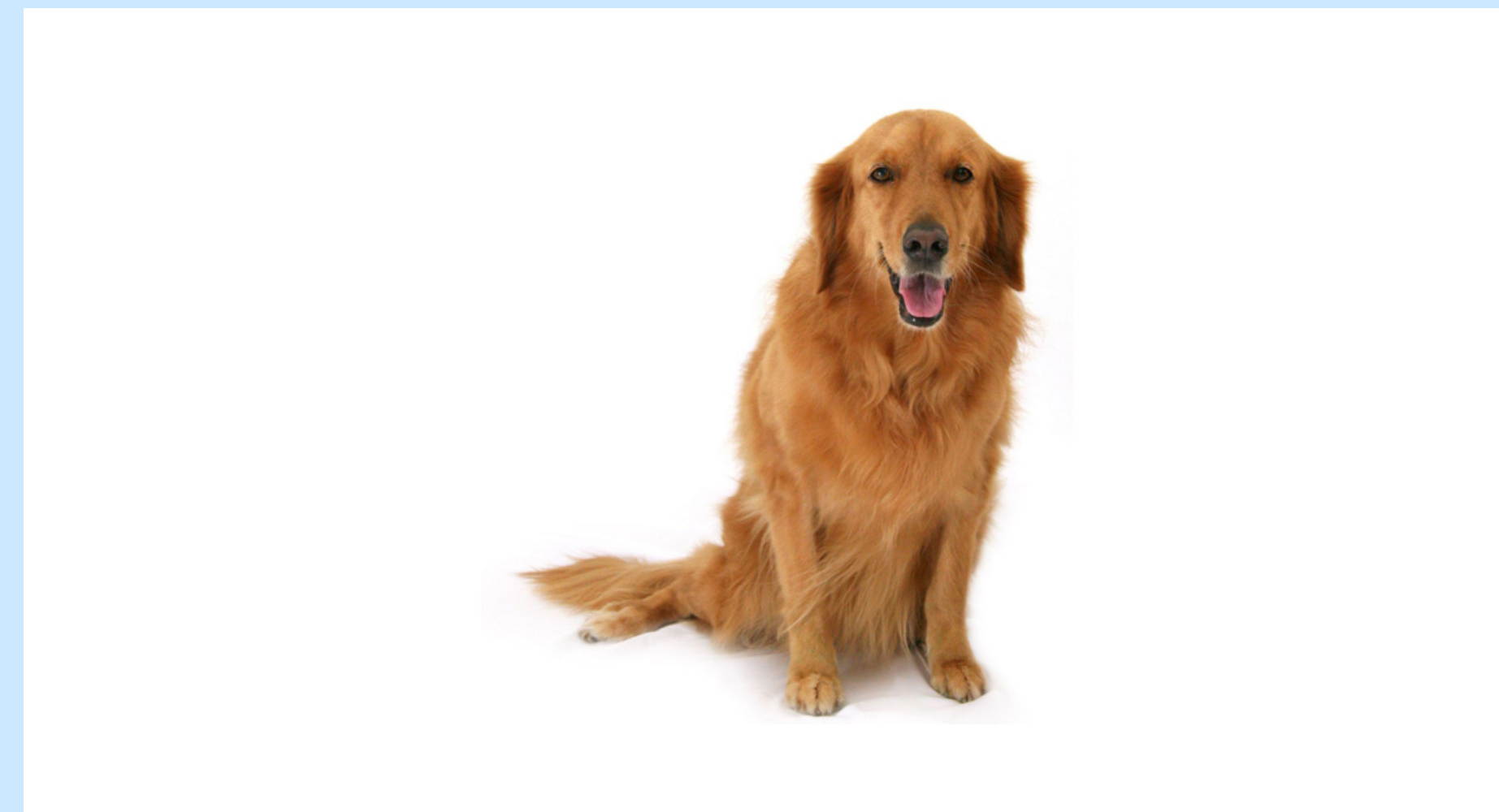


Introduction

It is a well-known fact that, usually, a concept lattice contains a considerable number of formal concepts. It is also known that, some of the concepts are more important (or natural) than others. This observation may be utilized for selecting only important concepts and filtering out the others. We propose an approach based on a phenomenon well-known in the psychology of concepts, namely the basic level of concepts.

What Do You See?



Why Dog?

Usually people answer a dog. But why dog? There is a number of other possibilities:

- :: Animal
- :: Mammal
- :: Canine beast
- :: Retriever
- :: Golden Retriever

So why dog?: Because "dog" is a basic level concept.

Basic Level Phenomenon

- :: When people categorize (or name) objects, they prefer to use certain kind of concepts.
- :: Such concepts are called the concepts of the basic level.
- :: Definition of basic level concepts: Are cognitively economic to use; "carve the world well".
- :: We use one of the first approaches, due to Eleanor Rosch (1970s):

Objects of the basic level concepts are similar to each other, objects of superordinate concepts are significantly less similar, while objects of the subordinate concepts are only slightly more similar to each other.

Our Approach to Basic Level

Formal concept $\langle A, B \rangle$ belongs to the basic level if it satisfies following properties:

- (BL1) $\langle A, B \rangle$ has a high cohesion.
- (BL2) $\langle A, B \rangle$ has a significantly larger cohesion than its upper neighbours.
- (BL3) $\langle A, B \rangle$ has a slightly smaller cohesion than its lower neighbours.

Cohesion of formal concept = measure of mutual similarity of objects.

Similarity

Similarity of objects x_1 and x_2 on $\langle X, Y, I \rangle$ can be viewed as similarity of their corresponding intents.

$$sim(x_1, x_2) = sim_Y(\{x_1\}^\uparrow, \{x_2\}^\uparrow). \quad (1)$$

$sim(x_1, x_2)$ denotes the degree (or index) of similarity of objects x_1 and x_2 .

For $B_1, B_2 \subseteq Y$ we use two measures of object similarity:

$$sim_{SMC}(B_1, B_2) = \frac{|B_1 \cap B_2| + |Y - (B_1 \cup B_2)|}{|Y|}, \quad (2)$$

$$sim_J(B_1, B_2) = \frac{|B_1 \cap B_2|}{|B_1 \cup B_2|}. \quad (3)$$

Cohesion

$coh(c)$ denotes the degree (or index) of cohesion of formal concept c .

For $\langle A, B \rangle \in \mathcal{B}(X, Y, I)$ we use two views on cohesion:

$$coh^\mathcal{A}(A, B) = \frac{\sum_{\{x_1, x_2\} \subseteq A, x_1 \neq x_2} sim(x_1, x_2)}{|A| \cdot (|A| - 1) / 2}. \quad (4)$$

$$coh^m(A, B) = \min_{x_1, x_2 \in A} sim(x_1, x_2), \quad (5)$$

Basic Level Degree

- :: We can compute for every formal concept $\langle A, B \rangle$ of $\langle X, Y, I \rangle$ degree $BL(A, B)$ to which $\langle A, B \rangle$ is a concept from the basic level.

- :: Concepts from the basic level need to satisfy conditions (BL1), (BL2), and (BL3).

$$BL(A, B) = \mathcal{C}(\alpha_1(A, B), \alpha_2(A, B), \alpha_3(A, B)), \quad (6)$$

where

- $\alpha_i(A, B)$ is the degree to which condition (BL $_i$) is satisfied, $i = 1, 2, 3$,
- \mathcal{C} is a "conjunctive" aggregation function

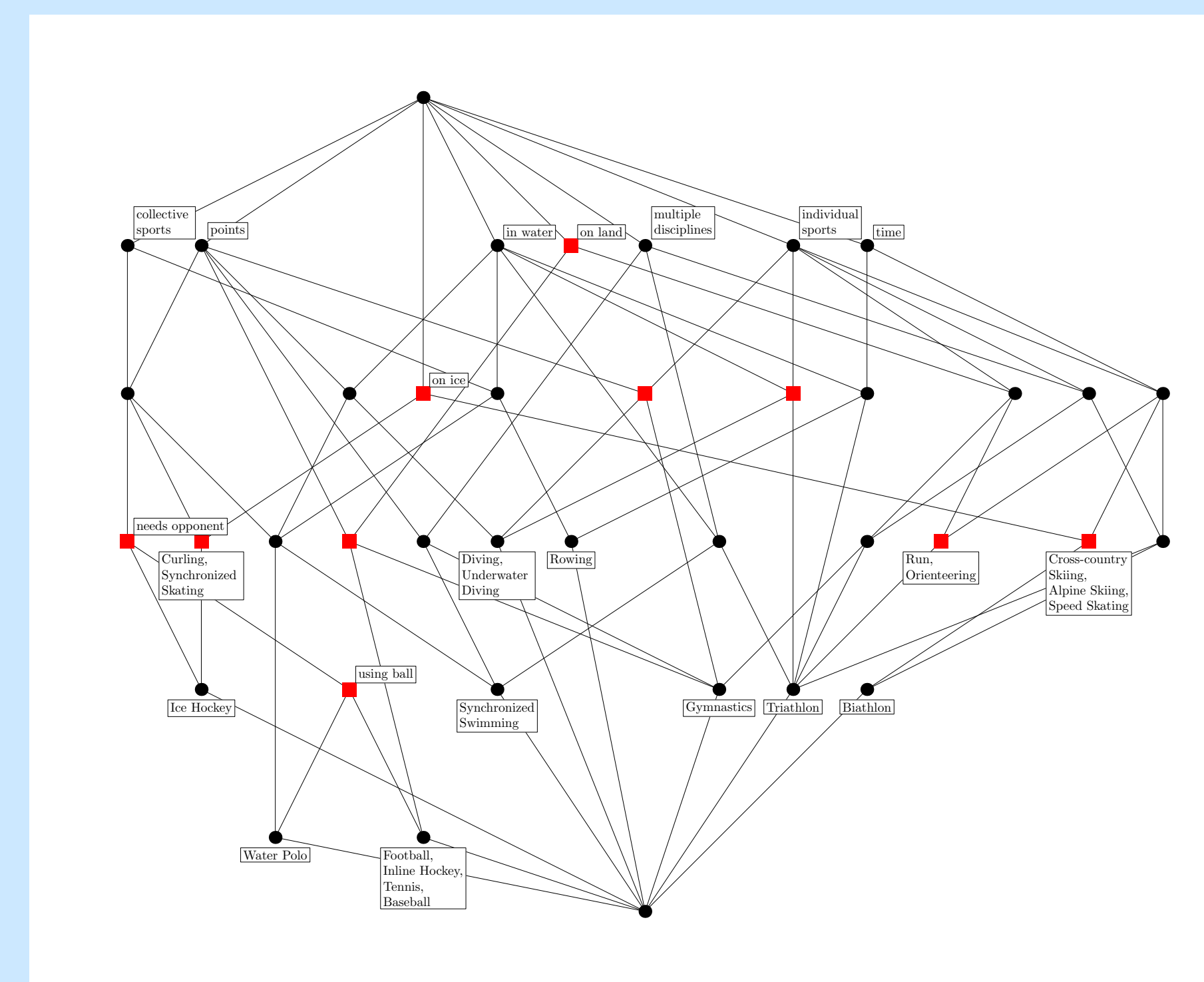
Experiments

For every dataset $\langle X, Y, I \rangle$ we compute the basic level degree of all concepts of the concept lattice $\mathcal{B}(X, Y, I)$.

Sports dataset

| | on land | on ice | in water | collective sport | individual sport | using ball | needs opponent | multiple disciplines | points | time |
|-----------------------|---------|--------|----------|------------------|------------------|------------|----------------|----------------------|--------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Run | 1 | x | | x | | | | | | x |
| Orienteering | 2 | x | | x | | | | | | x |
| Gymnastics | 3 | x | | x | | | | x | x | |
| Triathlon | 4 | x | x | x | | | | x | x | x |
| Football | 5 | x | | x | | x | x | x | | x |
| Inline Hockey | 6 | x | | x | | x | x | x | | x |
| Tennis | 7 | x | | x | | x | x | x | | x |
| Baseball | 8 | x | | x | | x | x | x | | x |
| Ice Hockey | 9 | x | x | | | | x | | | x |
| Curling | 10 | x | | x | | | | | | x |
| Cross-country Skiing | 11 | x | | x | | | | | | x |
| Synchronized Skating | 12 | x | | x | | | | | x | x |
| Alpine Skiing | 13 | x | | x | | | | | | x |
| Biathlon | 14 | x | | x | | | | x | | x |
| Speed Skating | 15 | x | | x | | | | | | x |
| Synchronized Swimming | 16 | x | x | | | | | x | x | |
| Diving | 17 | x | | x | | | | | | x |
| Water Polo | 18 | x | x | | | x | x | | | x |
| Underwater Diving | 19 | x | | x | | | | | | x |
| Rowing | 20 | x | x | | | | | | | x |

Concept Lattice of Sports



Basic level concepts are marked with red square.

Results

An important observation, basic level depends on the dataset and the selected attributes in particular. Typically, a human expert tends to take into account other information (not only the attributes present in the dataset).

Concepts Selected for Basic Level

- :: "ball games"
- :: "land sports"
- :: "individual sports"
- :: "winter collective sports"
- :: "individual water sports"

Arguably, all of them are likely to be considered natural, basic level concepts.

Concepts Not Selected for Basic Level

- :: "individual land sports with multiple disciplines"
- :: "sports performed in water with multiple disciplines"
- :: "individual winter sports with multiple disciplines evaluated by time"
- :: "collective winter sports with opponent evaluated by points"
- :: "sports evaluated by time"

Conclusion

- :: We proposed new method utilizing basic level of concepts to select possibility important concepts.
- :: Method seems to deliver reasonable basic level concepts.
- :: Simple, FCA-based formal framework for basic level study.

Feature Research

- :: Further experiments (drinks dataset, ...).
- :: Psychological experiments.
- :: Utilizing other approaches to the basic level developed in the psychology of concepts.
- :: Algorithmic aspects: Compute efficiently basic level concepts.
- :: Present the work to the community of the psychology of concepts.