

Establishing Correspondences between Attribute Spaces and Complex Concept Spaces Using Meta-PGN Classifier

Kr. Ivanova, I. Mitov, P. Stanchev,
Ph. Ein-Dor, K. Vanhoof

Institute of Mathematics and Informatics, BAS
Kettering University, Flint, USA
The Academic College of Tel Aviv-Yafo, Israel
Hasselt University, Belgium

Abstract

- One approach for extending the learning set of a classification algorithm with additional metadata used as a base for giving appropriate names of found regularities
- The analysis of correspondence between connections established in the attribute space and existing links between concepts can be used as a test for creation of an adequate model of the observed world
- Meta-PGN classifier is suggested as a possible tool for establishing these connections

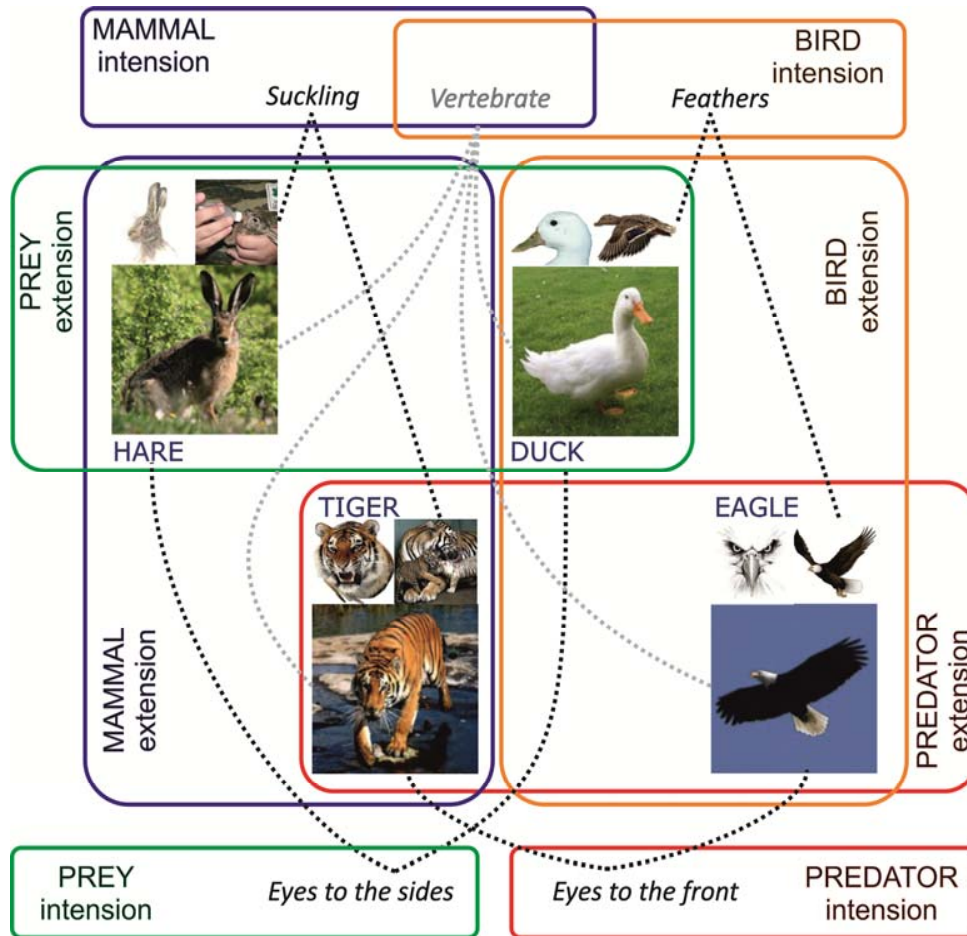
Concepts - communication



*When I use a word
it means just
what I choose it to mean -
neither more nor less*

Humpty Dumpty in Lewis Carroll's "Through the Looking-Glass"

Concept formation - Formal Concept Analysis



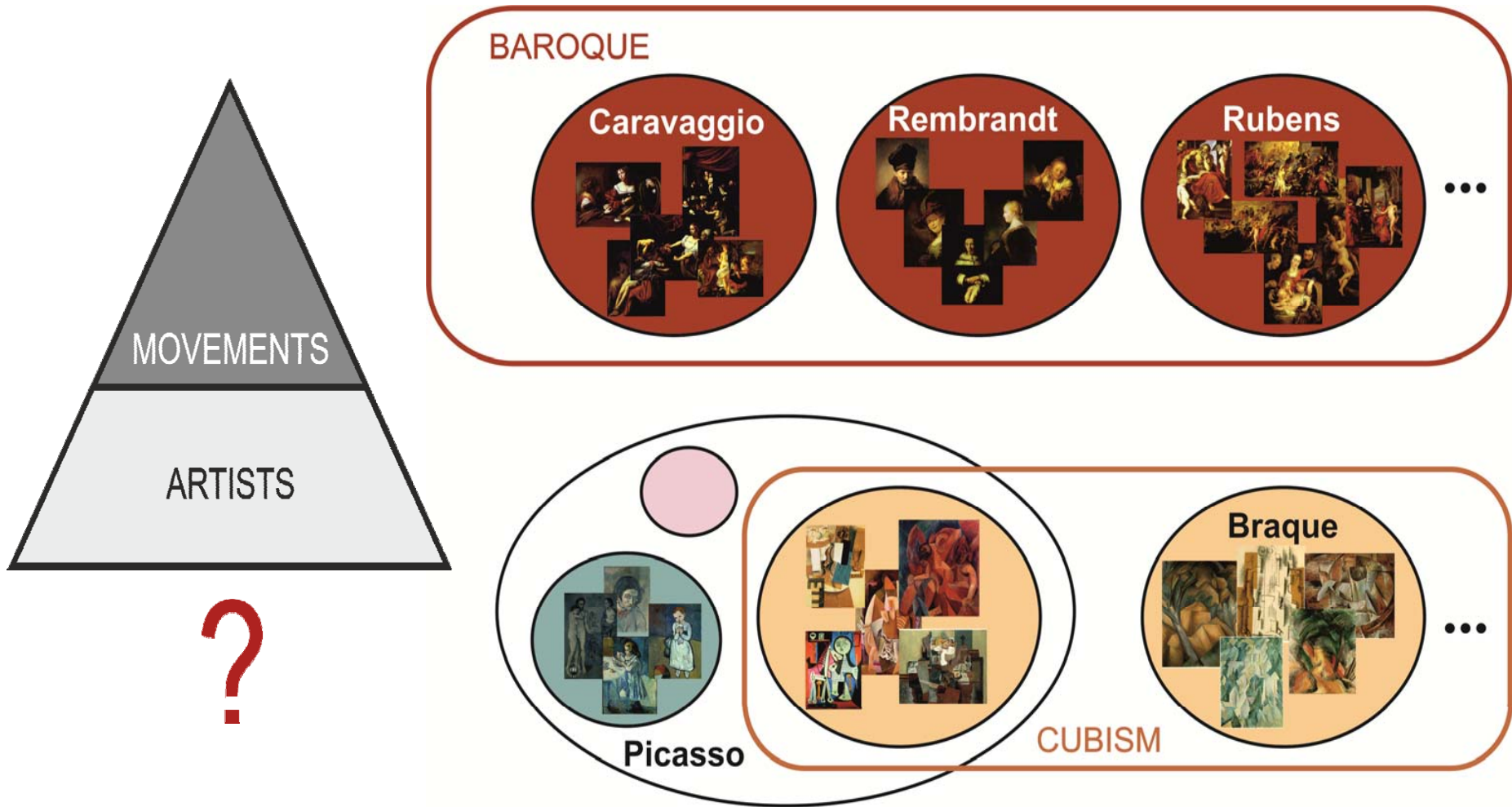
Concept:

- **extension** (covers all objects belonging to this concept)
- **intension** (comprises all attributes valid for all those objects)
- the possibility of using **attribute descriptions of the extension** of one concept to form its **intension definition**
- and to **uncover the relations** between different concepts.

"Understanding of the world" by the intelligent system

- We give the system some description of the real world and expect that it will follow our mental model and will choose appropriate names of concepts that arise.
- At the same time we expect that the system will explain the decision it has made and will show the parts of the model, which it has used to choosing the names of the concepts.
- This interaction with the intelligent system helps us to improve our mental model and to develop it by extending and/or changing any of its parts.
- As a result of this process of information interaction with an intelligent system one obtains the possibility of improving the attribute space of some observed area.

Algorithm description



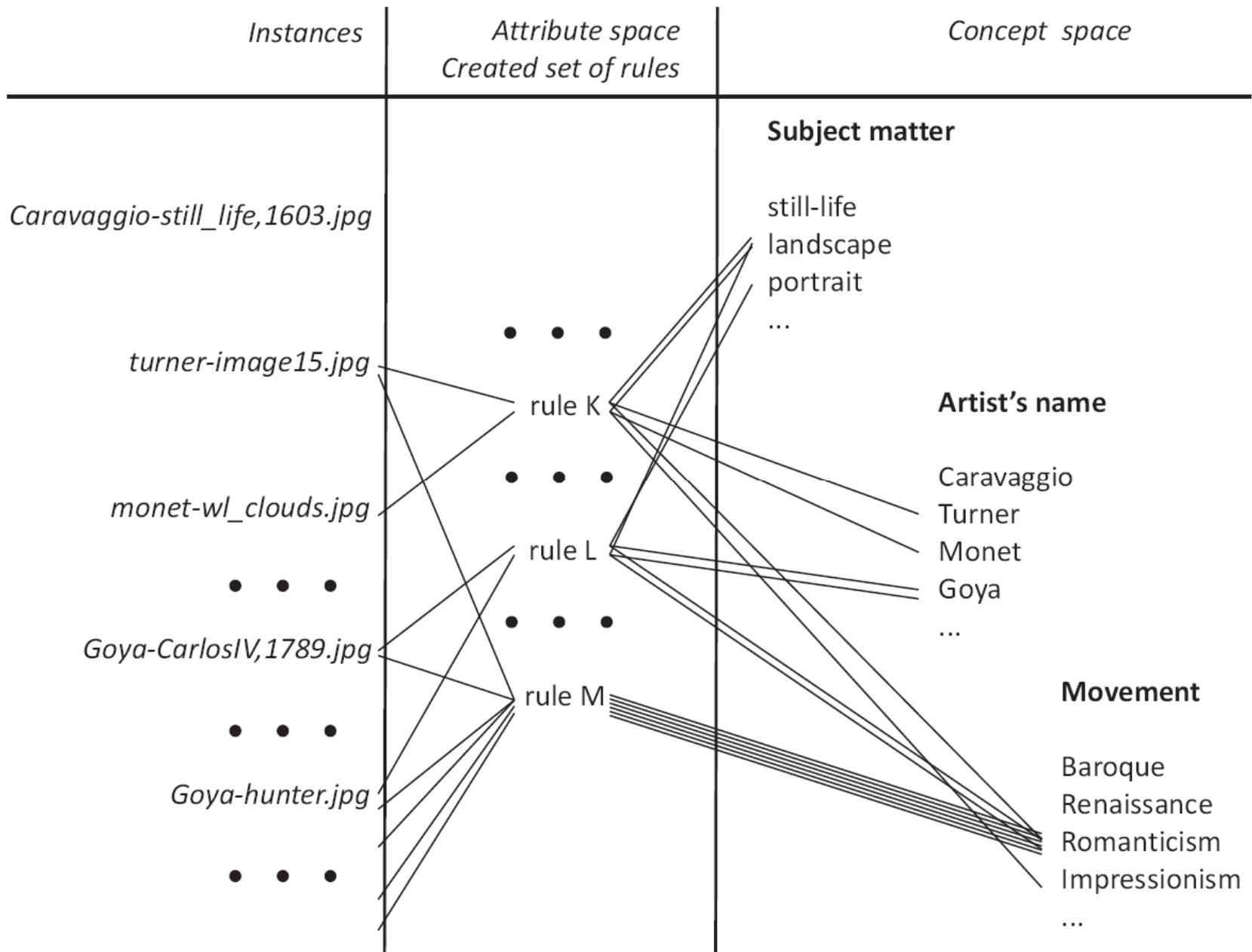
Dataset Construction

instance name	Descriptive part					Metadata part			
	A1	A2	A3	A4	A5	class= subject matter	class= artist' names	class= move- ments	
<i>Caravaggio-still_life,1603.jpg</i>	(v11)	(v12)	(v13)	(v14)	(v15)	still-life	Caravaggio	Baroque	
<i>turner-image15.jpg</i>	(v21)	(v22)	(v23)	(v24)	(v25)	landscape	Turner	Romanticism	
<i>monet-wl_clouds.jpg</i>	(v31)	(v32)	(v33)	(v34)	(v35)	landscape	Monnet	Impressionism	
		•	•	•			•	•	•
<i>Goya-CarlosIV,1789.jpg</i>	(vx1)	(vx2)	(vx3)	(vx4)	(vx5)	portrait	Goya	Romanticism	
		•	•	•			•	•	•
<i>Goya-hunter.jpg</i>	(vy1)	(vy2)	(vy3)	(vy4)	(vy5)	landscape	Goya	Romanticism	
		•	•	•			•	•	•

Identifying the Concepts

First step: generating the associative rules

- The learning set is processed by the standard classification algorithm PGN in the phase of generating the associative rules.
- As a result we receive a set of rules, which define specific frequently occurring combinations of attributes.
- Each rule is connected with the instances of the learning set, which were participated in its creation.



Identifying the Concepts

Second step: finding metadata values as potential names of the rules

- The rules, support for which is more than some threshold given as a parameter in each case, are examined.
- For each rule, the set of its instances is processed.
- The second parts of the vectors of these instances are analyzed.
- For each metadata position the normalized most frequently encountered value is determined.
- Each of these values is a potential candidate for the name of the concept, which is defined by the typical combination of attributes, contained in the examined rule.

Identifying the Concepts

Third step: defining the metadata values

- All metadata positions are traversed in order to find the rules that correspond to the value of this position.
- The rules, connected with the examined metadata value, are additionally processed in order to throw out the rules that are supersets of other rules in the group.

Identifying the Concepts

As a result, for each metadata value (that defines some concept) there are:

- **NO RULE** - The reason that corresponding rules do not exist usually lies in the fact that: (1) the threshold of this position was too high or (2) chosen primary attributes are not enough to correctly define this concept.
- **ONE RULE** – we can assume that this is the exact name of this rule. The content of this concept is represented as a conjunction of significant values of attributes, contained in the corresponding rule.

There is also a risk that primary attributes do not correctly represent the examined area. However, this is the problem of classification in general.

- **MORE RULES** - it is represented as a disjunction of these rules.

Finding Connections between Different Classes of Metadata

- **hierarchical dependencies**

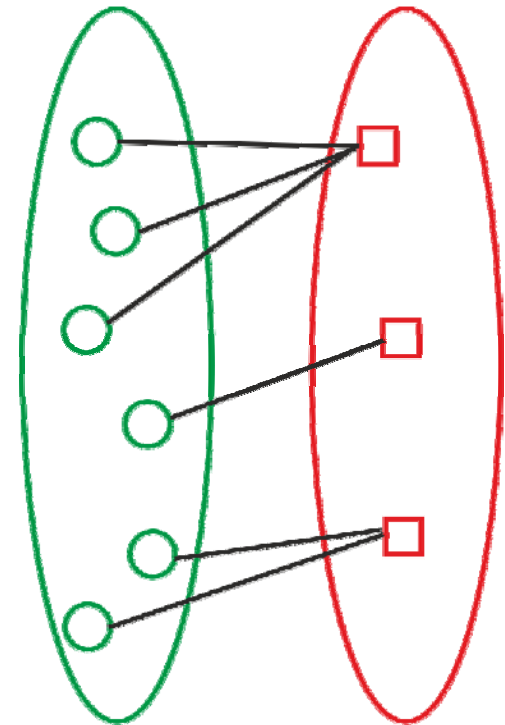
if

(1) each value in the first class corresponds to some value in the second class;

(2) different values in the first class correspond to one value in the second class;

(3) there are no different values in the second class that correspond to one value in the first class

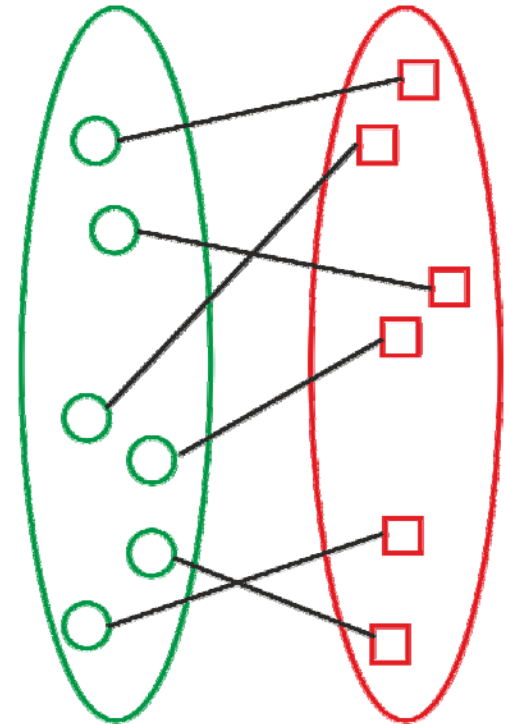
then the first class is hierarchically subordinated to the second class



Finding Connections between Different Classes of Metadata

- **equivalency between two classes**

when different values in the first class correspond to different values in the second class and vice versa.



!!! The analysis is focused not to establish the new connections between concepts, but to test the correctness of chosen attribute space and model of creating the rules in order to reproduce an **adequate model of the observed area**.

Conclusion

- Our attention in this position paper stopped on the boundary between the **processes of clustering** that **reveal the structure of the attribute space** and the **processes of categorization** that **build the bridges between the attribute space and concepts** on the way to **examine more complex concept spaces**.
- The analysis of correspondence between connections established in the attribute space and existing links between concepts can be used as a **test for creation of an adequate model of the observed world**.
- As an example of possible application, the attribute space created on the base of CBIR of paintings and the concept space of art collections is presented.

Thank you for your attention !

Krassimira Ivanova

`kivanova@math.bas.bg`