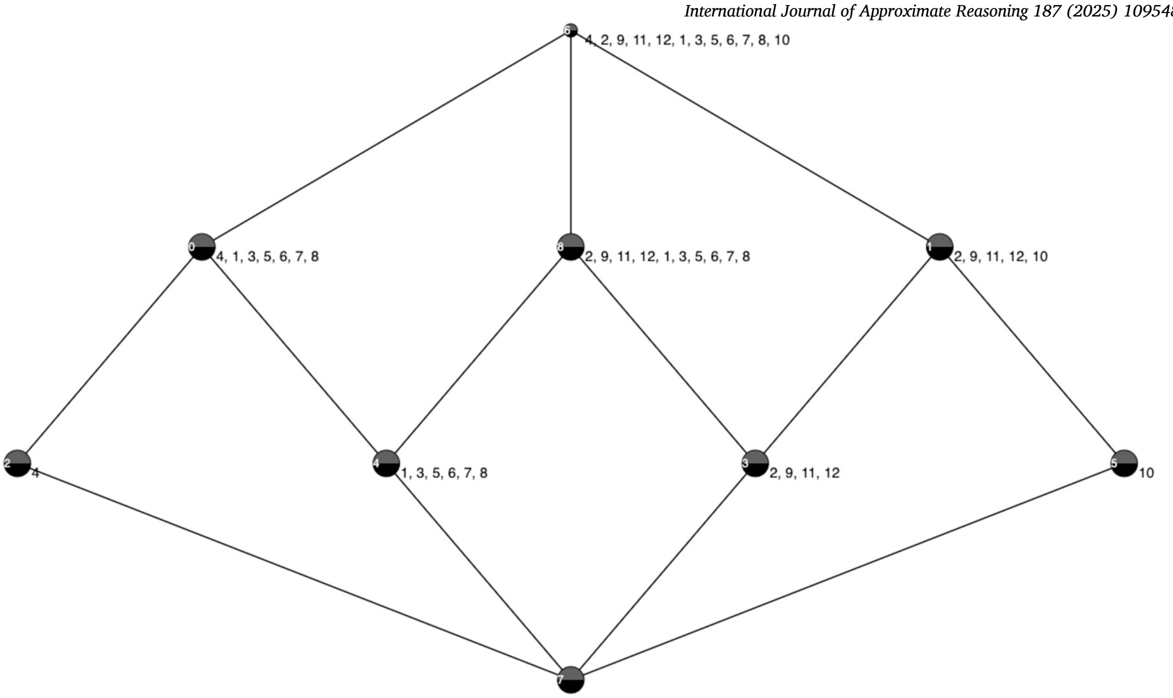

Fig. 3. Concept lattice associated with the crisp agenda $\{x_1, x_3\}$ of agent j_3.

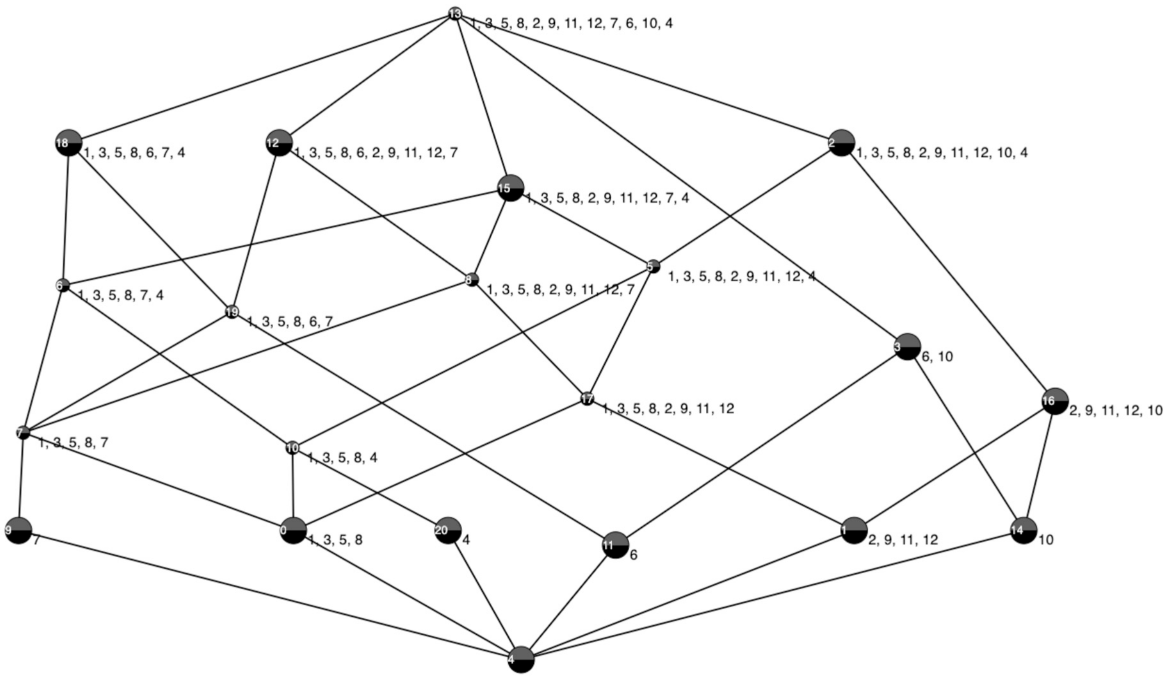

Fig. 4. Concept lattice associated with the crisp agenda $\rhd c = \{x_1, x_2, x_3, x_5\}$.

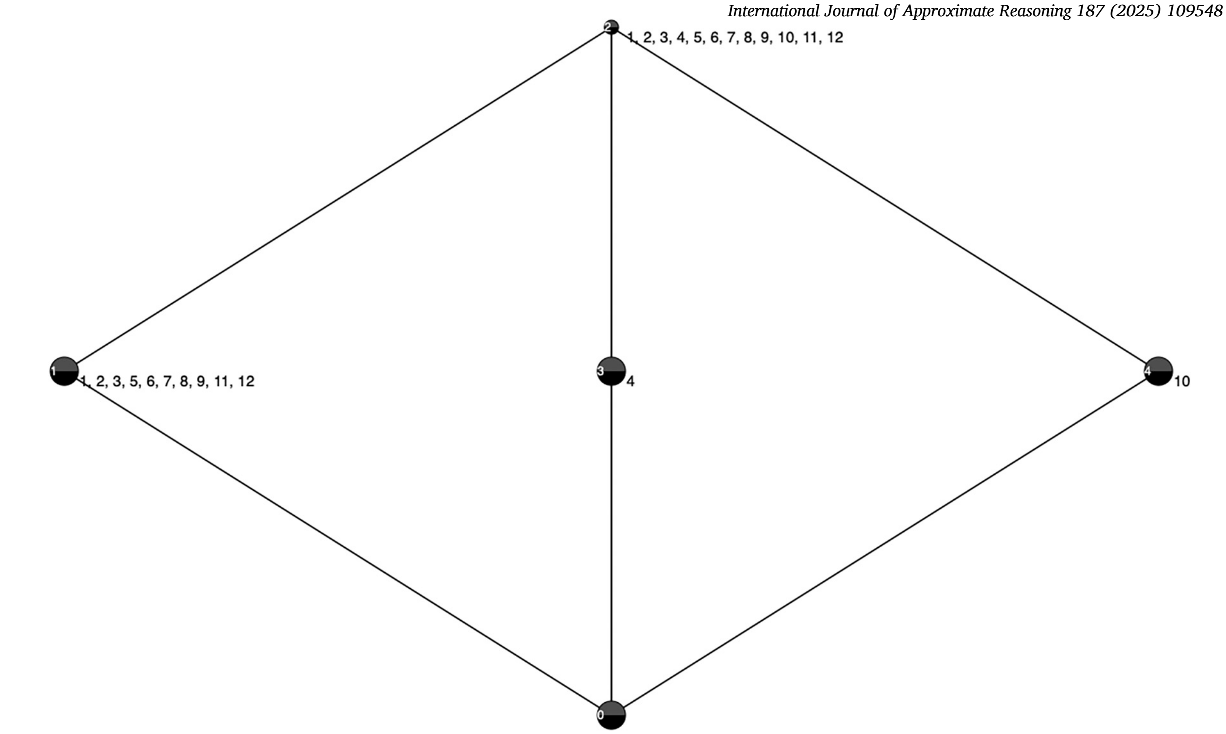

Fig. 5. This concept lattice is all of the following: (1) the concept lattice associated with the crisp agenda $\Diamond c = \{x_1\}$ in case, (2) most preferred categorization associated with the non-crisp agendas m_1 of j_1, m_2 of j_2, and $\oplus c$ (3) categorization obtained from non-crisp agendas m_1 and $\oplus c$ using stability-based method for $\beta = 0.5$.

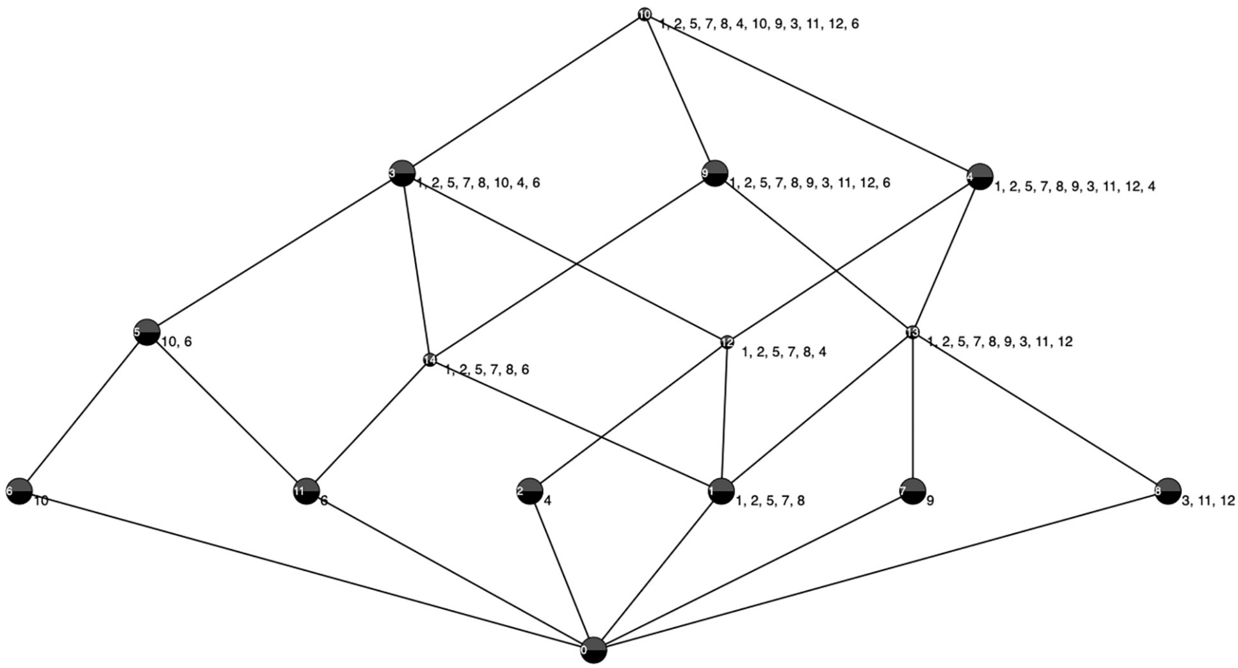

Fig. 6. This concept lattice is both (1) the most preferred categorization for the non-crisp agenda m_3 of j_3 and (2) the categorization obtained from non-crisp agenda m_3 using stability-based method for $\beta = 0.5$.

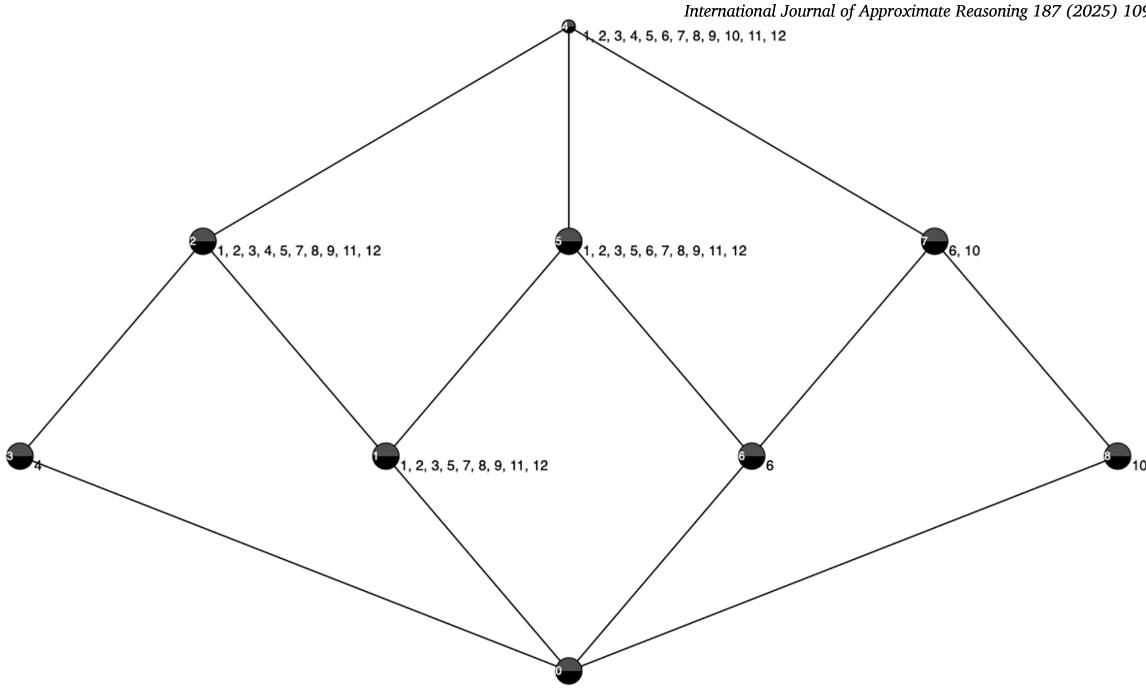

Fig. 7. Concept lattice representing the categorization system associated with non-crisp agenda m_2 using stability-based method for $\beta = 0.5$.

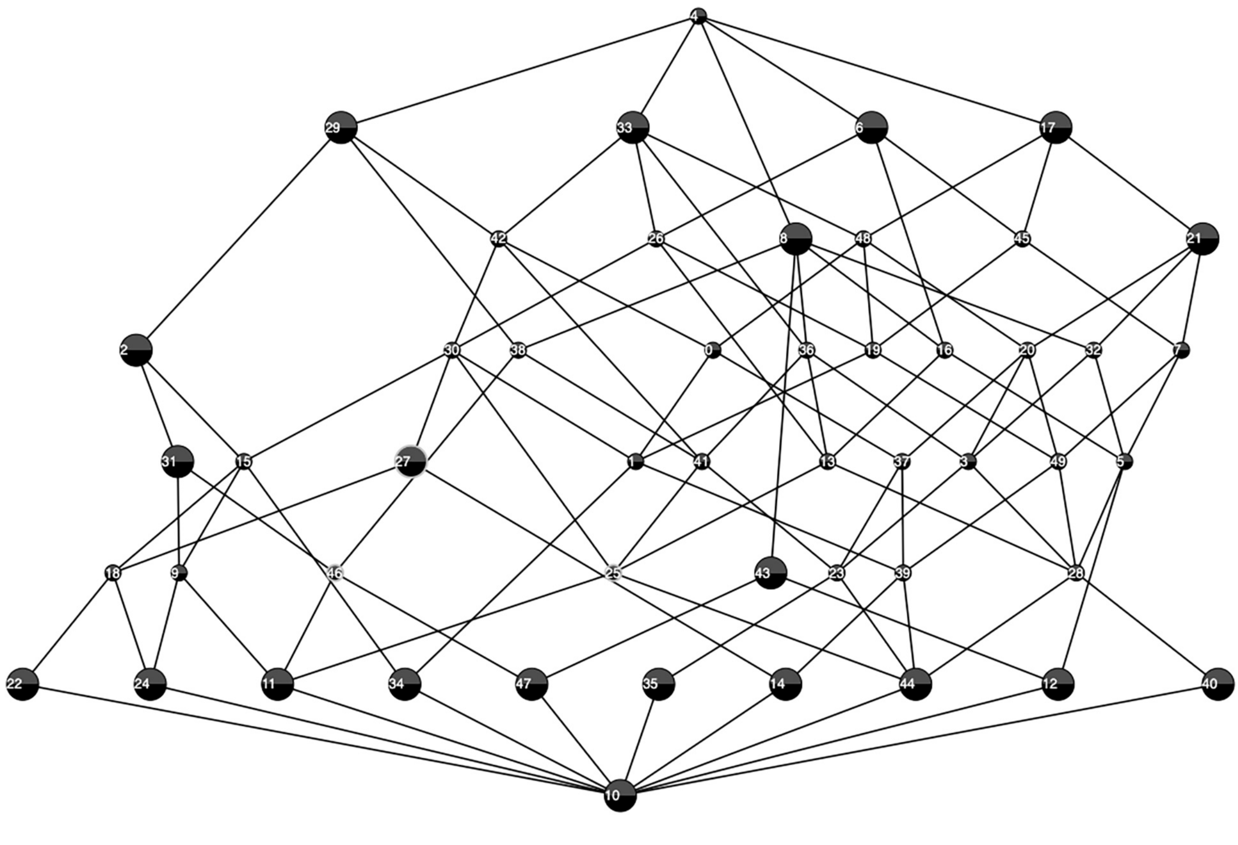

Fig. 8. The concept lattice representing both (1) the categorization obtained from non-crisp agenda $\rhd c$ using stability-based method for $\beta = 0.5$ (2) the concept lattice associated with the (crisp) agenda consisting of all the features in X.

Data availability

No data was used for the research described in the article.

References

[1] Erman Acar, Andrea De Domenico, Krishna Manoorkar, Mattia Panettiere, A meta-learning algorithm for interrogative agendas, arXiv preprint, arXiv:2301.01837, 2023.
[2] Mehwish Alam, Thi Nhu Nguyen Le, Amedeo Napoli, Latviz: a new practical tool for performing interactive exploration over concept lattices, in: CLA 2016-Thirteenth International Conference on Concept Lattices and Their Applications, 2016.
[3] Muhannad Alomari, Fangjun Li, David C. Hogg, Anthony G. Cohn, Online perceptual learning and natural language acquisition for autonomous robots, Artif. Intell. 303 (2022) 103637.
[4] Lawrence W. Barsalou, Ad hoc categories, Mem. Cogn. 11 (1983) 211–227.
[5] Radim Bêlohlávek, Fuzzy Galois connections, Math. Log. Q. 45 (4) (1999) 497–504.
[6] Yaxin Bi, Jiwen Guan, David Bell, The combination of multiple classifiers using an evidential reasoning approach, Artif. Intell. 172 (15) (2008) 1731–1751.
[7] Garrett Birkhoff, Lattice Theory, vol. 25, American Mathematical Soc., 1940.
[8] M. Boersma, A. Maliutin, S. Sourabh, L.A. Hoogduin, D. Kandhai, Reducing the complexity of financial networks using network embeddings, Sci. Rep. 10 (1) (2020) 1–15.
[9] Marcel Boersma, Krishna Manoorkar, Alessandra Palmigiano, Mattia Panettiere, Apostolos Tzimoulis, Nachoem Wijnberg, Outlier detection using flexible categorization and interrogative agendas, Decis. Support Syst. 180 (2024) 114196.
[10] Marcel Boersma, Sumit Sourabh, Lucas Hoogduin, Financial statement networks: an application of network theory in audit, J. Netw. Theory Finance (2018).
[11] Philipp Cimiano, Andreas Hotho, Steffen Staab, Learning concept hierarchies from text corpora using formal concept analysis, J. Artif. Intell. Res. 24 (2005) 305–339.
[12] Barry R. Cobb, Prakash P. Shenoy, On the plausibility transformation method for translating belief function models to probability models, Int. J. Approx. Reason. 41 (3) (2006) 314–330.
[13] George W. Cobb, Yung-Pin Chen, An application of Markov chain Monte Carlo to community ecology, Am. Math. Mon. 110 (4) (2003) 265–288.
[14] Joel B. Cohen, Kunal Basu, Alternative models of categorization: toward a contingent processing framework, J. Consum. Res. 13 (4) (1987) 455–472.
[15] Willem Conradie, Sabine Frittella, Krishna Manoorkar, Sajad Nazari, Alessandra Palmigiano, Apostolos Tzimoulis, Nachoem M. Wijnberg, Rough concepts, Inf. Sci. 561 (2021) 371–413.
[16] Willem Conradie, Sabine Frittella, A. Palmigiano, M. Piazzai, A. Tzimoulis, Nachoem M. Wijnberg, Toward an epistemic-logical theory of categorization, Electron. Proc. Theor. Comput. Sci. 251 (2017).
[17] Willem Conradie, Sabine Frittella, Alessandra Palmigiano, Michele Piazzai, Apostolos Tzimoulis, Nachoem M. Wijnberg, Categories: how I learned to stop worrying and love two sorts, in: International Workshop on Logic, Language, Information, and Computation, Springer, 2016, pp. 145–164.
[18] Willem Conradie, Krishna Manoorkar, Alessandra Palmigiano, Apostolos Tzimoulis, Nachoem Wijnberg, Questions as cognitive filters, arXiv preprint, arXiv:2506.22735, 2025.
[19] Willem Conradie, Alessandra Palmigiano, Claudette Robinson, Apostolos Tzimoulis, Nachoem M. Wijnberg, The logic of vague categories, arXiv preprint, arXiv:1908.04816, 2019.
[20] Manoranjan Dash, Huan Liu, Feature selection for classification, Intell. Data Anal. 1 (1–4) (1997) 131–156.
[21] Zhan Deng, Jianyu Wang, A novel decision probability transformation method based on belief interval, Knowl.-Based Syst. 208 (2020) 106427.
[22] Thierry Denœux, The cautious rule of combination for belief functions and some extensions, in: 2006 9th International Conference on Information Fusion, IEEE, 2006, pp. 1–8.
[23] Sebastian Enqvist, Modelling epistemic actions in interrogative belief revision, J. Log. Comput. 22 (6) (2012) 1335–1365.
[24] Linton C. Freeman, Douglas R. White, Using Galois lattices to represent network data, Sociol. Method. (1993) 127–146.
[25] Sabine Frittella, Krishna Manoorkar, Alessandra Palmigiano, Apostolos Tzimoulis, Nachoem Wijnberg, Toward a Dempster-Shafer theory of concepts, Int. J. Approx. Reason. 125 (2020) 14–25.
[26] Huaiyu Fu, Huaiguo Fu, Patrik Njiwoua, Engelbert Mephu Nguifo, A comparative study of fca-based supervised classification algorithms, in: International Conference on Formal Concept Analysis, Springer, 2004, pp. 313–320.
[27] Bernhard Ganter, Rudolf Wille, Conceptual scaling, in: Applications of Combinatorics and Graph Theory to the Biological and Social Sciences, Springer, 1989, pp. 139–167.
[28] Bernhard Ganter, Rudolf Wille, Formal Concept Analysis: Mathematical Foundations, Springer Science & Business Media, 2012.
[29] Robert L. Goldstone, The role of similarity in categorization: providing a groundwork, Cognition 52 (2) (1994) 125–157.
[30] Fei Hao, Geyong Min, Zheng Pei, Doo-Soon Park, Laurence T. Yang, k-clique community detection in social networks based on formal concept analysis, IEEE Syst. J. 11 (1) (2015) 250–259.
[31] Jonathan Hayes, Claudio Gutierrez, Bipartite graphs as intermediate model for rdf, in: International Semantic Web Conference, Springer, 2004, pp. 47–61.
[32] Anil K. Jain, Richard C. Dubes, Algorithms for Clustering Data, Prentice-Hall, Inc., 1988.
[33] Anil K. Jain, M. Narasimha Murty, Patrick J. Flynn, Data clustering: a review, ACM Comput. Surv. 31 (3) (1999) 264–323.
[34] Dominik Janzing, Lenon Minorics, Patrick Blöbaum, Feature relevance quantification in explainable ai: a causal problem, in: International Conference on Artificial Intelligence and Statistics, PMLR, 2020, pp. 2907–2916.
[35] Anne-Laure Jousselme, Dominic Grenier, Éloi Bossé, A new distance between two bodies of evidence, Inf. Fusion 2 (2) (2001) 91–101.
[36] Frank Klawonn, Philippe Smets, The dynamic of belief in the transferable belief model and specialization-generalization matrices, in: Uncertainty in Artificial Intelligence, Elsevier, 1992, pp. 130–137.
[37] Sergei O. Kuznetsov, Machine learning and formal concept analysis, in: International Conference on Formal Concept Analysis, Springer, 2004, pp. 287–312.
[38] Sergei O. Kuznetsov, Fitting pattern structures to knowledge discovery in big data, in: International Conference on Formal Concept Analysis, Springer, 2013, pp. 254–266.
[39] Sergei O. Kuznetsov, Jonas Poelmans, Knowledge representation and processing with formal concept analysis, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 3 (3) (2013) 200–215.
[40] LatViz, Lattice visualization, latviz.loria.fr, 2016.
[41] Stephan Lewandowsky, Mike Kalish, Thomas L. Griffiths, Competing strategies in categorization: expediency and resistance to knowledge restructuring, J. Exp. Psychol. Learn. Mem. Cogn. 26 (6) (2000) 1666.
[42] Feng Li, Shoumei Li, Thierry Denœux, k-cevclus: constrained evidential clustering of large dissimilarity data, Knowl.-Based Syst. 142 (2018) 29–44.
[43] Jundong Li, Kewei Cheng, Suhang Wang, Fred Morstatter, Robert P. Trevino, Jiliang Tang, Huan Liu, Feature selection: a data perspective, ACM Comput. Surv. 50 (6) (2017) 1–45.
[44] Pantelis Linardatos, Vasilis Papastefanopoulos, Sotiris Kotsiantis, Explainable ai: a review of machine learning interpretability methods, Entropy 23 (1) (2020) 18.

[45] Christoph Molnar, Interpretable Machine Learning, Lulu. com, 2020.
[46] Joel Nothman, Nicky Ringland, Will Radford, Tara Murphy, James R. Curran, Learning multilingual named entity recognition from Wikipedia, Artif. Intell. 194 (2013) 151–175.
[47] David W. Opitz, Feature selection for ensembles, AAAI/IAAI 379 (384) (1999) 3.
[48] Georgios A. Pavlopoulos, Panagiota I. Kontou, Athanasia Pavlopoulou, Costas Bouyioukos, Evripides Markou, Pantelis G. Bagos, Bipartite graphs in systems biology and medicine: a survey of methods and applications, GigaScience 7 (4) (2018) giy014.
[49] Gonçalo Pereira, Rui Prada, Pedro A. Santos, Integrating social power into the decision-making of cognitive agents, Artif. Intell. 241 (2016) 1–44.
[50] Jonas Poelmans, Sergei O. Kuznetsov, Dmitry I. Ignatov, Guido Dedene, Formal concept analysis in knowledge processing: a survey on models and techniques, Expert Syst. Appl. 40 (16) (2013) 6601–6623.
[51] Susanne Prediger, Logical scaling in formal concept analysis, in: International Conference on Conceptual Structures, Springer, 1997, pp. 332–341.
[52] Susanne Prediger, Rudolf Wille, The lattice of concept graphs of a relationally scaled context, in: International Conference on Conceptual Structures, Springer, 1999, pp. 401–414.
[53] Uta Priss, Formal concept analysis in information science, Annu. Rev. Inf. Sci. Technol. 40 (1) (2006) 521–543.
[54] Olga Prokasheva, Alina Onishchenko, Sergey Gurov, Classification methods based on formal concept analysis, in: FCAIR 2012–Formal Concept Analysis Meets Information Retrieval, 2013, p. 95.
[55] Abderrahim El Qadi, Driss Aboutajedine, Yassine Ennouary, Formal concept analysis for information retrieval, arXiv preprint, arXiv:1003.1494, 2010.
[56] Erzsébet Ravasz, Anna Lisa Somera, Dale A. Mongru, Zoltán N. Oltvai, A-L. Barabási, Hierarchical organization of modularity in metabolic networks, Science 297 (5586) (2002) 1551–1555.
[57] Deb Roy, Semiotic schemas: a framework for grounding language in action and perception, Artif. Intell. 167 (1) (2005) 170–205, Connecting Language to the World.
[58] Kari Sentz, Scott Ferson, et al., Combination of Evidence in Dempster-Shafer Theory, vol. 4015, Citeseer, 2002.
[59] Glenn Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.
[60] Frano Škopljanac-Mačina, Bruno Blašković, Formal concept analysis–overview and applications, Proc. Eng. 69 (2014) 1258–1267.
[61] Philippe Smets, Decision making in the tbm: the necessity of the pignistic transformation, Int. J. Approx. Reason. 38 (2) (2005) 133–147.
[62] Edward E. Smith, Andrea L. Patalano, John Jonides, Alternative strategies of categorization, Cognition 65 (2–3) (1998) 167–196.
[63] Mahito Sugiyama, Akihiro Yamamoto, Semi-supervised learning on closed set lattices, Intell. Data Anal. 17 (3) (2013) 399–421.
[64] Xiaochuan Tang, Yuanshun Dai, Yanping Xiang, Feature selection based on feature interactions with application to text categorization, Expert Syst. Appl. 120 (2019) 207–216.
[65] Rudolf Wille, Bernhard Ganter, Formal Concept Analysis, 1996.
[66] Qiang Wu, Zongtian Liu, Real formal concept analysis based on grey-rough set theory, Knowl.-Based Syst. 22 (1) (2009) 38–45.
[67] Yu Xue, Yihang Tang, Xin Xu, Jiayu Liang, Ferrante Neri, Multi-objective feature selection with missing data in classification, IEEE Trans. Emerg. Top. Comput. Intell. 6 (2) (2021) 355–364.
[68] Ronald R. Yager, Liping Liu, Classic Works of the Dempster-Shafer Theory of Belief Functions, vol. 219, Springer, 2008.
[69] Ruiping Yin, Kan Li, Guangquan Zhang, Jie Lu, A deeper graph neural network for recommender systems, Knowl.-Based Syst. 185 (2019) 105020.
[70] Jifu Zhang, Sulan Zhang, Kai H. Chang, Xiao Qin, An outlier mining algorithm based on constrained concept lattice, Int. J. Syst. Sci. 45 (5) (2014) 1170–1179.